

EuroMed2010

# Digital Heritage

3<sup>rd</sup> International Conference dedicated on  
Digital Heritage

Short Papers



*Before July, 1974...*



*After July, 1974...*

The Katholikon of the Monastery of Antiphonitis in Kalogrea, Cyprus (16<sup>th</sup> century)

**DON'T MISS THE *Hi-tech*-STORY...**

8<sup>th</sup>–13<sup>th</sup> of November 2010  
Limassol, Cyprus

Edited by

M. Ioannides, D. Fellner, A. Georgopoulos, D. Hadjimitsis

**M. Ioannides, D. Fellner, A. Georgopoulos, D. Hadjimitsis (Eds)**

*Marinos Ioannides*  
Editor-in-Chief

*Elizabeth Jerem*  
Managing Editor

*Elizabeth Jerem*  
Copy Editor

*ARCHAEOLINGUA*  
Cover Design

*Carol Usher, Gergely Hős*  
Desktop editing and layout

***Front and Back Cover Images:***

Christ *Antifonitis* near *Kalogrea* in the Kyrenia district in Cyprus is another church, which was built and decorated with wall paintings in approximately 1200 AD. In the early 16<sup>th</sup> century, the north and south wall of the church proper and the northwestern squinch were decorated with wall paintings from the life of the Virgin, the Stem of Jesse, the Last Judgment and the standing saints, while the cupola was decorated with Christ Pantocrator, the Deesis-Supplication, the Apostles seated on thrones and the Prophets. Unfortunately, after the Turkish invasion in 1974 and the occupation of the area by the Turkish army, the systematic destruction and removal of the wall paintings from the monument began. The faces of Archangels Michael and Gabriel in the conch of the apse were barbarically destroyed, while the wall painting of the Birth of Christ was removed. These wall paintings are dated from the end of the 12<sup>th</sup> or early in the 13<sup>th</sup> century. The great wall paintings of the Stem of Jesse and the Last Judgment (early 16<sup>th</sup> century) were fragmented into small pieces and removed from the south and the north wall on which they were painted. Several *fragments from the wall paintings (in yellow outlined)*, which had been exported by Turkish smugglers to Germany from where they were repatriated, are on display on the second hall of the Byzantine-Museum (BM.321-356) in Nicosia, Cyprus. Others are in Germany and it is hoped they will be returned to Cyprus. When all the pieces which have been found are gathered, an effort will be made to restore these two large wall paintings using different modern applications from the area of Digital Heritage.

Photos: Eliades Ioannis, Curator, Byzantine Museum in Nicosia, Cyprus

This work is subject to copyright.

Permission to make digital or hard copies of portions of this work for personal or classroom use is granted without fee, provided that the copies are not made or distributed for profit or commercial advantage and that the copies bear this notice and the full citation on the first page. Copyright for components of this work owned by others must be honored. Abstracting with credit is permitted. To otherwise reproduce or transmit in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage retrieval system or in any other way requires written permission from the publisher.

© 2010 by the individual Authors and Archaeolingua Foundation

ISBN 978-963-9911-16-1

Published by ARCHAEOLINGUA  
Printed in Hungary by PRIMERATE  
Budapest 2010



**ARCHAEOLINGUA**

## DISCOVERY AND USE OF ART IMAGES ON THE WEB: AN OVERVIEW

K. Ivanova <sup>a\*</sup>, M. Dobreva <sup>b</sup>, P. Stanchev <sup>a c</sup>, K. Vanhoof <sup>d</sup>

<sup>a</sup> Institute of Mathematics and Informatics, BAS, Sofia, Bulgaria – kivanova@math.bas.bg

<sup>b</sup> Centre for Digital Libraries Research, University of Strathclyde, Glasgow, UK – milena.dobreva@strath.ac.uk

<sup>c</sup> Kettering University, Flint, MI, 48504, USA – pstanche@kettering.edu

<sup>d</sup> Hasselt University, Hasselt, Belgium – koen.vanhoof@uhasselt.be

**KEY WORDS:** Art Images, Repositories, Image Retrieval, Content-Based Indexing, User Expectations, Interoperability, Europeana

### ABSTRACT:

This paper provides an overview of the area of resource discovery in the case of art images. Digitized art images are an important part of the online cultural heritage. Digitization and online accessibility were mainly materials focused on by libraries, museums, archives and audio-visual archives in the last decade. However, digitised art images as well as digitally-born art provide a major challenge for resource discovery and use. Better results in resource discovery could be achieved through methods combining queries executed on metadata with image retrieval methods. The paper gives an insight into the evolution of designated repositories for digital art, and then briefly presents some of the basic methods for intelligent resource discovery in this domain. Finally it looks at directions for further development within the digital repositories of art, which would enrich even more the user experiences.

### 1. INTRODUCTION

A key characteristic of visual arts' objects is that they are created by a cognitive process. The artwork is not merely an objective presentation, but also communicates one or more subjective messages "delivered" from the creator to the observer. Every touch to the artwork helps to build bridges between cultures and times. As suggested in (Chen et al, 2005) "research on significant cultural and historical materials is important not only for preserving them but for preserving an interest in and respect for them".

Since its first publication in 1962 Janson's History of Art (Janson, 2004) is one of the most valuable sources of information spanning the spectrum of Western art history from the Stone Age to the 20<sup>th</sup> century. It became a prominent introduction to art for children and a reference tool for adults trying to recall the identity of some familiar images. The colourful design and numerous illustrations of exceptional quality are far from being a means of providing dry information; they also contribute to experience a deep emotional fulfilment. But now online search engines have whetted web surfers' appetites for context and information, there are a host of digital databases and repositories offering easy access to digital items, presenting the colourfulness of art history as well as to connected metadata, giving all additional information from pure technical details, connected with the way of creating the artefacts, to deep personal details from the lives of the creators, which help the observers to understand input message in the masterpieces.

In Section 2 we look into the evolution of such systems and provide some examples. Then in Section 3 we focus on the specifics of access to art images. In section 4 we outline some of the influences of the Web on the organization of art resources in digital form. Finally, we provide some conclusions on the issues which are most challenging in this domain.

### 2. DIGITAL REPOSITORIES OF ARTWORKS

The digital repositories of cultural heritage objects can employ similar techniques as generic ones in order to provide standard functionality for searching objects. Cultural heritage objects are rich in content describing events, monuments, places, people; they are distributed across different locations. The users can formulate queries using different modalities such as free text, similarity matching, or metadata; one important current trend is the use of linked data (Gradmann, 2010). The specifics of observed objects profile some additional tasks, which are interesting to be considered. In the area of art paintings retrieval the sensitive, semantic and aesthetic gaps are a prevailing problem.

In 2004 David Mattison, named as a master of the online archive universe, published a series of lectures in the Searcher magazine focusing on the state-of-the-art of available Web resources and image databases, current techniques for image retrieval, and finally mentioning national collections that document the art history of Western civilization from medieval times to the 19<sup>th</sup> century (Mattison, 2004). The creators of this set of image databases, art collections and guides usually are academic, librarian, commercial, and private art museums and galleries, amateur and professional art historians, artist sites, commercial image agencies, auction houses (usually on a temporary basis), etc.

Several different applications in the field of Fine Arts have led to specialized digital image processing developments. During the years numerous successful projects in the field of processing very large high quality colour images had been funded by Esprit, Impact, Raphael, and IST programs of the European Commission. They provide the most conventional image processing for the museum, such as geometric correction, registration, mosaicing, etc. (Maitre et al, 2001). The European Union has funded a number of digital culture research and development initiatives.

Some of the projects, such as *Vasari* (1989-1992) and *Marc* (1995-1996), are focused on digital acquisition, storage and

---

\* Corresponding author.

handling of colorimetric high-definition images of paintings (up to 2GB per image) for different galleries and museums in Western Europe. The *Crisatel* project (2001-2004) developed equipment for the direct fast capture of paintings, with a new ultra-high definition multi-spectral scanner in order to make spectrometric analysis of varnish layers to allow the effect of an aged varnish to be subtracted from an image of a painting. The project *FingArtPrint* (2005-2008) aimed to make combinations of 3D surface scanning and multispectral imaging in order to create a unique data record of the object which can be compared to check its authenticity.

Several other projects and initiatives pursued the creation and population of art image repositories. One of the first projects in this direction was *NARCISSE* (1990-1992), which has created a very high digitized image bank, supervised by a multilingual text database. The objective of the project *Artiste* (2000-2002) was to develop and to prove the value of an integrated art analysis and navigation environment aimed at supporting the work of professional users in the fine arts. The environment has exploited advanced image content analysis techniques, distributed hyperlink-based navigation methods, and object relational database technologies using existing metadata standards and indexing schemes.

In recent years several projects and initiatives focused on the harmonization of activities carried out in digitization of cultural and scientific content for creating federated platforms for storage of and access to cultural heritage. Such project is *MINERVA+*, sponsored by 6FP, which intends to enlarge the existing thematic network of European Ministries in this direction. Since 2005 the Netherlands Organization for Scientific Research supports the research program *CATCH* (Continuous Access to Cultural Heritage) that finances teams which concentrate on improving the cross-fertilization between scientific research and cultural heritage. In the light of transferability and interoperability, the research teams execute their research at the heritage institutions.

The latest and biggest project in this domain is *Europeana* (<http://www.europeana.eu/portal/>), funded by the European Commission and the member states. The idea of *Europeana* was born in 2005, when the European Commission announced its strategy to promote and support the creation of a European digital library as a strategic goal within the European Information Society i2010 Initiative, which intends to foster growth and jobs in the information society and media industries. The European Commission's goal for *Europeana* is to make European information resources more available in an online environment. It will build on Europe's rich heritage, combining multicultural and multilingual environments with technological advances and new business models. Currently over 10 million digital items are available through the portal.

Some of very successful worldwide digital repositories are created independently as a result of collaboration between different stakeholders, most often state or public institutions; or by supporting concrete companies. Usually the funding for establishing and supporting such initiatives is received by donations; by sparing out of the budgets of the institutions; or by allocating funds from the profit of the companies gained with other activities (for instance Artchive is supported by art.com, which is a company that sells digital posters of fine art and decor items).

Here we will mention some examples of such repositories.

*Artcyclopedia* (<http://www.artcyclopedia.com/>) is an online database of museum-quality fine art founded by Canadian John Malyon. The site is the leading guide to museum-quality fine art on the Internet. The Artcyclopedia is a form of Internet search engine and deals with art that can be viewed online, and indexes

2 600 art sites with links to around 140 000 artworks by 9 000 renowned artists.

*Artchive* (<http://www.artchive.com/>) is a virtual art gallery website, coordinated by Mark Harden. The site is a leading example of an independently established collection of high-quality pictures, which are important in the history of art. Works include art from various periods, such as Abstract Expressionism, Baroque, Impressionism, Renaissance, Romanticism, Rococo, Surrealism and more.

*OCAIW* (Orazio Centaro's Art Images on the Web) (<http://www.ocaiw.com/>) is an educational and non-profit site for art-lovers, teachers, students, artists and collectors. The catalogue consists of information about painters, sculptors, architects and photographers. Every section of the catalogue includes a listing of the greatest artists in Art History. The authors' index lists about 1 500 artists from medieval times to the present.

*WebMuseum* (<http://www.ibiblio.org/wm/>) is one of the earliest examples of a virtual museum. Starting in 1994 as WebLouvre, now many mirror sites are established throughout the world. It provides an excellent archival and educational resource of good quality art images and information. WebMuseum is part of one of the largest "collections of collections" on the Internet *ibiblio* (<http://www.ibiblio.org/>), a conservancy of freely available information including software, music, literature, art, history, science, politics, and cultural studies.

*Web Gallery of Art* (<http://www.wga.hu/>) is a virtual museum and searchable database of European painting and sculpture of the Romanesque, Gothic, Renaissance, Baroque, Neoclassicism, Romanticism periods (1000-1850), currently containing over 23 000 reproductions. Picture commentaries, artist biographies as well as guided tours, contemporary music, catalogues, free postcards and other services are available.

*Bactoclassics* (<http://www.Bactoclassics.com/>) is a new virtual art gallery (since 2009) created by the Italian division of Microsystems MS Lab, which provides an insight into the creations of artists past and present, where paintings are classified not only by movements and artists, but also thematically (for instance Rembrandt's paintings are grouped into the following series: Portraits; Biblical Themes; Various Paintings; Self-Portraits; Etchings; Drawings; Landscapes).

*Olga's Gallery* (<http://www.abcgallery.com/>) is one of the largest online painting museums containing works and biographies of most of the world's best known artists. Olga's Gallery was founded in 1999 as a fine art-themed website under the domain name abcgallery.com (abc is part of the motto of the site "The abc of art") by the sisters-in-law Olga and Helen. Now Olga's Gallery contains over 12 000 works of art by more than 300 painters and receives over 30 000 visitors and 1 000 000 page views daily. Artists can be searched by name, country or genre, which you will find most useful. The site also includes brief biographies of significant artists with excellent links to more specified sites converting it into an indispensable research source for young students.

Many museums currently offer online galleries supplying access to their collections. The search engines of some of them are using only the metadata using categories such as artist, title of work, subject, chronology and reference number (for instance <http://www.museodelprado.es/>). In other sites attempts to implement content-based image retrieval (CBIR) techniques during the search process are included (such as <http://www.hermitagemuseum.org/>); these techniques are presented in section 3.2.

### 3. THE INTELLECTUAL WEB ACCESS TO IMAGES

The digital repositories of art images are created to facilitate resource discovery and provide access to numerous digital objects. In this section we will outline the basic methods which are currently used for this domain.

As mentioned in Chen and Rasmussen (1999), the newly available image capture techniques, inexpensive storage, and widely available dissemination methods have made digital images a convenient and easily available information format. This increased availability of images is accompanied by a need for solutions to the problems in indexing and retrieval. Traditional concept-based indexing uses controlled vocabulary or natural language to express what an image is or what it is about. Newly developed content-based techniques rely on a pixel-level interpretation of the data content of the image. The upper stage of indexing techniques – concept-based indexing is based on mixing of simple text-based and content-based tools taking into account additional information for interconnections between perceived information from the main player of this process – "the user".

#### 3.1 Text-based indexing

This trend of work has been analysed in numerous projects. The advancement in the processes of ordering and classifying the meta-information is very significant, although text-based indexing methods are vastly manual, effort- and time-consuming. Let's only mention the richness of gathered and structured information in Getty vocabularies<sup>1</sup>. The vocabularies in this program are:

- The *Art and Architecture Thesaurus – AAT* (containing around 34 000 concepts including 131 000 terms, descriptions, bibliographic citations, and other information relating to fine art, architecture, decorative arts, archival materials and material culture),
- The *Union List of Artist Names – ULAN* (containing around 127,000 records including 375,000 names and biographical and bibliographic information about artists and architects, including a wealth of variant names, pseudonyms and language variants),
- The *Getty Thesaurus of Geographic Names – TGN* (containing around 895 000 records including around 1 115 000 names, place types, coordinates and descriptive notes focusing on places important for the study of art and architecture), and
- The *Cultural Objects Name Authority – CONA* (forthcoming in 2011; it will include authority records for cultural works, featuring architecture and movable works such as paintings, sculpture, prints, drawings, manuscripts, photographs, ceramics, textiles, furniture, and other visual media such as frescoes and architectural sculpture, performance art, archaeological artefacts, and various functional objects that are from the realm of material culture and of the type collected by museums).

The use of these ontological structures in image retrieval processing leads to a decreasing metadata amount and expands the research scope utilising defined interconnections between concepts.

In addition, a number of text-based indexing initiatives deal with the development of metadata schemas and structures to classify image information. We could mention for example *Dublin Core* (Dublin Core, 1999), which is used primarily for retrieving resources on the web, *VRA Core* (VRA, 2002), which

has elements to describe both an original work of art and its surrogate, *CIDOC CRM* (CIDOC, 2006) that gives conceptual reference model intended to facilitate the integration, mediation and interchange of heterogeneous cultural heritage information.

#### 3.2 Content-based indexing

Content-based image retrieval (CBIR) is an area where knowledge is being extracted from the image content of the digital picture archives. Starting with searching by simple visual similarity between given picture or sketch, current CBIR systems pretend to be promising assistants in the processes of searching content by visual as well as by semantic similarities and became a helpful alternative for image retrieval.

CBIR, as we see it today, is any technology that helps to organize digital image archives by their visual content. By this definition, anything ranging from an image similarity function to a robust image annotation engine falls into the range of CBIR. This characterization of CBIR as a field of study places it at a unique juncture within the scientific community. While we witness continuing effort in solving the fundamental problem of robust image understanding, we also see people from different fields such as, computer vision, machine learning, information retrieval, human-computer interaction, database systems, Web and data mining, information theory, statistics, and psychology contributing and becoming part of the CBIR community (Wang et al, 2006). Image analysis processing techniques provide a powerful means of extracting useful information from content. The efforts are directed towards the use of extracted low level information such as colour and texture, or primitives, such as salient points, corners and shapes, and higher-level information, such as objects, scene content, subject description, for constructing upper levels of understanding the meaning of the images under different aspects. Content annotation based on pixels can be used to perform search operations from objective measures and descriptors of the visual content. Effective descriptors that agree with human perception and feeling are required, but also with particular attention paid to the computer science "semantics" of images and scenes among other things. Thomas Hurtut presents a detailed survey of the literature of 2D artworks analysis techniques (Hurtut, 2010). Following a content-based taxonomy, the transition from low-level features to high-level layers of concepts is discussed. He also suggests several kinds of abstraction, which are specific to art images and distinct from the generic semantic challenge.

#### 3.3 The User Needs and Expectations

The main focus in the creation of digital art resources has to be user-centred rather than system-centred since most of the issues around this content are related to making it accessible and usable for the real users (Dobрева and Chowdhury, 2010). Leo Konstantelos examined in his doctoral dissertation the needs for scholarly information retrieval within the context of Digital Art in Digital Libraries (Konstantelos, 2009).

In the image retrieval systems, an important parameter to measure user-system interaction level is the complexity of queries supported by the system. The queries can use different modalities such as: directly filling the values of the desired features; giving the image or sketch as example; keywords or free-text, and their combination.

Exploring user needs and behaviour is a basic and important phase of system development. Currently users are mostly involved in usability studies when a set of digital resources has already been created and is being tested (for an overview on usability evaluation methods in the library domain see (George,

<sup>1</sup> [http://www.getty.edu/research/conducting\\_research/vocabularies/](http://www.getty.edu/research/conducting_research/vocabularies/)

2008)). It would be really helpful to involve users on early stages of design and planning the functionality of the product which is being developed.

Amongst the various methods for involvement users in the evaluation of a product, relevance feedback is a key one in image retrieval, because it defines the goals and the means to achieve them. It provides a compromise between a fully automated, unsupervised system and one based on subjective user needs. Relevance feedback is a query modification technique which attempts to capture the user's precise needs through iterative feedback and query refinement. It can be thought of as an alternative search paradigm to other paradigms such as keyword-based search. In the absence of a reliable framework for modelling high-level image semantics and subjectivity of perception, the user's feedback provides a way to learn case-specific query semantics. A comprehensive review can be found in (Zhou and Huang, 2003) and (Crucianu et al, 2004). The goal in relevance feedback is: given the rather small amount of interaction with the user during a session, it is important to use all the available information to improve the retrieval results. Based on the user's relevant feedback, learning-based approaches are typically used to appropriately modify the feature set or similarity measure. In practice, learning instances are few. This circumstance has generated interest in novel machine-learning techniques to solve problems such as *one-class* learning, *active* learning, and *manifold* learning. Usually, classical relevance feedback consists of multiple rounds, which leads to losing the patience in the user. Recent developments are directed to find techniques for minimizing the rounds, keeping the history of earlier user logs in the system and on this basis creating user profiles.

The increasing number of digital repositories on the Web fosters research into consumer needs and expectations about the content and services provided by the repositories. Despite the increasing number of studies on consumer expectations and perceptions, yet these studies are not sufficiently advanced to offer easily applicable models and recommendations. For example, a recent study noted that "the focus of research on digital libraries is on technical issues (e.g. information search methods, software architectures, etc.)" rather than on consumers (Khoo et al, 2009).

The evaluation of digital repositories in terms of its efficiency and ease of use is not specifically aimed at understanding users' attitudes in dept (Europeana, 2009). Recently, studies are focused on information literacy and behaviour of individual groups, e.g. young people, or the "Google generation" (Google, 2008).

## 4. THE WEB SPACE

### 4.1 The interoperability

In the web space plenty of additional issues arise connected with creating a framework that includes the image, the delivery system and the users (Jørgensen, 2001). Numerous repositories are already available and many systems for reaching the content are implemented and used in the practice.

In current years, the problems of semantic, syntactic and profile interoperability and constructing reference layers are very topical (DCMI, 2009); additional areas to explore are linked data which allow to contextualise objects in the cultural heritage domain.

Today interoperability is considered a key-step to move from isolated digital repositories towards a common information

space that allow users to browse through different resources within a single integrated environment (Vullo et al, 2010).

Some problems are connected with the scientific study and development of innovative technological solutions for assembling multimedia digital libraries for collaborative use in the context of cultural heritage, supporting their semantic interoperability and developing new services for dynamic aggregation of their resources, improvement of access, personification, intelligent curation of content, as well as content protection and ensuring intellectual property rights.

An example for addressing the interoperability challenge exhaustively is the DL.org project. It adopts a multi-level approach, along the classification of the European Interoperability Framework for eGovernment services (IDABC, 2004):

- *organizational interoperability*: refers to the cooperation between and within digital library organizations, business goals and process modelling;
- *semantic interoperability*: refers to understanding the meaning of information in digital libraries;
- *technical interoperability*: refers to interconnection, presentation and exchange of digital objects within digital library, accessibility and security issues.

### 4.2 The rights

During the process of information exchanging thorough the Web, there arises inevitably necessity of keeping the rights of authors and users of digital information. By his nature, the objects, presented in digital space can be:

- *digitally born art-objects*. These new kind of masterpieces live in digital media, which is their primary environment;
- *digital images of analogue art-objects*. Here a digitized work of art is not the work itself but an image (instance) of this work;
- *objects, which are connected with other art-objects*. For instance, an essay over some painting period of the artists, or sketch representing main characteristics of the artwork.

The standard royalties are expanded with the new ones, concerning life-cycle of digital objects in Web space. One attempt for establishing legal framework for open multimedia supply and consumption to be used by all participants in the chain is made in MPEG-21 (MPEG-21, 2002). It defines a "Rights Expression Language" standard as a means of sharing digital rights/permissions/restrictions for digital content from content creator to content consumer. As an XML-based standard, MPEG-21 is designed to communicate machine-readable license information and to do so in a "ubiquitous, unambiguous and secure" manner. This open framework provides content creators, producers, distributors, representatives and service opportunities of the existing free market. This also benefits consumers by providing access to a wide variety of content in an interoperable way. MPEG-21 is based on two basic concepts:

- A basic *unit* of distribution and transaction – a digital item – which is "object" in the multimedia framework (e.g., video art, music album);
- The concept of *user* interacting with him, which is "subject" in this framework.

The main purpose of MPEG-21 is the defining technologies needed to support consumer access to digital items and their free exchange, use, trading and manipulation in an efficient and transparent manner.

### 4.3 Semantic Web

During the years, the ability of processing the information as well as expanding the ways of data exchanging is increasing in parallel. The development of computing and communication capacities allows to place the user in the central point of the process of information exchange and to enable him to use all power of the intellectualized tools for satisfying his wishes. Amit Agarwal (Agarwal, 2009) provides a simple and clear comparison between Web 1.0, Web 2.0 and Web 3.0 (Table 1).

Web 1.0	Web 2.0	Web 3.0
"the mostly read only web"	"the wildly read-write web"	"the portable personal web"
Focused on companies	Focused on communities	Focused on the individual
Home pages	Blogs	Lifestream
Owning content	Sharing content	Consolidating dynamic content
Britannica Online	Wikipedia	The semantic web
Directories ("taxonomy")	Tagging ("folksonomy")	User behavior ("me-onomy")
Netscape	Google, Flickr, YouTube	iGoogle, NetVibes

Table 1. "Comparison table" between Web 1.0, Web 2.0, Web 3.0 (excerpt from (Agarwal, 2009))

Starting from read-only content and static HTML websites in Web 1.0, where people are only passive receivers of the information, Web 2.0 became as participation platform, which allows users not only to consume but also to contribute information through blogs or sites like Flickr (<http://www.flickr.com/>), YouTube (<http://www.youtube.com/>), etc. These sites may have an "Architecture of participation" that encourages users to add value to the application as they use it. According to David Best (Best, 2006), the characteristics of Web 2.0 are: rich user experience, user participation, dynamic content, metadata, web standards and scalability. Further characteristics such as openness, freedom and collective intelligence by way of user participation, can also be considered as essential attributes of Web 2.0.

The pros and cons of using such paradigm as well as other one are many; for a good range of initiatives of social media outreach in the cultural heritage institutions see (WIDWISAWN, 2008). Let's mention alternatives, discussed from Eric Raymond in (Raymond, 1999) concerned two fundamentally different development styles, the "cathedral" model of most of the commercial world versus the "bazaar" model of the Linux open-source world, where the advantages of such social self-build systems are shown. Here, the situation is similar. For instance, while the Encyclopaedia Britannica Online (<http://www.britannica.com/>) relies upon experts to create articles and releases them periodically in publications, Wikipedia relies on anonymous users to constantly and quickly contribute information. And, as in many examples, the happy medium is the right position. Many art repositories and portals are used for educational purposes; consequently control over the main presented text is very important. On the other hand, they are natural places for users to share their own opinion and

to have a space for communication. The interest of users measured in number of hits and traces of their activity grows when they are able to add their own content or to comment on existing commentaries.

In the area of art images social networking sites can help extend the number of users consulting an image; for example the Library of Congress explained at the American Library Association annual conference in 2010 that the number of visitors consulting images which can be both accessed on the Library of Congress website and on flickr.com attracted higher number of visitors on flickr. The user generated comments on flickr also helped to improve the metadata records the Library maintained.

Not much time passed before the idea of "Web 3.0" appeared. Amit Agarwal states that Web 3.0 is about semantic web (or the meaning of data), personalization (e.g. iGoogle), intelligent search and behavioural advertising among other things. While Web 2.0 uses the Internet to make connections between people, Web 3.0 will use the Internet to make connections with information. The intelligent browsers will analyze the complex requests of the user given in natural language, search the Internet for all possible answers, and then organize the results for him. The adaptation to user specifics and aptitudes will be based on capturing the historical information thorough searching the Web. Many of the experts believe that the Web 3.0 browser will act like a personal assistant. The computer and the environment will become artificial subjects, which will pretend to communicate in real manner as real humans. Of course, the problems of applying rights policies in such a new atmosphere are crucial.

However, addressing the rights is an additional issue which needs to be solved. A core problem in this domain is which combination of retrieval methods and techniques can lead to high quality image discovery.

## 5. CONCLUSIONS

We presented in this paper a succinct overview of the development of repositories of digital art images and then highlighted the specialized search methods in this domain. Compared to other cultural heritage materials, to improve the accessibility of digitized art images, a transition in the methods is requested from approaches involving only textual metadata towards "hybrid" approaches of retrieval using content based image retrieval jointly with the metadata. Linked data is also an essential trend to integrate better; however it could rather help to contextualise art images within additional scattered information available on the web. Methods for analysis of the image content are the only current possibility to achieve accessibility which would also address characteristics of the images themselves besides those which had been given as metadata.

As in other cultural heritage domains, digital art images also require methods to resolve art issues and to experiment with and implement approaches for involving the users without compromising the trustworthiness of the resources.

We believe that areas which will develop with a priority in the very near future are:

- Further refining of specialized image retrieval techniques seeking to both improve the quality of the analysis and to overcome the semantic gap;
- Defining best practices in involving the users (individual users as well as communities of users);

- Sustaining trustworthiness of the resources when social media tools are used to add user generated content;
- Improving not only the information delivery but also the user experiences and expanding the delivery of information with immersing technologies.

The ultimate goal is to facilitate the access to art objects in digital form and to convert it to fun and a great experience.

## REFERENCES

### References from Journals:

Chen, C. – C., Wactlar, H., Wang, J. Z., Kiernan, K., 2005. Digital imagery for significant cultural and historical materials – An emerging research field bridging people, culture, and technologies. *International Journal Digital Libraries*, 5(4), pp. 275–286.

Jørgensen, C., 2001. Introduction and overview. In: *Journal of the American Society for Information Science and Technology*, 52(11), pp. 906–910.

Mattison, D., 2004. Looking for good art. In *Searcher - The Magazine for Database Professionals*. Vol. 12, Part I – N.8, pp. 12–35; Part II – N.9, pp. 8–19; Part III – N.10, pp. 21–32.

Zhou, X., Huang, T., 2003. Relevance feedback in image retrieval: a comprehensive review. *Multimedia Systems*, Vol. 8, No. 6, pp. 536–544.

### References from Books:

George, C., 2008. *User-Centred Library Websites. Usability Evaluation Methods*. Chandos publishing, 231 pp.

Janson, A., 2004. *History of Art*, Prentice Hall; 6th edition, 1032 pp.

### References from Other Literature:

Best, D., 2006. Web 2.0 Next big thing or next big Internet bubble? *Lecture Web Information Systems*. Technische Universiteit Eindhoven.

Crucianu, M., Ferecatu, M., Boujemaa, N., 2004. Relevance feedback for image retrieval: a short survey. In: *State of the Art in Audiovisual Content-Based Retrieval, Information Universal Access and Interaction Including Data Models and Languages (DELOS2 Report)*.

Dobрева, M., Chowdhury, S., 2010. A User-Centric Evaluation of the Europeana Digital Library. In: *ICADL 2010*, LNCS 6102, pp. 148–157.

Gradmann, S., 2010. *Knowledge = Information in Context: on the Importance of Semantic Contextualisation in Europeana*. Europeana White Paper 1, 19 pp.

Maitre, H. Schmitt, F., Lahanier, C., 2001. 15 years of image processing and the fine arts. In: *Proc. of ICIP*, Vol. 1, pp. 557–561.

Vullo, G., Innocenti, P., Ross, S., 2010. Interoperability for digital repositories: towards a policy and quality framework. In: *Fifth Int. Conf. on Open Repositories (OR 2010)*, Madrid, Spain.

Wang, J., Boujemaa, N., Del Bimbo, A., Geman, D., Hauptmann, A., Tesic, J., 2006. Diversity in multimedia information retrieval research. In: *Proceedings of MIR Workshop*, ACM Multimedia, pp. 5–12.

### References from websites:

Agarwal, A., 2009. *Web 3.0 concepts explained in plain English*. 30.05.2009. <http://www.labnol.org/internet/web-3-concepts-explained/8908/> (accessed 16.08.2010)

Chen, H.-L., Rasmussen, E., 1999. Intellectual access to images – image database systems. *Library Trends*, Fall. [http://findarticles.com/p/articles/mi\\_m1387/is\\_2\\_48/ai\\_59473802/](http://findarticles.com/p/articles/mi_m1387/is_2_48/ai_59473802/) (accessed 16.08.2010)

CIDOC, 2006. *CIDOC Conceptual Reference Model*. <http://www.cidoc-crm.org/> (accessed 16.08.2010)

DCMI, 2009. *Interoperability for Dublin Core Metadata*. <http://dublincore.org/documents/interoperability-levels/> (accessed 16.08.2010)

Dublin Core, 1999. *Dublin Core metadata element set*, ver.1.1. <http://dublincore.org/documents/dces/> (accessed 16.08.2010)

Europeana, 2009. *Europeana – Online Visitor Survey*. [http://www.edlfdoundation.eu/c/document\\_library/get\\_file?uuid=e165f7f8-981a-436b-8179-d27ec952b8aa&groupId=10602](http://www.edlfdoundation.eu/c/document_library/get_file?uuid=e165f7f8-981a-436b-8179-d27ec952b8aa&groupId=10602) (accessed 16.08.2010)

Google, 2008. *Google generation*. <http://www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/googlegen.aspx> (accessed 16.08.2010)

Hurtut, T., 2010. *2D artistic images analysis, a content-based survey*. 29 pp. [http://hal.archives-ouvertes.fr/hal-00459401\\_v1/](http://hal.archives-ouvertes.fr/hal-00459401_v1/) (accessed 16.08.2010)

IDABC, 2004. *European Interoperability Framework for pan-European eGovernment Services*. <http://ec.europa.eu/idabc/en/document/2319> (accessed 16.08.2010)

Khoo, M., Buchanan, G., Cunningham, S., 2009. Lightweight user-friendly evaluation knowledge for digital libraries, *D-Lib Magazine*, July/August 2009. <http://www.dlib.org/dlib/july09/khoo/07khoo.html> (accessed 16.08.2010)

Konstantelos L., 2009. Digital art in digital libraries: eliciting the needs of users for scholarly information retrieval. In: *Bulletin of IEEE Tech. Committee on Digital Libraries*, 5(2). <http://www.ieee-tcdl.org/Bulletin/v5n2/Konstantelos/konstantelos.html> (accessed 16.08.2010)

MPEG-21, 2002. *MPEG-21 Overview v.5*. [http://mpeg.chiariglione.org/standards/mpeg-21/mpeg-21.htm#\\_Toc23297976](http://mpeg.chiariglione.org/standards/mpeg-21/mpeg-21.htm#_Toc23297976) (accessed 16.08.2010)

Raymond, E., 1999. *The Cathedral & the Bazaar*. O'Reilly. <http://www.catb.org/~esr/writings/cathedral-bazaar/cathedral-bazaar/> (accessed 16.08.2010)

VRA, 2002. *VRA Data Standards Committee*. <http://www.vraweb.org/organization/committees/datastandards/index.html> (accessed 16.08.2010)

WIDWISAWN, 2008. *Special issue on Web 2.0*. Vol. 6 N. 1. [http://widwisawn.cdlr.strath.ac.uk/issues/vol6/issue6\\_1\\_1.html](http://widwisawn.cdlr.strath.ac.uk/issues/vol6/issue6_1_1.html) (accessed 16.08.2010)



**ACKNOWLEDGEMENT**

This work was supported in part by Hasselt University under the Project R-1875 "Search in Art Image Collections Based on Color Semantics", by the FP7-supported project SHAMAN, and by the Bulgarian National Science Fund under the Project D002-308 "Automated Metadata Generating for e-Documents Specifications and Standards".