

Progressing Picasso

Neal Audenaert, Alex Maslov, Adam Mikeal, Scott Phillips

{neal, alexey, adam, scott}@cs.tamu.edu

Texas A&M University

College Station, TX 77843-3112, USA

ABSTRACT

Prior work from the library sciences and art history perspectives has focused on developing schemes for indexing images and then searching those indices. A major component of this work focuses on providing metadata that map onto the needs and interpretive perspectives of different audiences. Scholars, educators, and students interact with large image collections (both digitally and in print), in complex and varied ways. In a series of informal interviews with faculty members from art education, history, Hispanic studies, art history, and with local K-12 art teachers we found a diverse set of needs, interests, and approaches to working with artworks in both teaching and research settings. Based on these investigations, we have approached the challenge of providing subject level access to scholarly art collections from a different perspective, namely, representing and visualizing the relationships in which artworks participate.

Categories and Subject Descriptors

J.5 [Arts and Humanities]: Literature.

Keywords

Online Picasso Project, humanities research, relationship expression, visualizations.

1. INTRODUCTION

For centuries humans have been creating and studying art. The first known rock paintings are dated from 40,000 years ago, in the upper Paleolithic. While humans have been analyzing and appreciating art for a long time, it is unknown when the first scholars started researching art as a subject in its own right. Certainly, though, by the times of the Greeks, the field was well developed. The scholarly model that has used to research this subject has received little to no change since its original inception.

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike License.

*To view a copy of this license, visit
<http://creativecommons.org/licenses/by-nc-sa/2.0/>
or send a letter to:*

*Creative Commons, 559 Nathan Abbott Way,
Stanford, CA 94305, USA*

Researchers and scholars ask questions such as:

- “How did the artist come to create his work?”
- “Who were his patrons? his teachers? his disciples?”
- “What historical forces shaped the artist?”
- “What impact did the artist produce?”
- “What is the artistic value of the work?”

Many of these questions can be answered with a digital archive. Beyond the noteworthy improvements of access speed and efficiency, digital technology can also be used to aid researchers, teachers, and students in discovery and understanding. By recording, calculating, and visualizing both explicit and implicit relationships, scholars can better understand existing art. They will be able to perform their analysis faster and with more access to quality information, allowing the production of better scholarship.

1.1 Background

Prior work from the library sciences and art history perspectives has focused on developing schemes for indexing images and then searching those indices. A major component of this work focuses on providing metadata that map onto the needs and interpretive perspectives of different audiences. Computational approaches have focused automatically recognizing image content.

To date, this work centers on the recognition of color, gradients, shapes, and textures. While this work has the potential to support image-based query interfaces and the clustering of visually similar images, it has yet to yield results that are adequate to support interactions with large image database based on the conceptual content of the images.

In this paper, we present a Web-based application, the Interactive Relationship Viewer (IRV) that supports visualization of and interaction with the relationships between artworks in a large image database. This work is based on our work with the *Picasso Project*.

The *Picasso Project* maintains a digital *catalogue raisonné*, or reasoned catalogue, containing more than digital 11,000 images of artworks created by Pablo Picasso (1881-1973). Descriptive metadata is provided for each artwork along with commentary noting sales of the item, exhibitions in

which it has been displayed, and scholarly literature in which it has been cited. The catalogue also includes more than 7500 biographical entries linked to artworks, photographs of key people, maps of places Picasso lived and worked, and related documents (for example, Picasso's birth certificate).

The catalogue provides tools for side-by-side comparisons of artworks and to build custom sub-collections from which an illustrated, color catalogue can be automatically generated for printing. This catalogue is currently the most complete and up-to-date catalogue of Picasso's work and new artworks are being added at a rate of approximately 2,000 per year.

1.2 User Identification

Picasso's works cover a broad range of themes, topics, and materials, thus presenting a rich substrate of artworks for building a network of semantically diverse, meaningful relationships. In addition to the image collection, the Picasso project includes extensive metadata related to these works, such as its place and date of creation, medium, dimensions, current location, as well as exhibitions and books in which it has appeared. We leverage much of this metadata to express relationships based on ownership, materials, patronage, or chronology.

Scholars, educators, and students interact with large image collections (both digitally and in print), in complex and varied ways. They analyze and critique individual works and compare and contrast these with other works. They identify similarities between pieces of art and trace threads of influence between artworks, artists, styles, materials, themes, and social, geo-political, or personal events. These scholars interpret artworks, identify missing links, and communicate their findings. A common thread in these tasks is the need to interact with artworks not as isolated artifacts, but rather in the artistic, cultural, and historical context in which they are embedded. In this context, artworks participate in multifaceted relationships with other artworks, historical people, places, and events, and with abstract concepts and themes.

In a series of informal interviews with faculty members from art education, history, Hispanic studies, art history, and with local K-12 art teachers we found a diverse set of needs, interests, and approaches to working with artworks in both teaching and research settings. One key theme running through each of these areas is the need to discover and present relationships between artworks, although the specific relationships of interest varied by discipline.

K-12 teachers are interested in identifying artworks that provide good examples of specific drawing or painting

techniques, such as the two-point perspective or the use of complementary color schemes. The art historian wishes to investigate relationships between artworks displayed together in an exhibition or to study works composed when an artist was with a particular lover. The historian wishes to view art in the context of significant historical events, for example, artworks created while Europe was anticipating World War I.

In addition to the interviews, we also attended sessions of two college-level art history survey courses. We observed that instructors typically showed one or two examples of artworks from different artists or art movements, discussing each for a few minutes. In subsequent interviews, the instructors explained that lack of time constrains their ability to include additional works. Creating thematic sub-collections based on the relationships discussed in class could alleviate this problem, enabling students to study additional examples of materials covered in the classroom. These observations of classroom interaction and feedback from educators and researchers have informed our enhancements for supporting the representation and visualization of diverse relationships.

Based on these investigations, we have approached the challenge of providing subject level access to scholarly art collections from a different perspective, namely, representing and visualizing the relationships in which artworks participate.

2. SYSTEM DESIGN

The Interactive Relationship Visualizer (IRV), an interactive, Web-based application, enables visualization of relationships. In addition to the displayed relationship, the interface presents the other relationships existing within the sub-collection as well as elements of the sub-collection and other artworks in the archive. This view enables users to navigate the intricately interconnected hypertextual web defined by these relationships.

The display of relationships changes to reflect the primacy of the relationship that has the user's attention. In order to express a rich set of relationships, we are augmenting existing metadata to include type (such as still life or portrait), art movement (cubism, fauvism, surrealism), and content (woman, nude, vase, mirror).

We determined that RDF (Resource Description Framework), a set XML schemas and related technologies developed for the semantic web effort, were the ideal way to express these relationships. The basic mode of expression in RDF is the *triple* consisting of subject, predicate, and object.

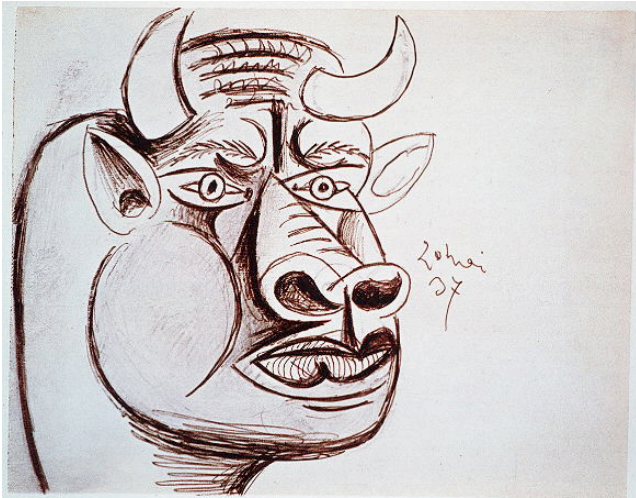


Figure 1: A study for Guernica

For example, a particular artwork might be claimed by art historians to be a preparatory work for a later painting. In RDF, this relationship would be expressed by making the artwork in question the subject, the predicate would be *isStudyFor*, and the object would be the later painting. In this way, both the standard descriptive metadata, as well as the more complex inter-work relationships can be natively described in RDF.

3. RELATIONSHIPS

The IRV distinguishes two broad categories of relationships, *inferred* and *specified*. Inferred relationships are those which can be expressed in terms of the metadata elements provided for each artwork. Some inferred relationships can be expressed in terms of a single metadata value, such as “artworks created with oil on canvas”. Others require mapping a range of metadata values onto a higher-level concept and require definitions involving multiple metadata fields.

For example, identifying “paintings created in Paris around the time of World War II” is a two-step process. The system must map the timeframe of World War II to a portion of the traditional calendar and locate paintings created during this time. It then selects from this set, those that were created in Paris. Finally, relationships such as “expensive paintings” involve subjective, theory-driven, and potentially variable definitions. A price that would be expensive in one art market might be comparatively inexpensive in another. Inferred relationships provide a powerful mechanism for exploring, discovering, and expressing relationships between artworks that leverages existing metadata.

In contrast to relationships inferred directly from existing metadata, other types of relationships must explicitly be stated. We refer to these as *specified* relationships. For example, Picasso sketched several rough drafts of large

works, interspersed with smaller works. Thus, a chronological (or timeline-based) view of artwork around the time the *Guernica* was painted results in a sub-collection that would include these preparatory works, as well as other, unrelated works. Hence, this “preparatory work” relationship must be expressed between early sketches of the *Guernica* and the final masterpiece.

Another specified relationship is images based on a shared subject, for example, Picasso’s interpretive series of works of Diego Velázquez’s *Las Meninas*. Specified relationships afford us the ability to define and represent relationships between artworks that are difficult to derive from the descriptive metadata associated with each work.

This category of relationships is critical for the expression of concepts based in established and novel analytical approaches to Picasso’s work—allowing relationships based on information beyond that which is encoded in the collection. The drawback is that participation in specified relationships must be encoded manually.

4. ARCHITECTURE

Figure 2 depicts the IRV system architecture. The Web interface employs specific visualizations for displaying different kinds of relationships. For example, the display of *Guernica* and its preparatory images uses a visualization that illustrates the centrality of *Guernica* relative to the other images displayed.

In contrast, the display of all artworks in the *Las Meninas* series uses a table-like view, since no image is clearly central to this sub-collection. We employ artwork images from the Picasso project, reinterpret and extend existing metadata to express myriad connections between these artworks, and facilitate visualization of these relationships to support art scholars from various disciplines.

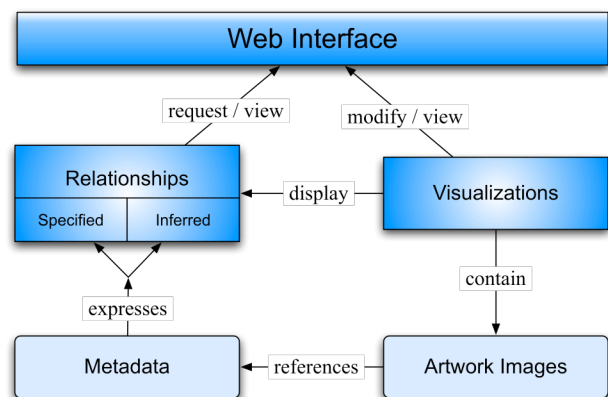


Figure 2: A representation of the Interactive Relationship Visualizer (IRV) system architecture

5. VISUALIZATIONS

There are many visualization possibilities that can be employed to depict the basic relationship types described previously. Three of the most commonly encountered patterns are:

- **Graph-based:** In the broad sense, a mesh network consisting of nodes that represent each subject and object in the RDF triple store, and vertices connecting the nodes that represent the predicates. A standard hierarchical tree visualization is a form of an acyclical, single-parent network mesh.
- **Timeline-based:** Linear representations of RDF subjects and objects where the spatial presentation of those nodes has been layered with an additional facet of information (typically date-based events or periods).
- **Set-based:** Involves visually grouping subjects with matching objects based on predicate relationships. For example, a rudimentary form of this pattern would be a Venn diagram that groups subjects across three predicate-object relationships.

Because of our decision to employ RDF as the basic storage and expression mechanism for our relationship data, there were several “ready-made” visualization products available to us. The SIMILE project at MIT produces several of these “RDF triple” visualizers; Welkin is a stand-alone Java-based application that utilizes the graph-based approach, and Longwell is a faceted RDF browser that is deployed as a web app using any modern web browser.

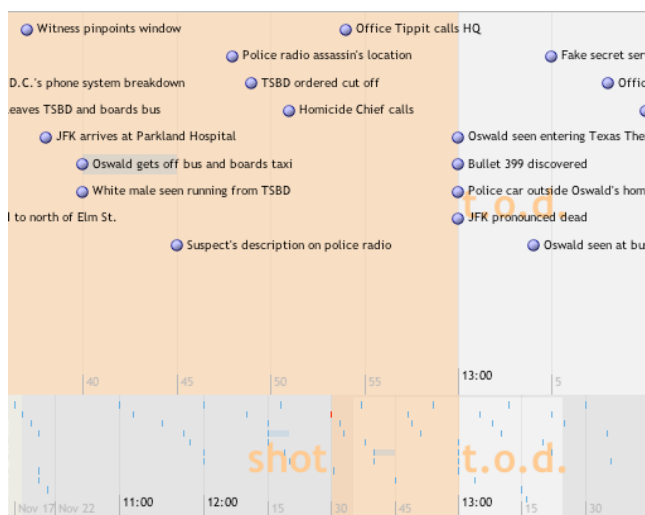


Figure 3: A depiction of a timeline-based visualization from the SIMILE project

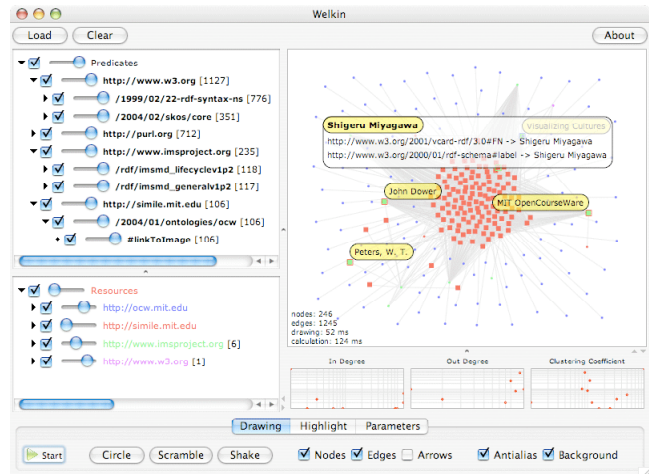


Figure 4: A depiction of the Welkin RDF visualizer from MIT

For the short-term duration of this class project, we opted to experiment with one of these pre-existing products, since building a new visualizer was not feasible within the time constraints of a semester. We were able to successfully interface our RDF data with Welkin, and able to exhaustively explore the limited sub-set of relationships we hand-selected for the development cycle.

Figure 3 depicts a possible timeline-based visualization pattern using the Timeline application from MIT’s SIMILE project, while Figure 4 shows a typical visualization produced by the Welkin RDF visualizer; predicates and subjects are given automatically-generated graphical control parameters which allow for the interactive filtering and adjustment of the RDF triples being currently displayed.

6. CONCLUSIONS

The single most time-consuming aspect of this project was the identification and analysis of the various user communities engaged in scholarly activities (both research and teaching) utilizing digital artworks collections. Our findings were perhaps not surprising to anyone involved in digital humanities research: there were a wide range of disparate user groups, each with distinct and individual needs for supporting their scholarly activities. A partial delineation of these user communities includes:

- **Students:** Students are frequently given assignments by their teachers to engage them further with the particular topic at hand. These assignments are designed to elicit exploration and analysis by the student. Their needs include the ability to discover relationships between artists and artworks outside those specifically identified by their teacher.

- **Educators:** Several of the educators with which we spoke mentioned the desire to provide their students with in-depth tools that would allow for further exposure that is possible in a traditional classroom environment. Additionally, the ability to understand and explore a collection to discover less-commonly used artworks as classroom examples was seen as a potential benefit of a digital collection.
- **Researchers:** The ability to understand and place artists and their works in the greater context of society and current thought is of paramount importance to these users. Systems that would allow for unrestrained discovery and social collaboration with colleagues would benefit this group.

Our approach to formally classifying and visualizing the relationships between artworks—both explicit and inferred—will aid in these tasks, and allow for the use of digital art collection in new ways. Discovery and analysis of the collection and its components will be enhanced by leveraging this explicit encoding and visualization of the relationships in which those artworks participate.

7. FUTURE WORK

We continue to add new metadata to enrich the relationships expressed in our archive. While new attributes enable us to express additional relationships, the growing number of relationships gets increasingly difficult to represent visually. We are investigating mechanisms to display secondary visualizations without overwhelming the presentation of the primary relationship views. As scholars analyze Picasso's works and life, the relationships of their interest are likely to increase in complexity as well as variety. It is not possible to express all possible relationships among these artworks a priori, nor is it possible to have all the necessary metadata. Enabling scholars to define useful metadata as well as supporting

them in forming new relationships will engage them as partners in this project rather than as mere users.

The IRV has potential for exploration of artwork relationships in the classroom as well as for evaluating student performance via homework and papers. For example, students could explore a relationship and write a short paper about the artworks it encompasses. An instructor could create a relationship and ask students to identify the relationship embodied by the included artworks. Educators need assistance in the form of targeted features for successful use of the IRV in the classroom setting. We continue our dialog with instructors to channel the IRV's expressive power for enriching education.

8. REFERENCES

- [1] Gladney, H. M., et al. (1998). 'Digital access to antiquities'. In *Communications of the ACM*. 41(4): 49-57.
- [2] Lanzelotte, R.S.G., et al. (1993). 'The PROTINARI project: science and art team up together to help cultural projects'. In *Proceedings of 2nd International Conference on Hypermedia and Interactivity in Museums (ICHIM'93)*, Cambridge, UK 1993.
- [3] Mallen, E., ed. (2006) *The Picasso Project*. Texas A&M University <http://picasso.tamu.edu/> [6 November 2006]
- [4] McGann, J. (1997). 'The Rational of Hypertext'. In Sutherland, K. (ed.) *Electronic Text: Investigations in Method and Theory*. New York: Oxford UP. 19-46.
- [5] Gemini G.E.L.: Online *Catalogue Raisonné*. Washington DC: National Gallery of Art <http://www.nga.gov/gemini/> [7 Sept 2006]
- [6] Resource Description Framework (RDF). <http://www.w3.org/RDF/>
- [7] The SIMILE Project. <http://simile.mit.edu/>
- [8] The Vincent van Gogh Gallery. <http://vggallery.com/>