Enabling Participatory Data Perspectives for Image Archives through a Linked Art Workflow

University of Basel

Julien A. Raemy[1], Tanya Gray[2], Alwyn Collinson[2], Kevin R. Page[3]

[1]Digital Humanities Lab, University of Basel, Switzerland

[2]Centre for Digital Scholarship, University of Oxford, United Kingdom

[3]Oxford e-Research Centre, University of Oxford, United Kingdom

Digital Scholarship @Oxford

Oxford e-Research Centre

University of Oxford

Motivation

The Participatory Knowledge Practices in Analogue and Digital Image Archives (PIA) [1] project, led by the University of Basel and the Bern Academy of the Arts, aims to encourage participation from scholars and the wider public through three collections from the photographic archives of the Swiss Society for Folklore Studies (SSFS): The *Atlas of Swiss Folklore*, *Ernst Brunner*, and *Kreis Family*. PIA aims to create multiple interfaces reflecting diverse perspectives by deploying community-developed Linked Open Usable Data (LOUD) specifications [3, 6] such as the International Image Interoperability Framework (IIIF) [5] and Linked Art [4], an RDF application profile (JSON-LD) based on the CIDOC Conceptual Reference Model (CRM) and the Getty vocabularies to describe cultural heritage resources, with a focus on artworks and museum-oriented activities [3, 7].

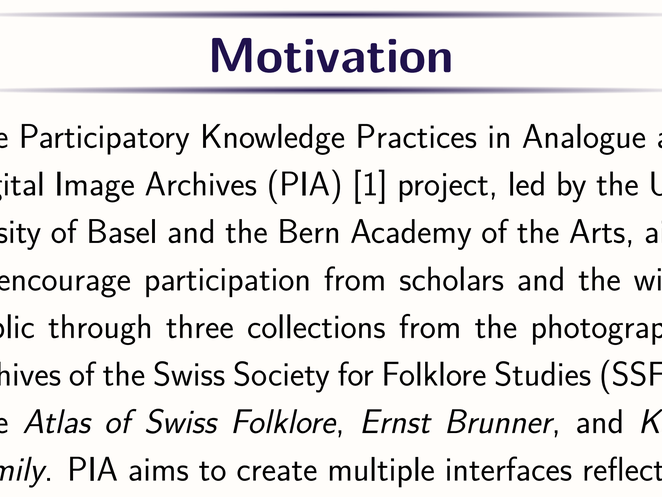

Figure 1: Linked Art from 50,000 feet. Adapted from a diagram by Robert Sanderson [8].

In collaboration with the University of Oxford via the Linked Art II project [2], PIA has transformed their cultural heritage collection data into Linked Art using templates, which encapsulate the data characteristics and cataloguing practices. A Linked Art API will provide an additional entry point, as a means of conveying semantically enriched events and as a benchmark against other collections leveraging this model.

Python-based Workflow

To generate Linked Art files for the combined PIA collection a data transformation workflow [9] has been created through which PIA templates [10] were encoded in Python for a given Linked Art API entity endpoint [11], currently: DigitalObject, HumanMadeObject, and Set. The workflow, described in a use case example with a photograph from the Ernst Brunner collection, provides a three-step software process for transforming data into Linked Art.

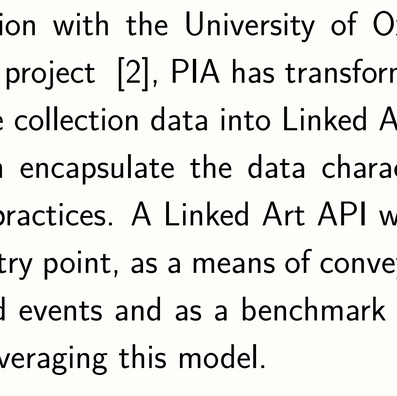

Figure 2: [Schwyzer Fasnacht]. Schwyz, 1937. Ernst Brunner. Original Medium: Black and White Negative, 6x6 cm. SGV_12N_00001. SSFS. CC BY-NC 4.0. https://archiv.sgv-sstp.ch/resource/422236.

1) Query

The first Python script extracts collection data from the (legacy) PIA JSON API. YAML front matters are used for script variables. The filepath for the relevant `.yaml` file is specified as a script argument throughout the workflow. For our example, the script queried data for all images that are DigitalObject entities and the object's metadata that are stored come specifically from https://json.participatory-archives.ch/api/v1/images/12033.

2) Map

Templates are used to map collection data to an intermediate JSON data format. The intermediate JSON data format means that the later transformation script that creates Linked Art JSON-LD does not necessarily need to be modified if a new data source is introduced.

```
{
"_label": "[Schwyzer Fasnacht]",
  "access_point": {
    "_label": "Image in full resolution",
    "id": "https://sipi.participatory-archives.ch/SGV_12/SGV_12N_00001.jp2/full/max/0/default.jpg"
  },
  "created_by": {
    "_label": "Digitisation of Photograph",
    "used": {
      "_label": "Negative of [Schwyzer Fasnacht]",
      "produced_by_event": {
        "begin": "1937-01-01",
        "display_title": "GREGORIAN:CE:1937-01-01:CE:1937-12-31",
        "end": "1937-12-31",
```

Listing 1: Snippet from the mapped data of SGV_12N_00001

3) Transform

The intermediate JSON data format is transformed to Linked Art with Python functions that define patterns (for example classified_as) for representing different aspects of photographic collection data as Linked Art.

```
{
    "@context": "https://linked.art/ns/v1/linked-art.json",
    "id": "https://data.participatory-archives.ch/digital/12033",
    "type": "DigitalObject",
    "_label": "PIA ID 12033 - [Schwyzer Fasnacht]",
    "classified_as": [
      {
        "id": "http://vocab.getty.edu/aat/300215302",
        "type": "Type",
        "_label": "Digital Image"
      }
    ],
```

Listing 2: Snippet from the Linked Art JSON-LD of SGV_12N_00001

Future Work

The SSFS will migrate its database into the DaSCH Service Platform (DSP) [13] and amend their data model which will affect PIA, requiring an upgrade of its infrastructure and APIs. The workflow for creating Linked Art representations will have to be repurposed with different data sources.

After the migration of the SSFS database, the PIA team will investigate remaining issues in the GitHub repository regarding the correct mapping of Linked Art entities and attribution of IDs. The team will work to ensure that the appropriate Linked Art modelling is achieved through the workflow.

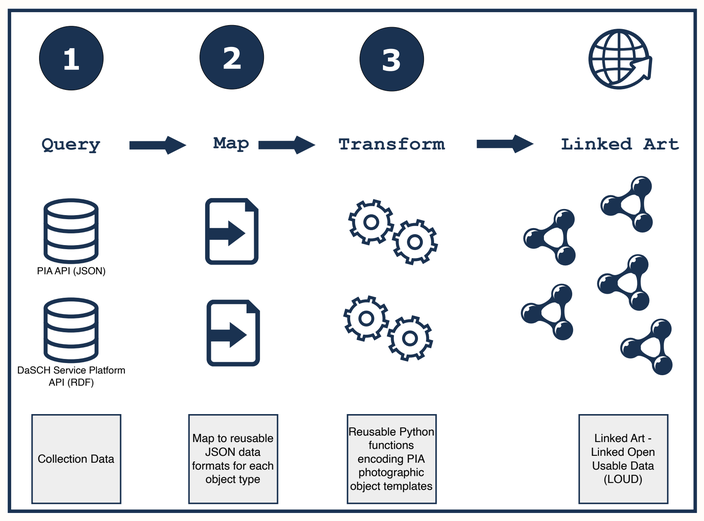

Figure 3: PIA Linked Art data workflow designed for reuse with different data sources and object types.

Conclusion

The reconfigurable Python-based workflow is able to transform cultural heritage data, initially photographic collections, into LOUD, as **a foundation for varied participatory interfaces supporting scholarship and beyond**. The adaptability and extensibility of the workflow allows for potential future transformations of data from other collections to Linked Art.

References

[1] https://about.participatory-archives.ch

[2] https://linked.art/community/projects/linkedartii/

[3] Newbury, David (2018): LOUD: Linked Open Usable Data and linked.art, in *2018 CIDOC Conference*, Heraklion, Greece: International Council of Museums, pp. 1-11.

[4] Page, Kevin R., Delmas-Glass, Emmanuelle, Beaudet, David, Norling, Samantha, Rother, Lynn & Hänsli, Thomas (2020): Linked Art: Networking Digital Collections and Scholarship, in *DH2020 Book of Abstracts. Digital Humanities 2020*, Online, pp. 504-509.

[5] Raemy, Julien Antoine & Demleitner, Adrian (2023): Implementation of the IIIF Presentation API 3.0 based on Software Support: Use Case of an Incremental IIIF Deployment within a Citizen Science Project, in *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection*. Cham: Springer International Publishing (Lecture Notes in Computer Science).

[6] Sanderson, Robert (2018): Shout it Out: LOUD. *EuropeanaTech Conference 2018*, Rotterdam, the Netherlands, 15 May 2018.

[7] Sanderson, Robert (2019): Keynote: Standards and Communities: Connected People, Consistent Data, Usable Applications, in *2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, Urbana-Champaign, Illinois, USA: IEEE, p. 28.

[8] Sanderson, Robert (2020): The Importance of being LOUD. LODLAM 2020, Los Angeles, CA, 5 February.

[9] https://github.com/tgra/Linked-Art-Collection-Data-Workflow

[10] https://github.com/Participatory-Image-Archives/linkedart

[11] https://linked.art/api/1.0/endpoint/

[12] https://github.com/thegetty/crom

[13] https://docs.dasch.swiss/

Raemy, Julien A., Gray, Tanya, Collinson, Alwyn & Page Kevin R. (2023): Enabling Participatory Data Perspectives for Image Archives through a Linked Art Workflow. *Digital Humanities 2023*, Graz, Austria, 12 July 2023. https://doi.org/10.5281/zenodo.7878358

This work has been supported by the PIA research project (2021-2025) which is funded by the Swiss National Science Foundation (SNSF) as part of the Sinergia funding scheme (contract no. CRSII5_193788) and by the Linked Art II project at the University of Oxford (Principal Investigator: Dr. Kevin R. Page, Oxford e-Research Centre) funded by the UK Arts and Humanities Research Council (AHRC project reference AH/T013117/1).