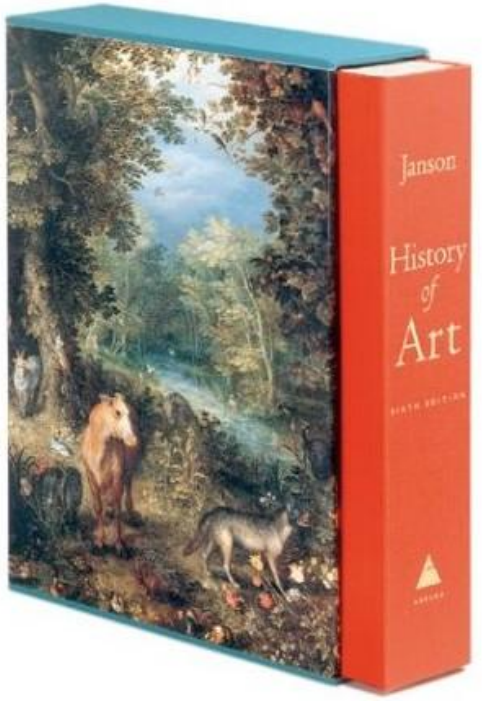
DISCOVERY AND USE OF ART IMAGES ON THE WEB: AN OVERVIEW

Krassimira Ivanova, Milena Dobreva, Peter Stanchev, Koen Vanhoof

Institute of Mathematics and Informatics, BAS, Sofia, Bulgaria
Center for Digital Libraries Research, University of Strathclyde, UK
Kettering University, Flint, MI, USA
Hasselt University, Hasselt, Belgium

kivanova@math.bas.bg

Introduction



Each touch to the artwork causes building the bridge between cultures and times.

“Research on significant cultural and historical materials is important not only for preserving them but for preserving an interest in and respect for them” [Chen et al, 2005] .

Computer analyzing, indexing and delivering techniques become irreplaceable assistance in many fields, related with cultural heritage.

This work was supported in part

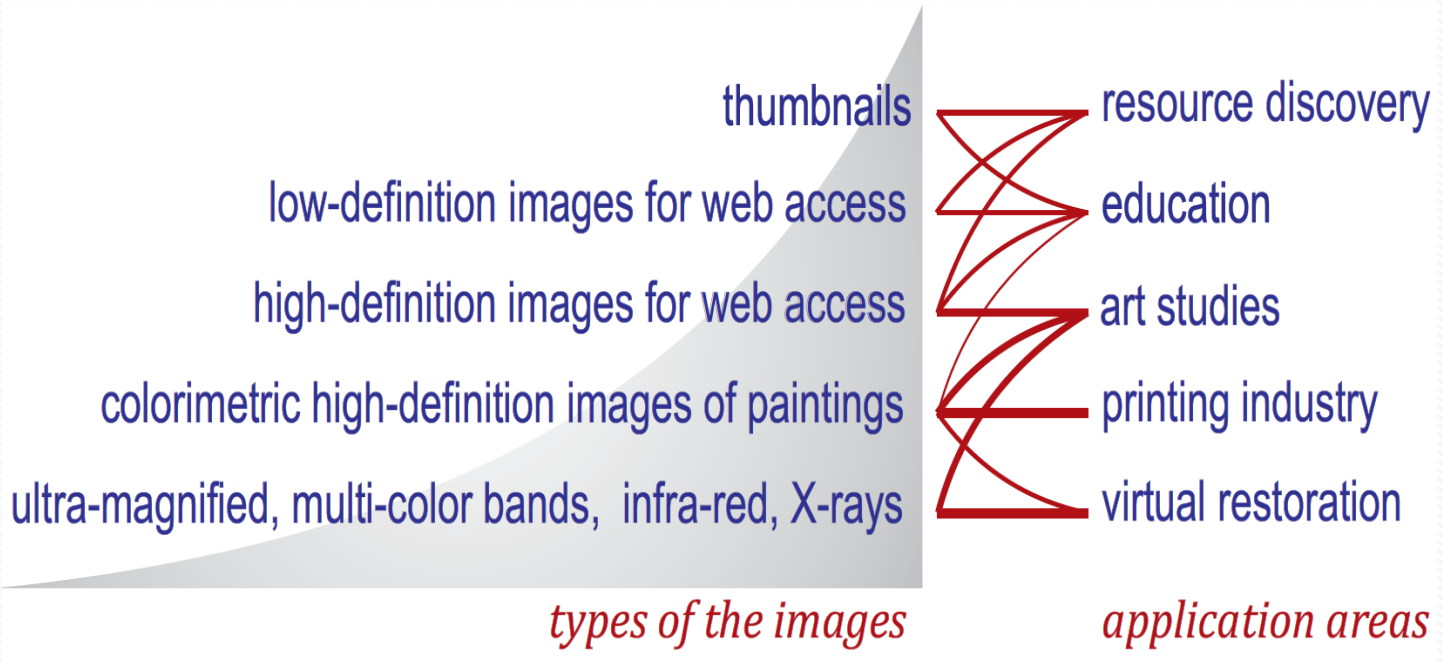
- by Hasselt University: Project R-1875 "Search in Art Image Collections Based on Color Semantics",
- by the FP7-supported project SHAMAN,
- by the Bulgarian NSF: Project D002-308 "Automated Metadata Generating for e-Documents Specifications and Standards".

Structure

- Digital art painting repositories for universal web users
- Intellectual image access – text based indexing; content-based techniques; concept based indexing
- User needs and expectations
- The web space – interoperability; rights; Web 2.0; Web 3.0
- Future trends

Digitised art images – quality and usage:

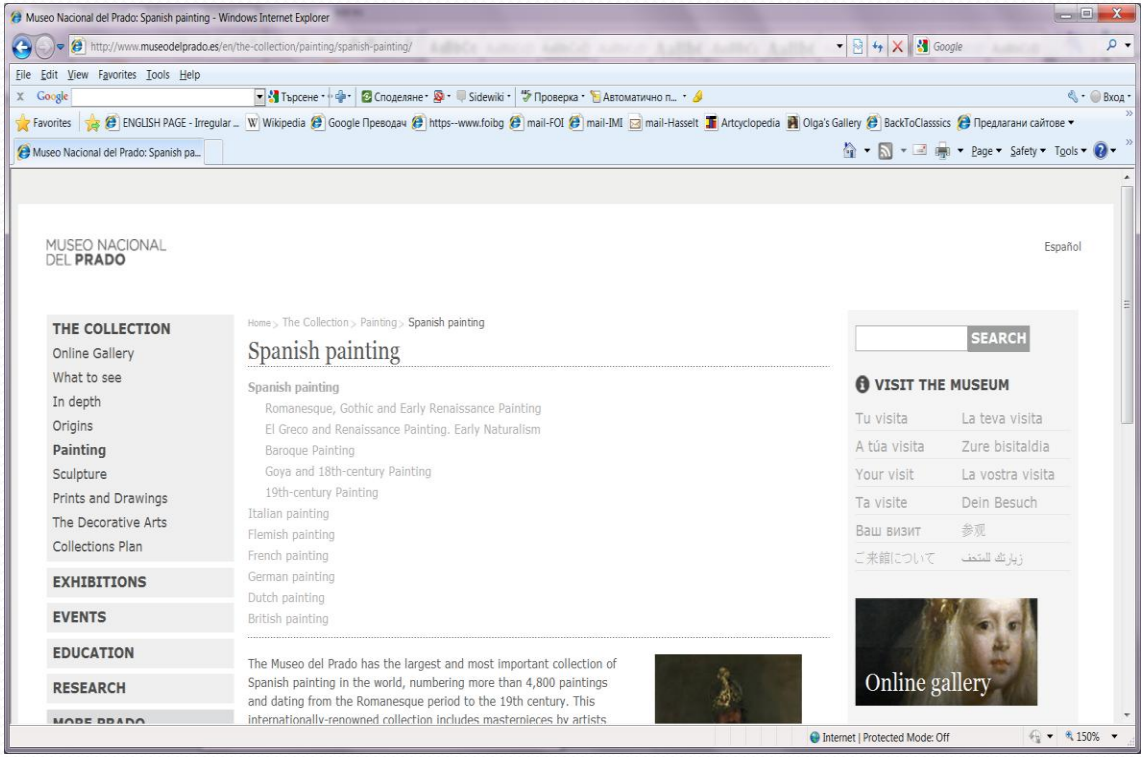
- image enhancement for publication purposes
- virtual restoration of artworks
- artistic practices studies
- art history investigation
- authentication
- watermarking
- …



Collections

- Text-based search:

  author
  title
  reference number



Museo National del Prado

http://www.museodelprado.es/

Collections

- Content-based image retrieval

Color search

Layout search

Implementation of QBIC in the Hermitage museum

http://www.hermitagemuseum.org/

Collections

Retrieval by:

features – color, texture, shape or spatial location;

abstract attributes – a step to bridging the semantic gap from low-level pixel information to high-level perceiving by the human being.

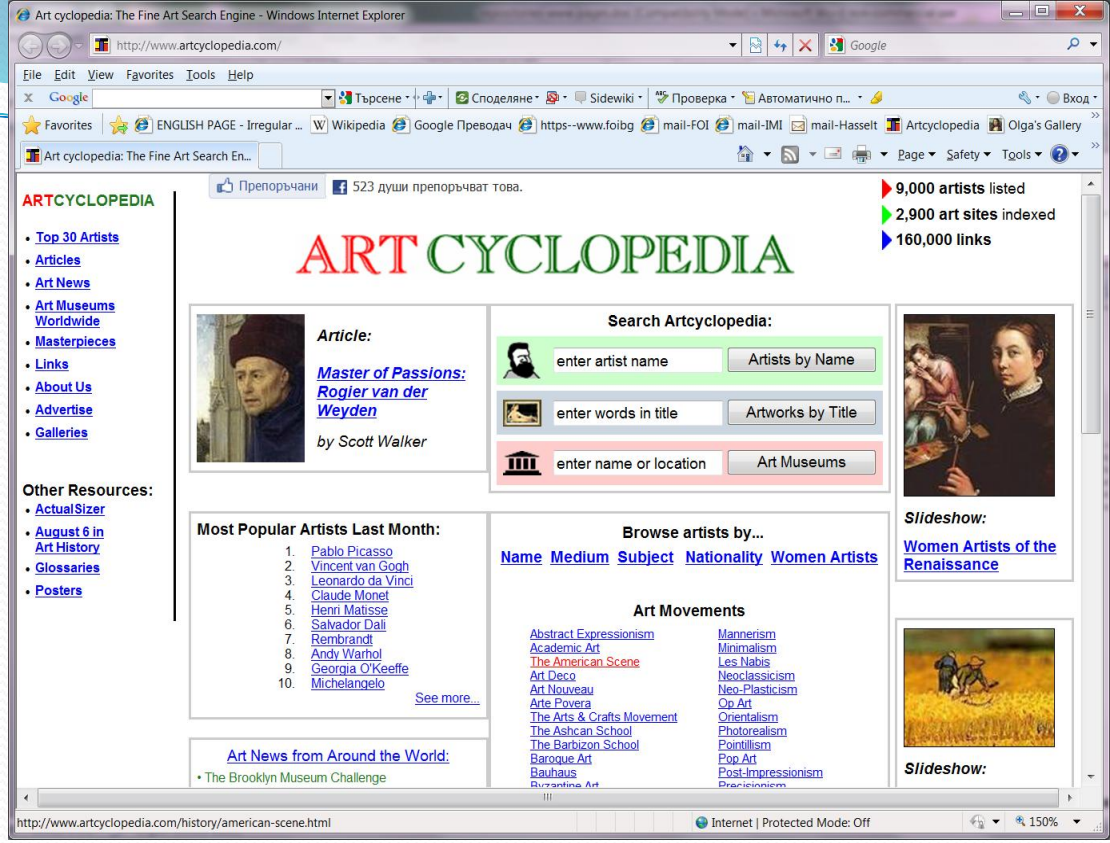
Collage for the London Guildhall Library

http://collage.cityoflondon.gov.uk/collage/

Portals

ArtCyclopedia

www.artcyclopedia.com

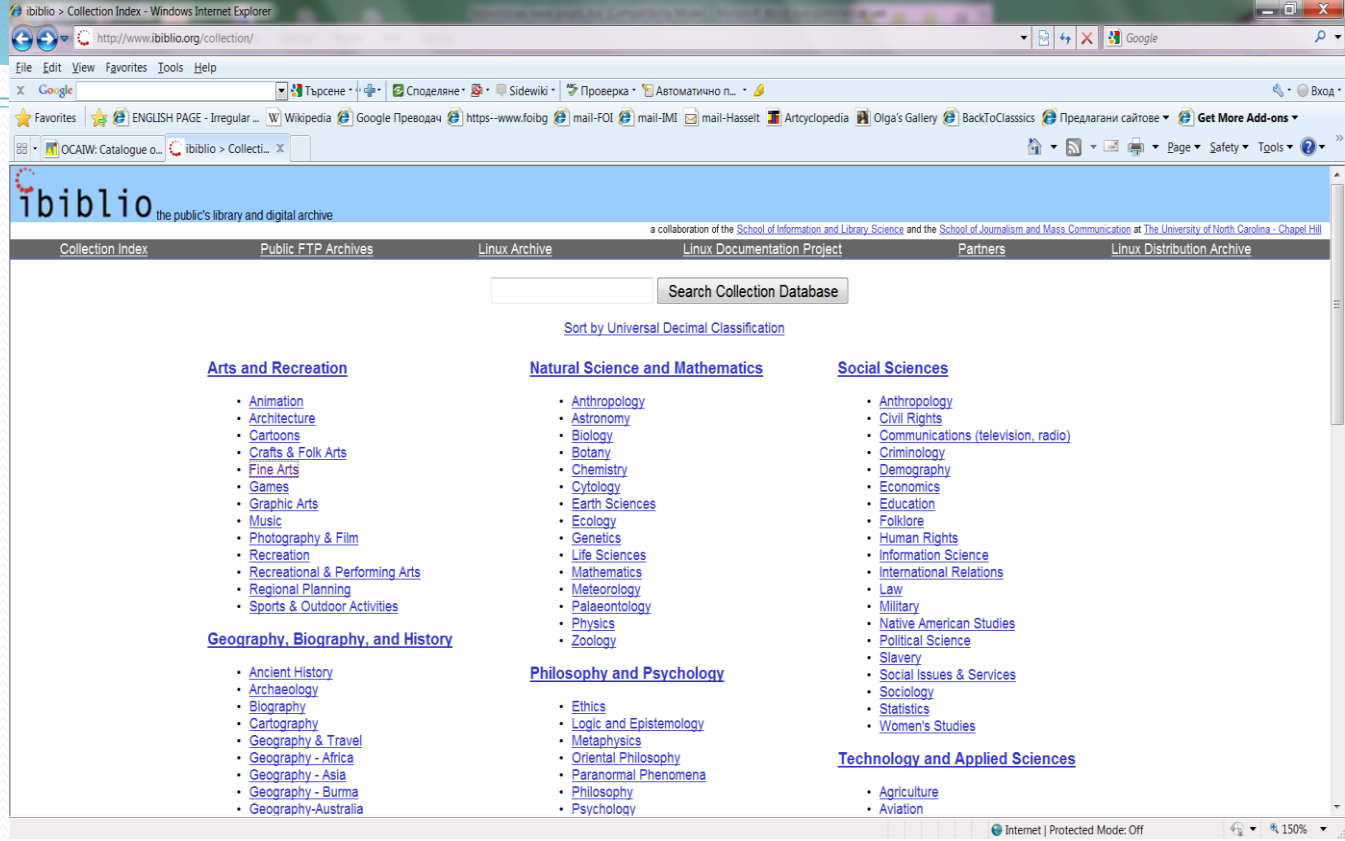


- Online database of museum-quality fine art founded by Canadian John Malyon;
- Leading guide to museum-quality fine on the Internet;
- A form of **Internet search engine**;
- Indexes 2 600 art sites (from museums, galleries and auction houses);
- Links to around 140 000 artworks by 9 000 renowned artists.

Portals

ibiblio

www.ibiblio.org

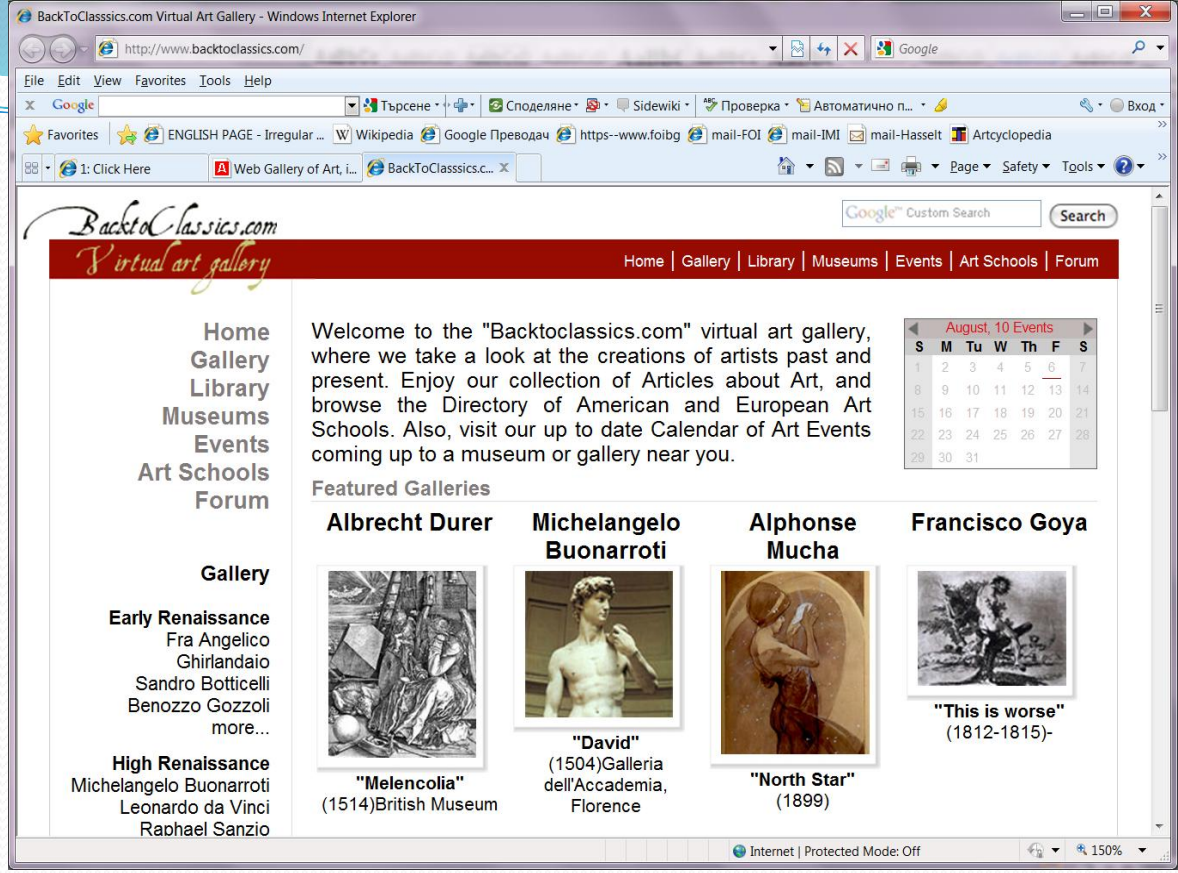


- WebMuseum is a part of one of the **largest "collections of collections" on the Internet** ibiblio a conservancy of freely available information, including software, music, literature, art, history, science, politics, and cultural studies;
- One of the earliest examples of a virtual museum;
- Starting in 1994 as WebLouvre, now many mirror sites are established throughout the world;
- excellent archival and educational resource of good quality art images and information.

Virtual Galleries

Backtoclassics

www.backtoclassics.com



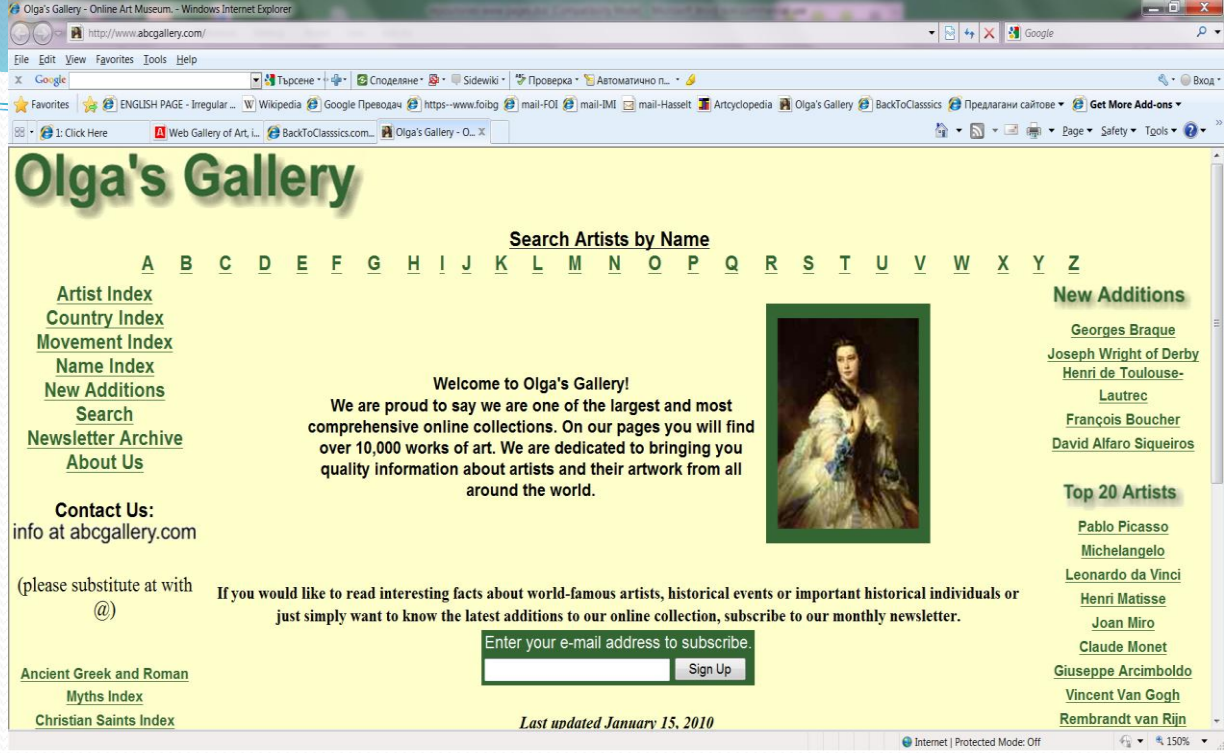
- New virtual art gallery (since 2009), created by Italian division of Microsystems MS Lab;
- The paintings are classified not only by movements and artists, but also **thematically**

  (Rembrand's paintings are grouped into: Portraits; Biblical Themes; Various Paintings; Self-Portraits; Etchings; Drawings; Landscapes)

Virtual Galleries

Olga's Gallery

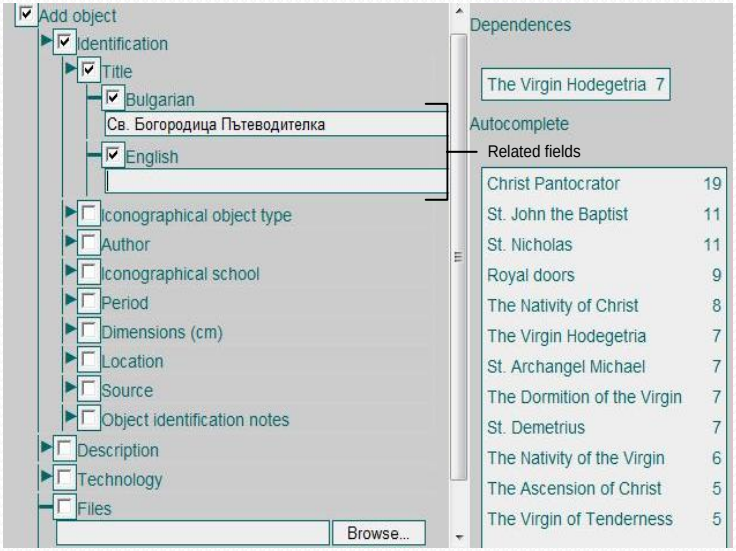
www.abcgallery.com



- One of the largest online painting museums; **Excellent educational site**;
- Works and biographies of most of the world's best known artists;
- Started by sisters-in-law Olga and Helen 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");
- contains over than 12 000 works of art by more than 300 painters and receives over 30 000 visitors and 1 000 000 page views daily;
- Searching by name, country or genre.

Example from Bulgaria: BIDL

Bulgarian Iconographical Digital Library



- tree-based annotation template for the semantic description of the iconographical objects;
- options for auto-completion, reuse of values, bilingual entering of data, automated media watermarking, resizing and conversing;
- special ontological model, describing the knowledge of East Christian Iconographical Art for annotation and semantic indexing of iconographical artifacts.

The global vision of BIDL is based on a long-term observation of the users' preferences, cognitive goals, and needs, aiming to find an optimal functionality solution for the end users.

EU initiatives for harmonization of the activities in digitization of cultural heritage

- NARCISSE (1990-1992) - creating a very high digitized image bank, supervised by a multilingual text database;
- Artiste (2000-2002) - developing an integrated art analysis and navigation environment for supporting the work of professional users in the fine arts;
- MINERVA+ (since 2004) - enlarging the existing thematic network of European Ministries of Culture in digitisation of cultural and scientific content;
- CATCH (since 2005) - improving the cross-fertilization between scientific research and cultural heritage;
- Europeana (since 2005) - building on Europe's rich heritage, combining multicultural and multilingual environments with technological advances and new business models.

Intellectual image access

- **Text-based indexing**: (Europeana - Linked data gives machines the ability to make associations and put search terms into context [St. Gradmann, 2010]) :
  - Ordering and classifying meta-information (Getty's AAT, ULAN, TGN, CONA; IconClass; …)
  - Developing metadata schemas and structures to classify image information (Dublin Core, VRA Core, CIDOC CRM, …)

- **Content-based indexing** (These techniques rely on a pixel-level interpretation of the data content of the image):
  - Fighting with existing **aesthetic**, **semantic** and **abstraction gaps** between human perceiving and machine interpreting of the visual content

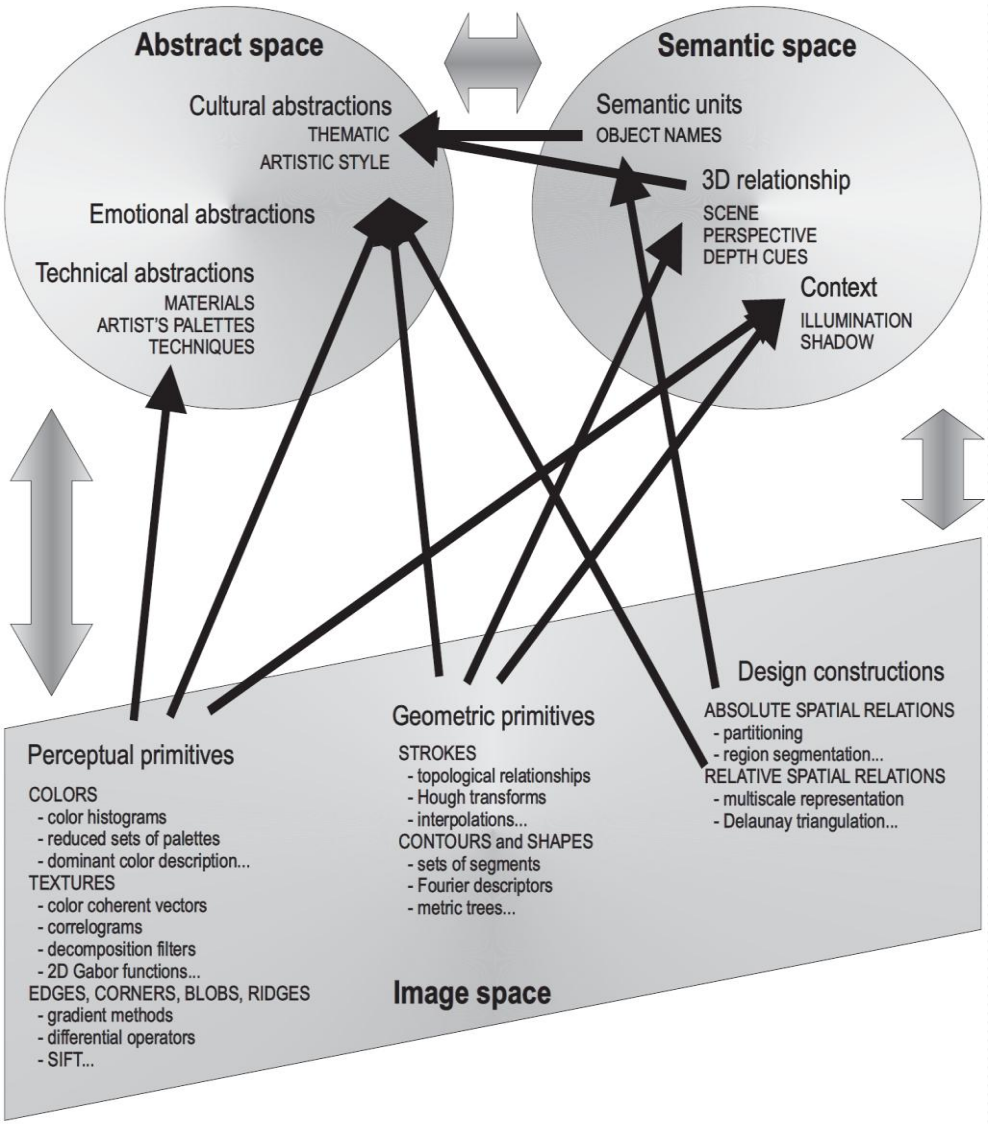
Taxonomy of art image content

Image space contains visual primitives, needed to record an image through visual perception.

Semantic space is related to the meaning of the elements, their potential for semantic interpretation.

Abstract space reflect cultural influences, specific techniques, and emotional responses evoked by an image.

** Inspired by*
[Burford et al, 2003] and [Hurtut, 2010]



User needs and expectation

- **complexity of queries** - important parameter to measure user-system interaction
- **relevance feedback** - defines the goals and the means to achieve them
- **understanding users' attitudes in depth** - crucial aspect, which must be taken into account in the process of creation and serving future digital repositories

  (studying of behaviour of individual groups,
  e.g. young people, or the "Google generation ")

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

Comparison table between Web 1.0, Web 2.0, Web 3.0,
excerpt from Amit Agarwal [2009]

- The characteristics of Web 2.0 [David Best, 2006]:

  rich user experience, user participation, dynamic content, metadata, web standards and scalability

- Britannica Online vs. Wikipedia
- Analogy – the "cathedral" model of most of the commercial world for software development versus the "bazaar" model of the Linux open-source world [Eric Raymond, 1999]

Web 2.0

For art repositories and portals, reached by universal citizens:

- controlling the main presented text (for educational purposes)
- supplying natural places for users to share their own opinion and to have a space for communication
- using social networks for improving metadata records (such as Library of Congress of USA use info from flickr.com for extending the number of users consulting images)

Web 3.0

Characteristics of Web 3.0 [Amit Agarwal, 2009]:

semantic web (or the meaning of data), personalization, intelligent search and behavioural advertising among other things.

- Protecting the rights:
  - for authors, holders, etc. - already standardized in MPEG-21, …
  - protecting the user profile which is kind of private personal information

- For researchers in the field of CBIR - Improving retrieval methods and techniques for achieving high quality image discovery

Conclusion

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 improve the quality of the analysis and to overcome the semantic and abstraction gaps;
- ✓ Defining best practices in involving the 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.

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

Krassimira Ivanova, Milena Dobreva, Peter Stanchev, Koen Vanhoof

Thank you for the attention

kivanova@math.bas.bg