

# Methods of Personalising a Collection of Images using Linking Annotations

Maristella Agosti, Giordana Mariani Canova, Nicola Orio, and Chiara Ponchia

University of Padua, Italy

agosti@dei.unipd.it, giordana.mariani.canova@unipd.it, orio@dei.unipd.it,  
ponchia@dei.unipd.it

**Abstract.** This paper reports the results of an ongoing project that deals with the development of a digital archive of drawings and illustrations of historic documents for research purposes, but also as a tool that can be used to disclose cultural heritage contents to the general public. The digital archive is currently used as a case study to provide innovative tools for researchers and scholars active in the preservation and dissemination of cultural heritage. In particular, we investigate the use of linking annotations to express the relations between images. We represent annotations with a hypergraph structure and we exploit this information to personalize the way images are presented to the final users.

## 1 Introduction

The starting point of the work reported in this paper is the digital archive of illuminated manuscripts that has been developed over the past decade within our research activities. This archive is called *IPSA*, which stands for *Imaginum Patavinae Scientiae Archivum* (archive of images of the Paduan science) [1,2]. The main focus of our initial project was to provide a tool for the analysis of the role played by the Paduan school during the Middle Ages and the Renaissance in spreading the new scientific method in different sciences, from medicine to astronomy to botany. The aims and the results of *IPSA* are well documented in the website that has been developed to report on the project<sup>1</sup>.

*IPSA* was designed and made available as a research tool for experts of the history of art and it provides advanced functions for studying illuminated manuscripts. In particular two functions are worth mentioning as they are meaningful examples of the way those experts operate: i) the possibility of recalling all images with a specific subject to study the different ways of its representation over the centuries, ii) the possibility of recalling different groups of images to compare them through images comparison, because art historians are able to understand the kind of relation existing between images by personally inspecting them.

The challenge that we now face is the envisaging of a new type of system able to open up its contents to the complete spectrum of users, ranging from

---

<sup>1</sup> URL: <http://www.ipsa-project.org/>

professional researchers or experts to members of the interested general public. To envisage such a system, a new collection of user requirements has started to be collected together with the identification of a set of relevant user questions.

## 2 IPSA Contents

IPSA is a digital archive of illuminated images taken from a series of manuscripts which were produced mainly in Padua and in the Veneto region during the XIV and the XV centuries. The archive includes 56 manuscripts belonging to some of the most important libraries in Europe and in the world, including: the British Library<sup>2</sup>, the Bodleian Library<sup>3</sup>, the Pierpont Morgan Library<sup>4</sup>, the *Biblioteca Marciana*<sup>5</sup>, the *Biblioteca Apostolica Vaticana*<sup>6</sup>, and the *Biblioteca Medicea Laurenziana*<sup>7</sup>.

All manuscripts included in IPSA are scientific books, of which 46 are astrological codexes and 10 botanical codexes. The digital archive includes about 3400 images. The entire collection has the merit of testifying the shaping of a new scientific mentality in the University of Padua during the XIV century. The collection has been digitised and marked up with key metadata information through the efforts of several research groups. The metadata include information on the content and the provenance of the digital images.

For this new challenge of opening up the archive to a new type of user, including members of the general public, it has been decided to concentrate attention in the first step of this new effort on the botanical codexes or herbal manuscripts. In fact, herbals are manuscripts which contain hand-drawn depictions of plants, such as trees, bushes or shrubs, and their parts, such as flowers or leaves. The IPSA collection contains herbals written and illustrated by the Paduan School, and successive herbals produced in Europe under its influence. Such manuscripts have the characteristic of containing high quality and very realistic botanical illustrations directly drawn from nature. The realism used in drawing the illustrations of plants makes it possible for members of the general public to use the illustrations more easily, by allowing them to link each illustration to one or more recent images or photographs of the plant, or to the description of its use in the past and in the present, and by allowing them to share this knowledge with other final users through the use of a network of links of comments and annotations.

A relevant example of the botanical codexes present in IPSA is the *Liber agreg de Serapiom* (London, British Library, Egerton 2020), which was made for Francesco II da Carrara, ruler of Padua (1390-1404), as testified by the presence of his coat of arms in two pages (ff. 4r, 267r). This is a translation of a botanical

---

<sup>2</sup> URL: <http://www.bl.uk/catalogues/manuscripts/INDEX.asp>

<sup>3</sup> URL: <http://www.bodleian.ox.ac.uk/dept/scwmss/wmss/medieval/>

<sup>4</sup> URL: <http://www.themorgan.org/collections/collectionsLitManu.asp>

<sup>5</sup> URL: <http://marciana.venezia.sbn.it/catalogazione.php?sst=40>

<sup>6</sup> URL: <http://193.43.102.72/gui/html/index.jsp>

<sup>7</sup> URL: <http://www.bml.firenze.sbn.it/index.htm>



**Fig. 1.** London, British Library, Egerton 2020, f. 27v, Grapevine.

medicine treatise that describes many different Mediterranean plants with their medical utilization and healing properties. So the users interested in the use of medical plants can gain interesting information and can decide to relate them to recent ones of interest. In fact the images of this manuscript depict the plants in a realistic way for the first time in the Middle Ages, because the illustrations are made by copying the plants from a real model and not from a previous painted one. This approach helps in recognising the plant also for the purpose of picking its fruits or leaves to make a drug. This new quest for realism in searching for objectivity was a distinctive feature of the University of Padua in the XIV century, especially in the medical school.

Figure 1 shows with great detail and realism some parts of a grapevine, while Figure 2 shows a part of a peach tree. Figure 3 is a screenshot of the present version of the IPSA Web application where the image of the grapevine is shown.

### 3 Linking Annotations

The analysis of user requirements that was conducted for the design of IPSA pointed out the need of tools for adding information about the relations between objects. Adding information can be achieved through the use of an annotation system built over an existing image digital archive. The main service provided by the annotation system consists of a set of tools for image annotation.

Annotations are an effective means to enable interaction between users and a digital archive, since their use is a diffuse and very well-established practice.

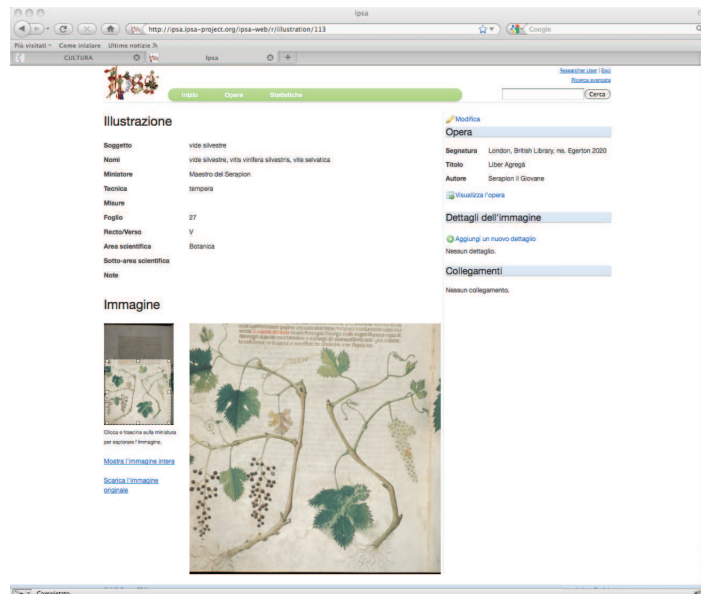


Fig. 2. London, British Library, Egerton 2020, f. 166v, Peach tree.

Annotations are not only a way of explaining and enriching an information resource with personal observations, they are also a means of transmitting and sharing ideas to improve collaborative work practices [3]. Furthermore, annotations allow users to merge and link personal contents with the information resources provided by a digital archive. In [4] a taxonomy has been proposed for linking annotations, which is divided in two classes. The first link class reflects a *hierarchical* relation between two images, where one image depends on an earlier one. According to user requirements, there are three typologies of hierarchical relations between two images, which are expressed by the following types:

- **A has\_progenitor\_in B**: image **B** is the first exemplar of a given representation of an object, from which a number of images, including **A**, descend as direct copies or as copies of intermediate representations.
- **A is\_copy\_of B**: the author of **A** used image **B** as his direct source of inspiration, both from the stylistic and from the pictorial point of view. **B** can either be the progenitor of **A** or an intermediate representation between **A** and its progenitor.
- **A is\_elaboration\_of B**: image **A** has been inspired by **B**, but there are a number of differences between the two that show that the author of **A** added personal changes to the original representation.

The second link class reflects a *relatedness* relation between two images, because they share similar properties even though they have been created independently. Also in this case, the system automatically adds a symmetric annotation,



**Fig. 3.** Screenshot of the IPSA Web application where the grapevine is depicted.

in order to emphasize the relation in both directions. According to researchers, there are three motivations by which two images can be related, even if there is no hierarchical relation between them:

- **A has same model of B**: images **A** and **B** descend from the same progenitor, even though they have been independently created by their authors. This kind of link tells us that **A** and **B** belong to two chains of derivations, which share the same progenitor.
- **A is similar to B**: although the two images are visually similar, it is not possible to draw a hierarchical relation between the two or to state that they descend from the same model. The existence of this link may help researchers and students to discard apparent hierarchical relations and it has been proved that they do not exist.
- **A is connected to B**: sometimes a possible connection is not completely clear from the images themselves. Instead we need to exploit additional information, such as knowledge on authors background and connections or their belonging to similar schools of scientific representation. This kind of link, which express a generic relation, helps to point out that the relation between **A** and **B** needs further investigation.

The presence of a hierarchical structure can be exploited by automatically discovering *paths* across the hypertext given by the link structure. For instance, paths may describe the creation over the years of copies of a renowned author, and the subsequent creation of copies of the copies and so on. At the same

time, paths can describe the dissemination of a particular approach to scientific representation over time and space. On the other hand, relatedness links may be useful for disclosing similarities between images, because together with hierarchical links they may introduce a notion of distance between each pair of images. Researchers may discover new similarities because two images, even if not directly connected, are very close in the graph structure induced by links.

The proposed taxonomy has been derived from the conducted analysis and design of users requirements. The effort involved in this research has been to identify the primitive types of link necessary to make all the relations that are of interest for the researchers and students explicit. In particular, it has been considered that the annotation of a digital archive is an ongoing process, which depends on the results of scientific research, and the link structure can be continuously modified by a researcher. This taxonomy is sufficient to build a directed graph structure on the digital archive which can be used to discover the chains of derivation among images. A positive side effect of this design choice is that a reduced amount of link typologies helps researchers to be consistent in their use of link types.

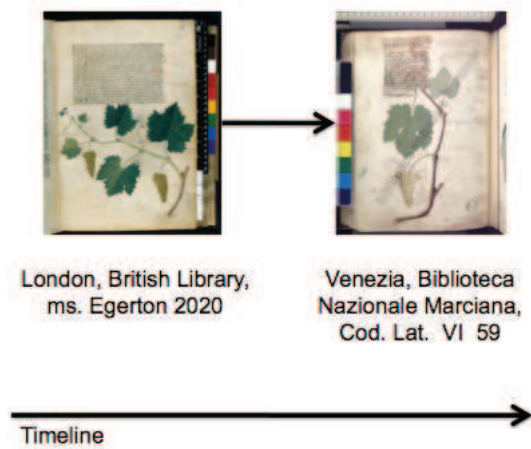
#### **4 Personalized User Views**

After a user has added his personal knowledge on the application domain by drawing typed links between images, the digital archive is dynamically enriched by a hypertextual structure. Each researcher may then access the digital archive in different ways, because each user may have his own view on the archive, which is given by the hypertextual structure. The existence of such a structure can integrate direct search through navigation inside the image collection. If a mechanism of information sharing inside a user group is provided, as the ones applied to collaborative environments, researchers may cooperate in their study on illuminated manuscripts by sharing the information on relations between different sources. The annotations can be private, shared among a group of collaborators or public.

As an example, let us suppose that a group of users is formed and it also contains an art historian and a botanist. The botanist may have identified a relation between two different images of plants, because the leaves drawn on one image are not realistic since they are too similar to those of another plant. The art historian can use this information to disclose a new historical path connecting the two images. Moreover, motivation of links between images needs to be made explicit for all users, and thus depends on the expertise of the researcher. The final goal is to highlight the mental process of the researcher who previously highlighted a relation.

#### **5 The Work Under Development**

As already mentioned, IPSA has already been released and used as a research tool by scholars in the history of art in our team. Work on its dissemination



**Fig. 4.** Two representations of *vitis vinifera* (Common Grape Vine) over time.

has begun in order to present it to other categories of users with the purpose of disclosing its contents not only to experts.

The previously proposed taxonomy of linking annotations was proposed after the study of requirements of professional users, whereas the requirements analysis that is presently under way aims to capture the desires of final users. While the analysis is still under way, initial clues indicate that the taxonomy of interest for the final users can be simplified, in fact the final user often does not have the knowledge that can support him in expressing advanced link types. Instead, would need to be found to new ways of automatically building links among images prepare paths that could be of use for the user to browse and acquire information from the digital archive. One link that can be automatically detected and inserted between two images is a link that expresses a simplified version of the `has_progenitor_in` type of link connecting two images that represent the same subject but that have been done in different moments over the timeline. Figure 4 shows two representations of *vitis vinifera* (Common Grape Vine) that were produced in manuscripts compiled in two different periods of time: the first in 1390-1404 and the second in 1445-1448, so the images are represented in chronological order.

The same type of link can be used to automatically link images of manuscripts to present days photographs that depict the same subject. Figure 5 shows the link to a photograph of a common grape vine plant. Furthermore, a different interpretation of the same type of link can be used to link the photograph to a document or a piece of text that describes what is represented in the photograph, thus allowing the user to relate the old image of a plant to the illustration of the characteristics of the present day plant.

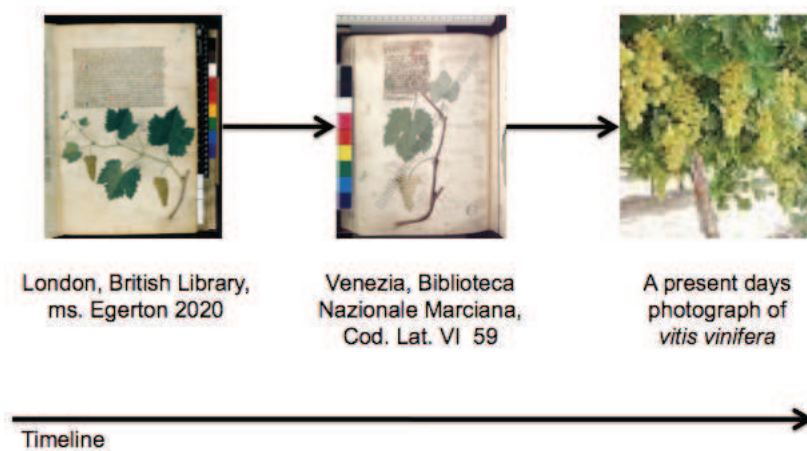


Fig. 5. Two representations of *vitis vinifera* (Common Grape Vine) over time.

## Acknowledgements

The work reported has been partially supported by the CULTURA project<sup>8</sup>, as part of the Seventh Framework Programme of the European Commission, Area “Digital Libraries and Digital Preservation” (ICT-2009.4.1), grant agreement no. 269973, and by the EuropeanaConnect Best Practice Network funded by the European Commission within the area of Digital Libraries of the eContentplus Programme (Contract ECP-2008-DILI-52800).

## References

1. Agosti, M., Ferro, N., Orio, N.: Graph-based Automatic Suggestion of Relationships among Images of Illuminated Manuscripts. In Haddad, H., ed.: SAC, ACM (2006) 1063–1067
2. Mariani Canova, G.: Hyginus De Astronomia. In: The Splendor of the Word. Medieval and Renaissance Illuminated Manuscripts at the New York Public Library. Harvey Miller, New York, USA (2006) 337–339
3. Agosti, M., Ferro, N.: A Formal Model of Annotations of Digital Content. ACM Transactions on Information Systems (TOIS) **26** (2008) 3–57
4. Agosti, M., Ferro, N., Orio, N.: Annotating Illuminated Manuscripts: an Effective Tool for Research and Education. In Marlino, M., Sumner, T., Shipman III, F.M., eds.: Proc. 5th ACM/IEEE Joint Conference on Digital Libraries, (JCDL 2005), ACM Press, New York, USA (2005) 121–130

<sup>8</sup> CULTURA Project Website, URL: <http://www.cultura-strep.eu/>