Community and Interoperability at the Core of Sustaining Image Archives

Ulrike Felsing\textsuperscript{1,3}, Peter Fornaro\textsuperscript{2}, Max Frischknecht\textsuperscript{1,3} and Julien Antoine Raemy\textsuperscript{2,4}

\textsuperscript{1}HKB - Bern Academy of the Arts, Bern University of Applied Sciences, Fellerstrasse 11, 3027 Bern, Switzerland
\textsuperscript{2}Digital Humanities Lab, University of Basel, Spalenberg 65, 4051 Basel, Switzerland
\textsuperscript{3}Walter Benjamin Kolleg, University of Bern, Muesmattstrasse 45, 3012 Bern, Switzerland
\textsuperscript{4}DaSCH - Swiss National Data and Service Center for the Humanities, Gewerbestrasse 24, 4123 Allschwil, Switzerland

Abstract
In our paper, we discuss how the digital domain extends the sustainability of analogue archives through communication with the public. Our interdisciplinary research project "Participatory Knowledge Practices in Analogue and Digital Image Archives" (PIA) is funded by the Swiss National Science Foundation (2021–2025) and developed in cooperation with the photographic archives of the Swiss Society for Folklore Studies (SSFS). It aims to increase the use of image-based research data by developing participatory tools and deploying shared application programming interfaces (APIs) such as standards that adhere to the Linked Open Usable Data (LOUD) design principles. By involving the public, the project aims to increase the overall use of image-based research data. This makes data more sustainable in interaction with the analogue archive and increases the attractiveness of digital infrastructures.

Keywords
Citizen Science, Cultural Heritage, Digital Infrastructure, Interoperability, Linked Open Usable Data, Participatory Design, Sustainability

1. Introduction
We are engaged in an ongoing research project that explores open collaboration in the context of photographic archives. Our objective is to develop interactive tools and interfaces that foster the utilisation of image-based research data through active participation. By facilitating engagement from various communities, we aim to encourage the collective production of knowledge and democratise decision-making processes within archival settings. In our paper, we investigate how an interconnected and usable digital realm extends the sustainability of analogue archives through interaction and public involvement. We also emphasise the importance of standards in enabling data sharing and improving the resilience of cultural heritage data.
We first provide a short state of the art in terms of building communities of citizen scientists in section 2. It is followed by section 3 that emphasises on the research project Participatory Knowledge Practices in Analogue and Digital Image Archives (PIA). We then discuss Linked Open Usable Data (LOUD) in section 4 and their communities that develop and maintain shared application programming interfaces (APIs) for semantic interoperability purposes. In sections 5 and 6 we address the participatory design of the graphical user interface (GUI) and the overall technical architecture of PIA. Section 7 outlines suggestions for further research and finally in section 8, we conclude our analysis.

2. Sustaining Archives with Citizen Scientists

An increasing amount of research is being done in open collaboration with a crowd, with some of these projects being understood as Citizen Science which is characterised by openness in terms of participation and thus offers diverse perspectives for engagement within different fields of knowledge. Similar projects include Ajapaik [1] for crowdsourcing additional visual heritage metadata, Corley Explorer [2] for collecting stories, sMapshot [3] for georeferencing images, or Historypin [4] and notreHistoire.ch [5] for sharing local history.

According to Ridge, moderating contributions is one of the bigger challenges. For example, one participant reported feeling frustrated about moderation delays on her posts, which indicates that a lack of timely feedback on a completed task has a detrimental effect. In contrast,

“others […] felt that the changing source material they were working with helped them stay motivated.” [6, 131]

These and other issues of curating collaboration will always be the responsibility of the group that initiates the “Call for images”. On our platform, however, we must provide the adequate functionalities to do so. This includes, for example, the login and the possibility to report problematic content.

3. Participatory Knowledge Practices in Analogue and Digital Image Archives

Participatory Knowledge Practices in Analogue and Digital Image Archives (PIA)¹ is a four-year research project (2021-2025). The project is led by the University of Basel (Institute for Cultural Anthropology and European Ethnology as well as the Digital Humanities Lab) and the Bern Academy of the Arts (HKB).

Our research is based on three cultural heritage collections from the photographic archives of the Swiss Society for Folklore Studies (SSFS)²: one focusing on scientific cartography (Atlas of Swiss Folklore, published from 1950 until 1995), a second from the estate of the photojournalist Ernst Brunner (1901–1979), and a third collection consisting of vernacular photography which was owned by the Kreis Family (1860–1970).

¹https://about.participatory-archives.ch
²https://archiv.sgv-sstp.ch/
• The Atlas of Swiss Folklore was commissioned by the SSFS as part of a long-term scientific project. By mapping cultural patterns, the atlas constructs a comprehensive picture of Switzerland based on the links between culture and geography. The data was compiled by both academic professionals and non-academics, called “laymen” in the language of the time. The goal of this extensive survey was to document “Swiss folk culture” in the 1930s and 1940s through questions on a wide range of topics such as everyday behaviour, local laws, festivals, work and trade [7].

• Ernst Brunner consists of 48,000 negatives and 20,000 prints. Brunner, a professional photographer, published numerous artistic and documentary photographs on a wide variety of folkloric subjects, primarily in popular magazines e.g. “Swiss Home” and “Swiss Family” [8].

• Kreis Family is a private collection of a Basel-based family of physicians and printers that connects many strands of the cultural imagination. It is a typical example of urban bourgeois culture. The collection includes 20,000 loose photographic objects. A quarter of them were organised in 93 photo albums.

In a decisive interdisciplinary approach between digital humanities, cultural anthropology and design research, we work together to connect practical and theoretical issues, especially the user perspectives identified in the workshops and the theoretical findings. We are committed to collaborating with the scientific community and the wider public, facilitating the preservation and dissemination of knowledge, and encouraging users to engage together with their own histories and contemporary practices.

Overall, our main goal is to increase the use of the images and their metadata of the three collections. For this purpose, we develop digital tools that support contextualising, linking and contrasting images. By fostering exchange and collaboration in digital communities, we contribute to strengthening the photographic archives of the SSFS, and in order to achieve this we rely as much on the development of APIs maintained by cultural heritage practitioners as on the creation of a GUI following participatory design guidelines.

4. Semantic Interoperability through Linked Open Usable Data

Collaboratively designed standards build the base by which individuals and institutions can participate by having unmitigated access to the data. We consider, in particular, that Linked Open Usable Data (LOUD) specifications [9] are adequate not only in the cultural heritage field but more broadly for all participatory efforts within Citizen Science practices as it is an approach to serialise and expose data for different target groups.

“One of the first intentions of LOUD is to provide access to the data for both academic and software developers. An appropriate balance must be found between the requirements of data completeness and accuracy, driven by the ontological construct, and the pragmatic concerns of scalability and usability”. [10]

Similar to Tim-Berners Lee’s 5-star deployment scheme for Open Data³, five design principles

³https://5stardata.info/
underpin LOUD⁴:

- **The right Abstraction for the audience**: use cases rather than ontological purity should be favoured to determine the level of interoperability.
- **Few Barriers to entry**: the data, and the underlying model, must be easy to leverage. Having such systems in place will encourage more people to actively use them.
- **Comprehensible by introspection**: the data must be largely understandable simply by looking at it, without requiring external help.
- **Documentation with working examples**: comprehensive documentation should be produced to clarify the implementation of the use cases.
- **Few Exceptions, instead many consistent patterns**: patterns should be able to accommodate as few exceptions as possible to avoid adding rules that require the creation of tailor-made leeway fields on a case-by-case basis.

These design principles were inspired by those of the International Image Interoperability Framework (IIIF)⁵, a community-driven initiative made up largely of academic and memory institutions. Since 2012, IIIF has developed and maintained shared APIs for representing and annotating digital resources [11] and it has fundamentally shaped how libraries, archives and museums disseminate digital surrogates and digital-born objects over the past few years.

If we consider the relationship between the FAIR data principles (*Findable, Accessible, Interoperable, Reusable*) [12] and LOUD, it appears that the former refers to the environment in which the data is situated, while the latter pertains to the content itself. Breaking down the LOUD

---

⁴https://linked.art/loud/
⁵https://iiif.io/api/annex/notes/design_principles/

**Figure 1**: Example of a LOUD ecosystem which illustrates how the IIIF APIs (here the Image, Presentation and Change Discovery APIs), the Web Annotation Data Model and the Linked Art API can be integrated in the same environment. This diagram was inspired by the one on https://iiif.io/get-started/how-IIIF-works/.
acronym, the terms **Linked**, **Usable** and **(machine-readable) Data** can be seen as characteristics of the data (and its usability once transferred to an environment) [10]. **Open** can correspond somewhat to the principle of reusability outlined in FAIR [13].

Among the standards that adhere to the LOUD design principles are of course those conceived by the IIIF community, especially the IIIF Presentation API 3.0 which has been made with the latest updates of their design principles and for easier integration with JSON-LD 1.1 and other Web standards [14] such the Web Annotation Data Model [15], as well as Linked Art, a Resource Description Framework (RDF) application profile of CIDOC-CRM — a high-level ontology to enable information integration for cultural heritage data [16] — for semantically conveying assertions in a event-based paradigm [17]. All of these specifications were made and created collaboratively. While the Web Annotation Data Model is the result of a World Wide Web Consortium (W3C) working group that had a limited lifespan, the other two are established and open communities with participants mainly from the cultural heritage field and the Digital Humanities.

As shown in Figure 1, these specifications can be implemented separately and in conjunction with each other, since they rely on the same technological foundation as they are serialised in JavaScript Object Notation for Linked Data (JSON-LD), which allows some mapping of ontological constructs into JSON, the lingua franca for most of current web applications. Above all, these standards are predominantly built and maintained from the ground up while adhering to the architecture of the World Wide Web [18], reusing existing standards where possible. For example, the IIIF Change Discovery API relies on the W3C Activity Streams 2.0 standard to describe changes to resources and facilitates crawling to build search indexes [19].

As an exemplary model for other organisations, LUX<sup>6</sup>, the Yale Collection Discovery platform, has been upgraded to rely on LOUD standards, notably by leveraging Linked Art for facilitating data sharing across domains, and IIIF for seamless image delivery. It brings together over 40 million cultural heritage records (objects, works, people and organisations, places, concepts, and events) from libraries and museums within Yale University, making it a valuable reference for other institutions. While providing access to Yale’s collections, LUX also reconciles its data with external sources like Wikidata. This integration not only allows users to explore a vast range of cultural artefacts and information within Yale’s collections but also sets a precedent for other organisations aiming to enhance their own cultural heritage platforms. In short, LOUD could be considered as a pragmatic perspective for presenting or describing as well as sharing cultural heritage data on the Web that responds to the usability requirements of both scholars and developers working jointly through use cases. Thus, we strongly believe that implementing LOUD standards in the PIA digital infrastructure is a way to communicate specific values about sharing and open data practices, as well as to improve the resilience of cultural heritage data.

However, it is worth noting that one of the still largely unexplored areas of study concerning LOUD standards, in particular for the IIIF and Linked Art specifications which are exposed to a greater extent than the Web Annotation Data Model normally baked into other standards, is to carry out an assessment of these APIs in terms of usability factors, precisely to appraise one of the drivers of the perspective<sup>7</sup>.

---

<sup>6</sup>https://lux.collections.yale.edu/<br>
<sup>7</sup>What has already been done within the IIIF community are, for example, usability tests of IIIF-compatible
5. Participatory Design of Graphical User Interface

In this section, we discuss our approach to the participatory design of the graphical user interface. We believe that participatory development is essential in creating helpful and comprehensible GUI’s. We understand participatory research as the attempt to involve individuals and groups affected by the respective topic and question as active decision-makers [22]. Such participation can occur at different stages of the interface development process, for example, during conceptualisation, design, collaborative usage or evaluation of the prototypes. For our user research we apply a combination of Activity Centred [23] and Goal Directed [24] design methods that aim at understanding users goals and motivations in relation to activities and tasks performed in the interface. In doing so, we aim to develop a coherent user experience (UX) and interface principles that facilitate an active engagement beyond the project duration.

In the following, we highlight experiences from an early stage of the interface development where we worked with different humanities scholars from the fields of History and Cultural Anthropology that work with the photographic archives of the SSFS in their projects. The goal of the workshops was to define some of the visual, communicative, and functional requirements specific to humanities scholars for searching and interpreting digital cultural photographic collections. While Humanities scholars are one of the platform’s main target groups, we plan to diversify our workshop participants throughout the project duration to involve a broader public.

One of the two use cases of the 2022 workshops was the Images of Swiss Commons project. It aims to document and explore the mutability and innovation potential of the Swiss commons and the collective action of alpine farmers. This is performed by means of historical and contemporary photographs. Different actors of today’s collective will be involved, i.e. current or former officials of civil communities, businesses, alpine cooperatives, etc., as well as local historians.

With the support of the PIA platform, the researchers would like to use Ernst Brunner’s photo collection to enter into dialogue with representatives of interested organisations through “Calls for images”. The specific communities addressed will be invited to contextualise these historical photographs with their own visual documents on their collective forms of ownership and management.

Photographs are at the heart of the project because they offer non-historians a direct view of the past and a low-threshold access to the diversity of material and immaterial culture. Photographs provide important clues to cultural practices because they can convey implicit knowledge. This makes them valuable starting points for discussions that can be held on the digital platform, but also in real workshops. Hereafter are some of the challenges and questions that were discussed during the workshop:

- How should the comments and discussions with the addressed groups be curated and moderated?
- Is a login necessary or should we provide a low-threshold access to increase the chances of participation?

Software [20, 21] or the creation of personas by the IIIF Design Community Group (see https://iiif.io/guides/guides/personas/), but not yet an assessment at the specification level.
• Should the addressed groups be able to react to specific photographs or do we formulate open “Calls for Images”?

One concrete outcome of the 2022 workshop was a more detailed envisioning of how users can create a sub-collection through our platform (see Figure 2). This also includes the differentiation of a collection into specific themes, for example on the resources water, forest and alpine pasture in Images of Swiss Commons.

For the second use case, we collaborated with a member from the Mensch & Haus\(^8\) research project that documents the cultural evolution of historic farmhouses and their inhabitants across Switzerland. In this scenario the PIA platform is used to organise and analyse historic photographs from the landscape of Adelboden in Switzerland to support on-sight ethnographic surveys with inhabitants. In the workshop, a user journey (see Figure 3) was developed collaboratively to facilitate discussion and define the researcher’s motivation, tasks and challenges in using our GUI.

This collaboration helped us to understand two fundamental aspects that further informed the development of our GUI:

1. Our platform is used in different phases of research and is one of several tools with which researchers work. This highlights the emphasis on user-friendly import and export

\(^8\)https://data.snf.ch/grants/grant/189398
The overall goal of the PIA platform is to allow scholars and citizen scientists to engage with, functionalities that allow data to be transferred in and out of the platform in standardised formats such as JSON or CSV. This is in line with the principle of open systems and usable data, which gives users sovereignty over their own data [25]. It further shows that the LOUD design principles described earlier, especially the first two (*The right abstraction for the audience* and *Few barriers to entry*), are relevant for users.

2. As the researcher in this scenario works with a large image corpus of 300 images, it’s important that the interface provides the appropriate tools to explore the corpora intuitively. This includes, for example, the ability to group and organise images and to visualise them in different contexts, such as a timeline or a map. This led to the idea of generating multiple views in the interface (see Figure 4). Each group of images can be displayed in different visual contexts, such as a grid, table, map, timeline or network visualisation. Each view provides the user with specific interaction possibilities. For example: While the table provides advanced sorting functionality, the map allows searching for similar images within a certain radius of a given location.
edit, enrich and curate data of cultural heritage collections like the SSFS archive. On the one hand, this aims at enabling crowdsourcing in the usual sense by moderated metadata enrichment [6, 26], for example, when residents of an area make comments on historical photographs. On the other hand, the interface offers more far-reaching possibilities of participation by allowing users to create individual sub-collections and launch open “Calls for Images”. These user-generated collections are compilations of archival images and documents, as well as personal uploads and descriptions, that can be edited by individual or groups of users.

The challenge here is to find a good balance between control over content and low-threshold access. Related to this are in particular questions around the login-policy of the platform. The first prototypes prioritised easy access and most functionalities didn’t require a login. As the platform scales we are gradually introducing login requirements while trying to preserve easy access. For example: logins are not necessary to create collections in the GUI or annotating images in Mirador⁹, a IIIF-compliant viewer, as they are temporarily saved in the browser cache. To save collections and annotations in the long term and to edit them in detail, a login is required. Similarly, with “Calls for Images”, while a login is not required to submit entries, it is necessary for editing and curating user submissions. Other functions are planned with login:

⁹https://projectmirador.org/
editing one’s own sub-collection, initiating individual “Calls” as well as editing annotations or receiving appropriate credit for this.

The design and development of the PIA interface is enabled by a flexible software architecture using a headless approach, separating the back-end from the front-end and connecting them through the API. Inspired by the atomic design approach [27], each interface view (e.g. Grid, Table or Map) can be developed as an independent front-end component based on the same data structure, coding conventions and design system. On the one hand, this allows us to implement future views derived from new user requirements, on the other hand, it allows other projects to adopt parts of our codebase into their own. This approach further contributes to the GUI’s sustainability beyond the projects duration. It allows to maintain front-end and back-end separately, making it easier, for example, to find suitable partners for further development and make changes to parts of the used technology. The component based architecture enables the exchange of parts of the GUI without having to revise the overall architecture. We are currently discussing different models to ensure the sustainability of the GUI after the end of the project in 2025. We are working together with the SSFS photo archive, which could take over the maintenance. However, the costs incurred would have to be covered by the projects that want to carry out the "Calls for images".

6. PIA Overall Architecture

Within this section, we will first cover the high-level overview of the infrastructure and its environment, and then go into more detail about the generation of LOUD resources.

As mentioned previously, the software architecture is strictly separated into a headless back-end that provides the potential to represent data in complex digital data models, various interfaces for communication, and a front-end framework that is capable of embedding tools for different types of applications.

![Figure 5: High-level overview of the PIA Infrastructure and its connection to the DaSCH Service Platform (DSP) and the official SSFS Photo Archive Website.](image-url)
As can be seen in Figure 5, the PIA infrastructure runs parallel to the SSFS photo archive website, but in-between there is the DaSCH Service Platform (DSP), which manages the long-term data and metadata of the SSFS and provides its own APIs (IIIF Image API and DSP API); this is where the ground truth of the digital data lies. Our database is both a selection, as it is made up of only three collections from the archives, and an extension, due to the participatory needs of the platform, of the DaSCH database. In terms of APIs, we have currently deployed the core IIIF APIs (Image and Presentation APIs) as well as a bespoke Omeka S API.

To create IIIF resources that can be displayed and compared into any IIIF-compliant viewers, we have a specific workflow based on Laravel (see Figure 6), an open-source PHP framework, which generates IIIF Presentation API 3.0 resources (Manifests and Collections) as well as associated annotations (a series of AnnotationPage) that comply with the Web Annotation Data Model. Our main database is managed by Omeka S, which provides its own JSON-LD API. The data we collect through this API is combined with the machine learning Object Detection from vitrivr, which is hosted in a separate SQLite database. Next, we populate our temporary metadata collection with image-specific information provided by SIPI, our image server, such as pixel dimensions. The application then translates this data construct into the IIIF-specific format and serves the generated file to the client.

![Diagram of PIA IIIF Resources and Annotation Workflow](https://example.com/diagram.png)

**Figure 6:** Overview of the PIA IIIF Resources and Annotation Workflow. To create IIIF resources that can be displayed and compared into any IIIF-compliant viewers, we have a specific host application based on Laravel, an open-source PHP framework.

As IIIF takes an agnostic approach to the provision of descriptive metadata, and has no particular intention of imparting semantics to the content it carries [28], we have decided to rely on the seeAlso property to link to structured metadata. Currently, our IIIF Manifests point to the Omeka S API, which is loosely based on the SSFS data model and relies heavily on

---

10https://docs.dasch.swiss/
11https://codeberg.org/PIA/pia-iiif-manifest-host
12https://vitrivr.org/
13https://sipi.io/
the Schema.org vocabulary. It is also planned that the IIIF Manifests will point to a forthcoming Linked Art API, which is currently under preparation\(^{14}\).

The purpose of offering as many mutually agreed and shared APIs as possible from dedicated communities is to enable reuse of the data beyond the scope of PIA as well as to ensure improved data stewardship. Therefore, participation is not limited to the GUI, and we strive to provide thorough API documentation and working examples\(^{15}\).

### 7. Future Work

In the near future the GUI will be further developed and extended. The next steps particularly focus on the elaboration of the different visual access points, such as map, timeline and graph. Each of these approaches brings its own possibilities and limitations that are to be tested. The use of the map, for example, to explore the photographic archives and display cartographic material is of special importance to the project because one of its three collections, *Atlas of Swiss Folklore*, has genuine geographical qualities.

We are also planning two workshops at the *Bern Academy of Arts* to test the platform’s UX with different target groups. In the first workshop we will discuss with graphic and interface designers the participation functions and the visual appearance of the platform. The focus will be on how to motivate a broad audience to participate. In the second workshop, we will develop themes for future “Calls for images” with art education students who will later work, for example, in a museum or a high school.

Further action is needed in the development and integration of the metadata API and its documentation into the GUI as a user-friendly access point for programmatic interaction with the photographic archives.

For the evaluation of our APIs, and in particular for LOUD specifications, we still need to define a list of evaluation criteria. We believe that the necessary and important factors of the study are those of *learnability, efficiency, understandability, effectiveness, satisfaction*, and *readability* as there are factors commonly used to gauge an API’s usability. [29]. This evaluation is expected to be carried out as part of one of the PhD theses associated with the PIA project and will take into account the specificity of the various LOUD APIs.

As the interface and its technical infrastructures matures the building of communities for citizen scientists needs to be assessed from a communicative perspective and in collaboration with the SSFS to enable sustainable usage of the collections beyond the project duration.

Last, but not least, synchronisation aspects between PIA and DaSCH will need to be clarified. It is quite possible that the data co-exist in parallel with each other while implementing a data curation and reconciliation strategy with all stakeholders, which could last after the project. Although this effort will not be straightforward, it will be streamlined by having interoperable and open systems.

\(^{14}\)https://github.com/Participatory-Image-Archives/linkedart/

\(^{15}\)For instance, on the Observable HQ platform: https://observablehq.com/@participatory-archives
8. Conclusion

Within PIA, we develop an environment enabling a digital workflow that starts at the original printed source and ends where experts and citizens enrich the data with their knowledge. The close dialogue between humanities researchers, archivists, and experts in design and software development, ensures a highly applicable solution upon which to engage in constructive criticism. By the transfer of archival methodologies and processes from the analogue to the digital domain, we create a sustainable aura for stored data. The innovative GUI and the integration of APIs encourage collaboration with the public and, thus, a variety of open-ended interpretive perspectives.

The participatory design approach taking place at different stages of the process – from early development to final evaluation – allows for the inclusion of a variety of perspectives in the use and reception of the SSFS archives. We hope that by giving many different groups a voice in the curation of the collections, we can contribute to a democratisation of the decision-making power of archives [30].

As we show, the sustainability of data and digital tools is closely related to the application; we go beyond open data by demonstrating the power of standardised APIs, namely those that adhere to the LOUD design principles. The possibility to enrich data makes data sustainable and increases the attraction of digital infrastructures.

Acknowledgments

We would like to express our sincere gratitude to all the colleagues, partners, and individuals involved in the PIA research project, whose invaluable contributions, support, and collaboration were instrumental in the successful completion of this paper.

Participatory Knowledge Practices in Analogue and Digital Image Archives (PIA) is supported by the Swiss National Science Foundation (SNSF) as part of the Sinergia funding scheme (contract no. CRSII5_193788): https://data.snf.ch/grants/grant/193788.

References


