

# Towards an Annotation Data Model for a Scholarly Semantic Annotation Platform in Visual Heritage: A Case Study Using the Murten Panorama

Tsz Kin Chau<sup>1,\*</sup>, Daniel Jaquet<sup>1</sup> and Sarah Kenderdine<sup>1</sup>

<sup>1</sup>Laboratory for Experimental Museology, EPFL, Switzerland

## Abstract

Historically, the ability to reproduce visual works in print for comparison and detailed illustration has been crucial for art historical research and its dissemination. While this has always been possible, the study of big monolithic visual heritage can be greatly streamlined by an interactive digital research environment. The goal of our research is to develop an annotation platform that leverages linked open data to facilitate a thorough and scholarly description of big monolithic visual heritage. We will deploy our platform to craft a scholarly edition of the Murten Panorama (1894), offering high-quality, well-provenanced knowledge graph to both the public and scholars, giving them a window into the historical context of the creation of the 19th c. panoramic masterpiece.

## Keywords

Semantic Web Applications for Cultural Heritage, Virtual Research Environment, Semantic Annotation, Panorama, Cultural Heritage, Digital Art History, Digital history, Ontology, CIDOC-CRM

## 1. Introduction

The scholarly semantic annotation platform proposed in this paper is conceived for, though it is not limited to, the study of big monolithic visual heritage. There are two dimensions to the concept of big monolithic visual heritage. Firstly, big in extent, either in physical size, exemplified by painted panoramas or mural paintings, or in digital size, as seen in macroscopic and microscopic images. Secondly, big in density, signifying visual heritage with rich composition, motifs, decorations, cultural references, and localized features or narratives, making it a complex and composite image. These two facets of “big” are interconnected. Certain objects might be small in size but contain dense features, requiring gigapixel imaging to capture their content fully, thus rendering them digitally big.

Historically, the ability to reproduce visual works in print for comparison and detailed illustration has been crucial for art historical research and its dissemination. While this has always been possible, the study of big monolithic visual heritage can be greatly streamlined by an interactive digital research environment. Recent advances in gigapixel imaging and the International Image Interoperability Framework (IIIF) allow researchers to pan and zoom into large images, examining multiple layers of X-ray and RGB data in real-time. This enables close examination and comparison with other visual and textual sources. Furthermore, linked open data and ontologies, such as CIDOC-CRM, provide fine-grained descriptions of the knowledge created during the study of visual heritage objects, facilitating the documentation of processes, theories, and citations that are employed in scholarly researches.

The goal of our research is to develop an annotation platform that leverages linked open data to facilitate a thorough and scholarly description of big monolithic visual heritage. There are two distinctive features in our method. First, we dive deep into the visual material, isolating and describing local features (also referred to as Point of Interests (POIs)), that make annotation the first-class citizen in our research. Second, distinct from mainstream annotation tools, we also focus on data trustworthiness by documenting the scholarly process involved in an annotation.

---

SWODCH'24: *International Workshop on Semantic Web and Ontology Design for Cultural Heritage*, October 30–31, 2024, Tours, France

\*Corresponding author.

✉ [tszkin.chau@epfl.ch](mailto:tszkin.chau@epfl.ch) (T. K. Chau); [daniel.jaquet@epfl.ch](mailto:daniel.jaquet@epfl.ch) (D. Jaquet); [sarah.kenderdine@epfl.ch](mailto:sarah.kenderdine@epfl.ch) (S. Kenderdine)

🆔 0009-0006-7939-3151 (T. K. Chau); 0000-0003-1515-3619 (D. Jaquet); 0000-0002-7190-9946 (S. Kenderdine)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

This paper presents parts of our ongoing research, focusing on the development of the data model for our annotation platform. We begin by introducing the concept of our scholarly semantic annotation platform, followed by a discussion on the formalization of our annotation data model. This paper contributes the following: 1) an outline to leverage semantic web technology to enable scholarly annotation and analysis in visual heritage, and 2) a proposal for the reuse and further development of existing visual interpretation ontologies to create a standards-compliant ontological data model that drives a complex knowledge graph management platform application.

## 2. Background: Digitising and Augmenting the Panorama of the Battle of Murten (DIAGRAM)

The visual heritage item at the centre of our research, the Murten Panorama (1894) by Louis Braun (1836-1916) is an illustrative example of visual heritage characterized by both its vast scale and intricate detail. Depicting the Battle of Murten on June 22, 1476, during the Burgundian Wars fought between the old Swiss Confederacy and Charles the Bold, the Duke of Burgundy, the Murten Panorama stands as both a Swiss national treasure and a visual heritage of international significance.

Physically measuring approximately 10 x 100 meters, the panorama was digitized in 2023 at 1,000 dpi and fully processed in June, 2024 [1] under the initiative of the DIAGRAM project, creating a 1.6 Terapixels digital twin. To commemorate the 550th anniversary of the Battle of Murten and to advocate for the panorama's recognition as a UNESCO Memory of the World, a series of augmented immersive installations and a scholarly edition of the Murten Panorama will be produced in 2024-2026.

The content within the panorama is exceptionally diverse, encompassing a wide range of named geographical locations, historical characters, heraldic representations, recognized historical events, about 5,000 people (including 26 women and 1 child) in various costumes and arms, and 700 horses. Additionally, it is rich in cultural-historical references, comprising an array of visual elements such as weapons, flags, costumes, and narratives that can be traced in museum collections, illustrated chronicles, and historical documents, providing an immense opportunity for linked data annotation.



**Figure 1:** Louis Braun, *The Panorama of the Battle of Murten*, 1893, approximately 10 x 100 metres

## 3. Related Work

A platform conceptually similar to our conceived platform is Geovistory [2], a collaborative, web-based research and data publication environment. It includes Geovistory Toolbox that allows researchers to collect, curate, and evaluate data conforming to the methodologies of historical science. The toolbox offers strong semantic support from CIDOC-CRM, FRBRoo, and a community of data profiles for handling different types of source materials.

While Geovistory primarily focuses on text-based materials, in the visual domain, open-source annotation libraries that support IIIF images include Mirador Annotations [3] and Annotorious [4]. Both libraries comply with the Web Annotation Data Model [5] and are widely adopted for online annotation tools, crowd sourced annotation, as well as for creating guided virtual exhibitions.

Several annotation tools are specifically designed to document the deep scholarly context. Pliny was developed to document the process of scholarly interpretation [6]. Pundit offers a “Triple Composer” feature, which allows annotators to describe content in a named graph [7].

Some tools are tailored for art historians' analytical needs. HyperImage Virtual Research Environment [8] and ARIES (ARt Image Exploration Space) [9] provide a light table virtual environment for visual studies, allowing manipulations such as rearranging, resizing, and comparing images to study visual relationships. Tropy [10] is a desktop-based, open-source personal research image management tool that supports POI-based image annotation and metadata customization.

During our review process, we concluded that while tools like Geovistory meet our scholarly needs, the lack of visual annotation support renders it unsuitable for our purposes. Alternatively, we identified ResearchSpace [11], an open-source, template-driven, and highly customizable knowledge graph management platform built on the CIDOC-CRM framework and its extensions, as particularly well-suited for the experimental development of our proposed annotation method. Notably, its unique focus on image annotation knowledge graph authoring makes it an ideal fit for our use case.

## 4. Development of the Annotation Data Model

### 4.1. Goal

To fully benefit from the ontological resources available on ResearchSpace, we aim to express our annotations using classes and properties from CIDOC-CRM and its extensions. This approach requires us to develop our data model apart from the widely adopted Web Annotation Data Model. Our model is designed to encompass not only the content produced by an annotation but also the workflow involved.

### 4.2. Review of existing ontologies

For annotation specific ontology, the Web Annotation Data Model (WADM) [5], is a standard-setting data model powering countless applications. It is also integrated into important data models for cultural heritage, such as the IIF Presentation API and the Europeana Data Model (EDM) [12].

For visual interpretation ontologies, VIR [13] is developed as a CIDOC-CRM extension for describing the visual recognition process in visual heritage, grounded in visual recognition and communication theory. Additionally, ICON [14] offers a fine-grained approach specifically tailored to represent Panofsky's three-level iconographical analysis.

For representing scholarly assertions, particularly in historically-oriented humanities, models such as Factoid [15], symogih.org [16], and STAR (Structured Assertion Record) [17] provide a common pattern to structure fact/event extractions. HiCO [18] extends this assertion pattern by introducing interpretation criterion, which allows for the documentation of interpretative activity involved in the extraction process. Finally CRMInf [19] provides a framework for describing argumentation and inference activity.

### 4.3. Characterizing the Visual Annotation Domain

Our ontology design process is inspired by established ontology design methods including the SAMOD methodology, ontology design patterns (ODP) [20], and the middle-out approach [21]. Our formalization team is composed according to the roles described in the SAMOD methodology [22], namely Ontology Engineer (OE) and Domain Expert (DE). The team consists of two members:

- **Member 1:** Serves as both ontology engineer and domain expert in art history, specializing in hierarchical iconographical analysis.
- **Member 2:** Represents the domain expert in medieval history, specializing in military material culture and martial arts.

From our ground source studies, we identified so far the following motivating scenarios (Table 1).

**Table 1**  
Motivating scenarios identified through ground source studies

Motivating scenario	Description	Example
Image-Image Comparative Annotation	Describes visual connections between related historical images or image segments.	See 4.4.1.
Image-Text Historical Event Annotation	Contextualizes and compares the annotation subject with a real historical event based on historical documents.	An example of this pattern is the death of John of Luxembourg (1437-1476) described by Panigarola's letter. The case was discussed in a previous presentation [23].
Image-Object Comparative Annotation	Describes visual relationships between related images or image segments and a museum collection item.	A pavise depicted in the panorama, with its equivalent item (KZ-386) found at the Swiss National Museum.
Image-Image Similarity Assessment	Evaluates the type and degree of similarity between related images or image segments.	Evaluate the similarity of the coat of arms representation in the panorama versus the historical sources.
Image-Real-World Entity and Concept Linking	Most common annotation task that identify and classify images or image segments to the represented person, concept, or place.	Identify Charles the Bold (1433-1477) by attributes, such as the collar of Order of the Golden Fleece.

### 4.3.1. Challenges

A key challenge in formalizing our data model lies in accommodating the multi-domain nature of the users on our proposed platform, while our modeling team is limited to two domain experts, which may not be sufficient to generate all motivating scenarios. To address this limitation, we plan to develop the annotation platform and data model with the community of historians through a series of workshop sessions. The feedback and additional case studies gathered during these sessions will be used to evaluate and refine our annotation data model.

## 4.4. Result from the first iteration

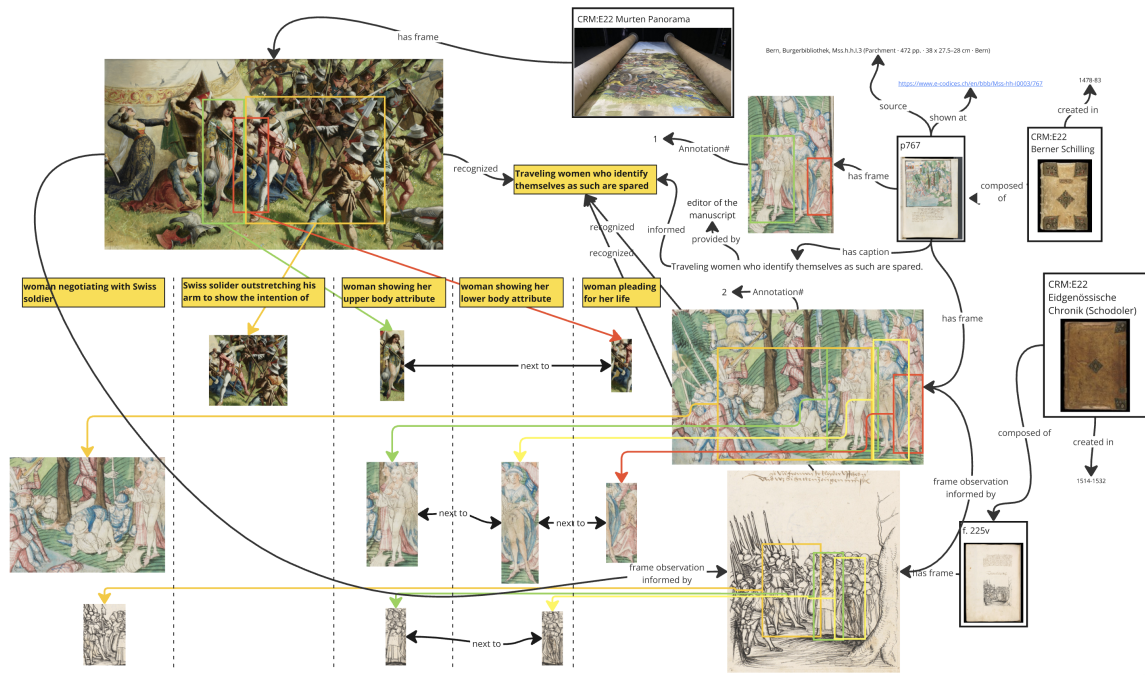
### 4.4.1. Case study

The case study presented in this section is an example extracted from the motivating scenario: image-image comparative annotation. Table 2 provides a detailed description of the scenario. This case was selected for discussion because of its complexity and its significant contribution to the development of our current annotation data model.

**Table 2**  
Motivating scenario: image-image comparative annotation

Description	Upon receiving an image that appears visually connected to the annotation subject, the first step is to perform source criticism. This involves evaluating the image's temporal extent, creator, provenance, and other relevant details. Afterward, describe the visual relationship between the image and the annotation subject. Even if direct evidence of an influence cannot be found, we assume these images are linked through a visual transmission process, similar to what is seen in manuscript transmission, where prototypes or representations are transmitted through both tangible and intangible pathways. Our goal is to capture and describe the deep visual relationships between images in our system, enabling future analysis of how the prototype evolves over time.
Competency questions (CQ)	What are the changes over time for a particular representation?

Developed through the collaboration of both the OE and DE, a sample annotation presented as a knowledge graph mock-up is shown in Fig. 2. The three selected images/image segments include an illustration from the Berner Schilling Chronicle (15th c., Mss.h.h.I.3, Burgerbibliothek Bern), an illustration from the Werner Schodoler Chronicle (16th c., ZF 18, Aargauer Kantonsbibliothek), and a segment from the Murten Panorama (19th c.). All three depict the same scene: a group of women affiliated to the Burgundians being spared by Swiss soldiers during the Battle of Murten. These depictions



**Figure 2:** The knowledge graph mock-up of our selected case study

differ in detail, with the most notable variation being the portrayal of the women’s role. In the former two, a woman is shown negotiating with the Swiss soldiers. In contrast, the 19th c. Murten Panorama depicts her as passively protected by a Swiss soldier.

#### 4.4.2. Takeaway from the case study

The first takeaway is the layers of knowledge produced in the image-to-image comparative annotation process. The first layer of knowledge involves the arbitrary selection of an image region as the boundary for a visual interpretation. We refer to this process as “framing”, and the selected region itself as a “frame”, borrowing terminology from photography. Framing is a critical scholarly process. As illustrated in the mock-up, the original frame provided by Member 2 (annotation #1) omitted a part of the image, which could lead to significant changes in subsequent visual recognition.

The second layer of knowledge produced involves associating a representation with a frame through visual recognition. In the example, the primary representation is “Traveling women who identify themselves as such are spared”. To deeply annotate the visual relationship, a frame can be subdivided into subframes, each associated with a subframe representation that contributes to the overall meaning of the parent frame’s representation. Such methodology is widely adopted in iconographical method and is formalized in VIR and ICON.

The third layer of knowledge produced is the arbitrary appellation assigned to a representation affected by the annotator’s domain knowledge and perspective.

The second takeaway is scholarly assertion justification by referencing external sources or knowledge. While some frames are spontaneously observable, other frames require external knowledge, often by using another frame as a reference. This is exemplified by both the Berner Schilling Chronicle and the Murten Panorama. Another example of external knowledge use is in naming the primary representation, which is informed by the editor’s caption.



### 4.4.3. Representating Visual Recognition and Hierarchical Analysis

Our visual recognition model follows the CIDOC-CRM pattern of E24 Physical Human-Made Thing → P65 shows visual item → E36 Visual Item. Initially, we modeled our case study according to the VIR ontology. We chose VIR over the more fine-grained ICON ontology due to its domain-neutral design, which better aligns with our use case.

When applying the VIR ontology (Fig. 3) to our case study, two specific issues arise. First, VIR’s formalization of the relationship between the physical visual heritage item (crm:E22) and its subregion (vir:IC1) does not align with our concept of framing and the frame. Both vir:IC1 and its superclass, crm:E25 Human-Made Feature, define the feature region as “purposely created by human activity”, which does not capture the that a frame is an intentional product of the annotator.

The second issue concerns the relationship between vir:IC9 Representation and vir:IC10 Attribute. In our case study, the main representation “Traveling women who identify themselves as such are spared” is assigned to vir:IC9 while subframe representations, such as “woman negotiating with Swiss soldier” are assigned to vir:IC10. In VIR, a representation inherently incorporates its subframe representations. However, since the subframe representations differ across the three instances in our case study, it becomes impossible to aggregate them under a common vir:IC9.

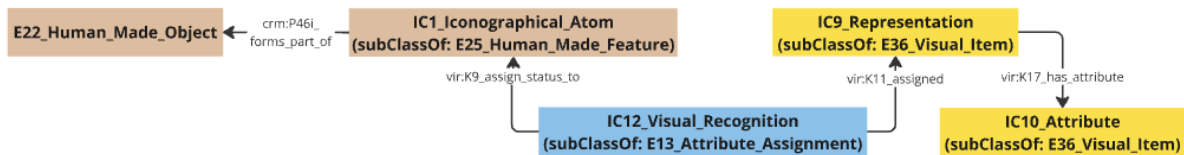


Figure 3: Visual recognition classes the VIR ontology [13] (graph created by the authors)

Fig. 4 presents our proposed alternative to VIR. To represent the framing of a visual heritage item, we introduce :Framing, a subclass of crm:E13 Attribute Assignment, which assigns crm:P46 to :Frame (subclass of crm:E25). The use of a crm:E13 characterizes :Frame as a product of some agent’s activity. Following the framing process, our :Visual\_Recognition adhere to VIR’s original formalization except we rename vir:IC9 to :Recognized\_Visualitem for a more domain-neutral terminology.

For representation and its subframe representations, we use the recursive :Framing and :Frame pattern. A subframe is essentially another :Frame that can be assigned a Recognized Visualitem. To aggregate all subframe representations, one simply needs to traverse to the :Frame nodes and access all associated subframes.

Traversing the frame hierarchy can be done efficiently through the property path of crm:P141i\_was\_assigned\_by/crm:P140\_assigned\_attribute\_to. We tested the model using SPARQL (appendix A) to answer the CQ: “What are the changes over time for a particular representation?” and obtained the expected results.

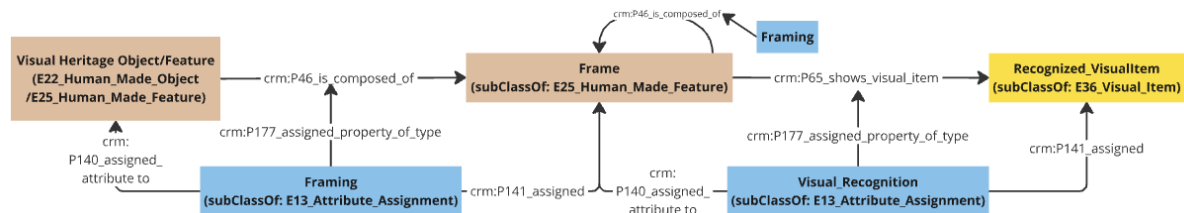


Figure 4: Visual recognition classes from our data model

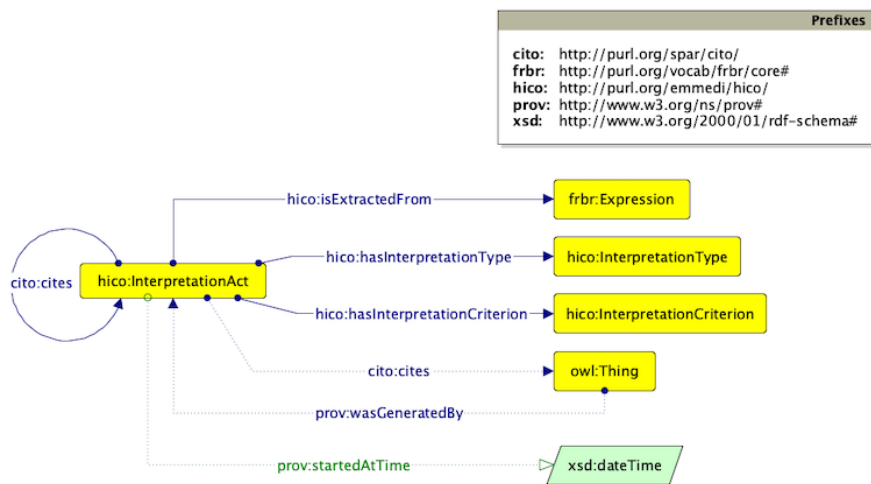
#### 4.4.4. Representing scholarly assertion

At this stage, we have identified three classes of scholarly assertion which are summarized in Table 3, along with their definitions and associated details.

**Table 3**  
Classes of scholarly assertions identified for the annotation platform

Assertion Class	Definition	Subject (crm:P140)	Object (crm:P141)	Predicate (crm:P177)
:Framing	Observation of a frame from a visual heritage item or other frame	E22, E25 or :Frame	:Frame	crm:P46
:Visual_Recognition	Association of a representation to a visual heritage item or feature	E22, E25 or :Frame	:Recognized_VisualItem	crm:P65
:Appellation_Assignment	Naming of a recognized visualItem	:Recognized_VisualItem	E33_E41	crm:P1

The three classes of scholarly assertion are formalized as subclasses of `:Scholarly_Assertion`, an adaptation of HiCO's `hico:InterpretationAct` (Fig. 5) utilizing `crm:E13`. `hico:InterpretationAct` documents the provenance of statement-extracting (such as actor roles in historical events) scholarly hermeneutical activities. It is lightweight yet adequate to capture holistically the various aspects of an interpretation act, including sources from which the statement is extracted, citations, interpretation methods, and relationships to other interpretation acts.



**Figure 5:** The `hico:InterpretationAct` [18]

Although `crm:E13` and `hico:InterpretationAct` connect to asserted statements differently, with `crm:E13` connecting directly to the statement triple through its properties and `hico:InterpretationAct` connecting indirectly to the generated statements via a `prov:Entity` node, the documentation properties of `hico:InterpretationAct` can be sufficiently represented using `crm:E13` (Table 4).

Our modification to HiCO involves using a reification node `crm:PC16` to connect sources and external knowledge to their respective assertion criterion. This adjustment accommodates the edge case that a scholarly assertion utilizes multiple sources or pieces of external knowledge. In our case study, to represent the reference of another frame in an `:Framing` activity, we use `crm:PC16→crm:P02_has_range` to link to the reference `:Frame` and `crm:PC16→crm:P16.1_mode_of_use` to an interpretation criterion vocabulary (`crm:E55`), such as “infer from another frame”. In an edge-case scenario, we want to further support our assertion by citing a historical document that describes our observed frame as a single unit. We will need a second pair of “Source / External Knowledge” and “Assertion Criterion”, and only through the reified `crm:PC16` that such pairing can be structurally ensured.

**Table 4**  
Our model of scholarly assertion and the related HiCO and CIDOC-CRM properties

Data element	HiCO properties	Our selection of CIDOC-CRM property (path)
Assertion Creator Additional Type Source/External Knowledge	prov:wasAssociatedWith hico:hasInterpretationType cito:cites	crm:P14_carried_out_by crm:P2_has_type crm:P01←crm:PC16_used_specific_object→crm:P02_has_range
Assertion Criterion	hico:hasInterpretationCriterion	crm:P01←crm:PC16_used_specific_object→crm:P16.1_mode_of_use
Relation to Other Assertion	sub-properties of cito:cites	sub-properties of cito:cites

The scholarly assertion pattern plays a pivotal role in aggregating knowledge in a multi-domain annotation contexts. In our visual recognition component, we do not make any domain-specific assumptions about the `:Recognized_Visualitem`. For instance, a sword could be the simplest recognition or a pre-iconographical recognition that contributes to a complex visual symbol. In our proposed formalization, domain knowledge can be expressed through vocabulary by using `:Visual_Recognition→crm:P2_has_type` to characterize a `:Recognized_Visualitem`, for example, as “pre-iconographical” (`crm:E55`). In this formalization, we aim to aggregate knowledge from different domain annotation using the domain-neutral `:Recognized_Visualitem` node.

We tested the model using SPARQL (appendix A) to answer the competency question: “What is the extended recognition from other domains for a particular frame?” and obtained the expected results.

#### 4.5. Limitation and discussion

Our proposed `:Recognized_Visualitem` serves as the aggregating node to connect the historical instances of `:Frame` that embody it. Structurally, `:Recognized_Visualitem` should be mapped to the level of `crm:E89/frbroo:F1` (`frbroo` is now `LRMoo` [24]), which is defined as the common idea and underlying prototype (`crm:E89`) that evolves over time (`frbroo:F1`), while `crm:E36`, a subclass of `crm:E73`, is already realized as an identifiable immaterial item.

Mapping `:Recognized_Visualitem` to `frbroo:F1` may remove its visual quality and turn it into modal-neutral which can then serve as cross-modal aggregation, such as the textual mention of “Traveling women who identify themselves as such are spared”. Let’s call this class the `:Cross-modal_Recognized_Item`.

However, `:Cross-modal_Recognized_Item` requires a new property other than `crm:P65` for connecting a `:Frame`. One possible option would be `:Frame→crm:P130_shows_features_of→frbroo:F1`. Yet, in CRM-base, `crm:P130` refers to the generalization of the notions of “copy of” and “similar to”, which might not be a perfect match. A more formal path would be `:Frame→crm:P128_carries→frbroo:F2→frbroo:R3_realises→frbroo:F1`, together with a shortcut property to bypass documenting the `frbroo:F2` level, following the pattern from `crm:P62`.

## 5. Conclusion and outlook

We have demonstrated our motivation, design concept, and work-in-progress data model for a scholarly semantic annotation platform. We believe that our proposed method will contribute to art-historical research, GLAM data curation, and scholarly editing across various types of visual heritage.

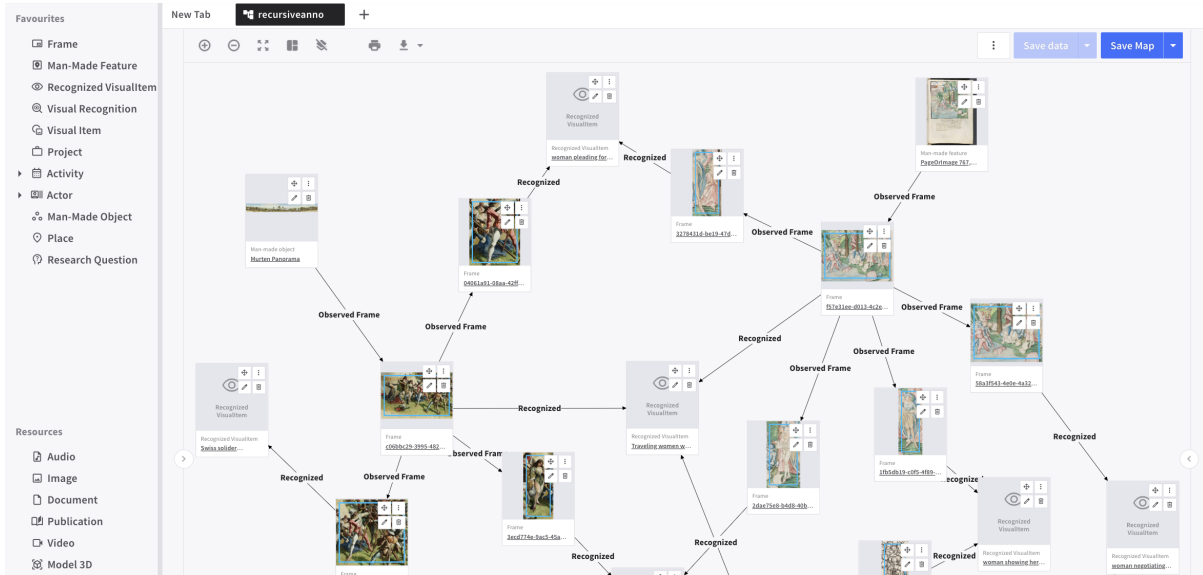
Looking ahead, we will continue to expand our use case examples through workshop sessions and expand our annotation data model with additional features. In the next iteration, in particular, we will explore how to enable argumentation among scholarly assertions, with the potential integration of CiTO [25] or CRMInf. We will also explore the impact of AI, such as object detection, on our data model, which could become a key feature in a human-in-the-loop, AI-assisted annotation environment.

We are actively developing the annotation platform to deploy our proposed data model and designing the corresponding workflows and user interface. Additionally, we have customized ResearchSpace’s



image annotation feature to support our recursive framing methodology (Fig. 6). Users of our system can link digitized images from the vast IIF resources available on the web.

We will deploy our platform to craft a scholarly edition of the Murten Panorama, offering high-quality, well-provenanced knowledge graph to both the public and scholars, giving them a window into the historical context of the creation of the 19th c. panoramic masterpiece.



**Figure 6:** Screenshot of recursive frame annotation in the ResearchSpace Knowledge Map view

## References

- [1] P. Bourke, The TERAPIXEL Panorama Project, 2024. URL: <https://www.epfl.ch/labs/emplus/projects/murten-panorama-digital-twin-scanning-project-the-making-of/>.
- [2] V. Alamercury, F. Beretta, F.-J. Favey, D. Ferhod, S. Hart, D. Knecht, G. Muck, A. Perraud, M. Pica, J. Schneider, A. Stebler, Open Research Practices with the OntoME-Geovistory environment, 2023. URL: <https://shs.hal.science/halshs-04162294>. doi:10.5281/zenodo.8107384.
- [3] ProjectMirador, mirador-annotations, 2015. URL: <https://github.com/ProjectMirador/mirador-annotations>, original-date: 2020-05-06T12:54:15Z.
- [4] R. Simon, annotorious-openseadragon, 2013. URL: <https://github.com/annotorious/annotorious-openseadragon>.
- [5] R. Sanderson, P. Ciccarese, B. Young, Web Annotation Data Model, 2017. URL: <https://www.w3.org/TR/annotation-model/>.
- [6] J. D. Bradley, Pliny: A model for digital support of scholarship, *Journal of Digital Information* 9 (2008). URL: <https://jodi-ojs-tld.tdl.org/jodi/index.php/jodi/article/view/209>, number: 1.
- [7] M. Grassi, C. Morbidoni, M. Nucci, S. Fonda, F. Piazza, Pundit: augmenting web contents with semantics, *Literary and Linguistic Computing* 28 (2013) 640–659. URL: <https://academic.oup.com/dsh/article-lookup/doi/10.1093/llc/fqt060>. doi:10.1093/llc/fqt060.
- [8] J.-M. Loebel, H.-G. Kuper, *HyperImage: Of Layers, Labels and Links.*, Riga, Latvia, 2013.
- [9] L. Crissaff, L. Wood Ruby, S. Deutch, R. L. DuBois, J.-D. Fekete, J. Freire, C. Silva, ARIES: Enabling Visual Exploration and Organization of Art Image Collections, *IEEE Computer Graphics and Applications* 38 (2018) 91–108. URL: <https://ieeexplore.ieee.org/document/8059795/>. doi:10.1109/MCG.2017.377152546.
- [10] S. M. Robertson, A. Mullen, Tropy: A Tool for Research Photo Management, 2017.
- [11] D. Oldman, D. Tanase, Reshaping the Knowledge Graph by Connecting Researchers, Data and

