Workshop on Computational Methods
in the Humanities - COMHUM2018
University of Lausanne, Switzerland, 4-5 June 2018

The Confluence in Digital Humanities:

the Computer Scientist, the Digital Humanist, and the Final User

(Invited Talk, 4 June 2018)

Maristella Agosti

Information Management Systems Research Group (IMS)

Department of Information Engineering

University of Padua, Italy

The Confluence in Digital Humanities

- How to learn a lesson
- A posteriori analysis of the architecture and skills
- The birth of the digital library area
- Why we need to devise new models and systems
- Relevant case studies

How to learn a lesson

The Italian National Project of Library Automation

- The Italian national project of library automation, called SBN - *Servizio Bibliotecario Nazionale*, is an advanced library automation project started in 1970s
  - Different library automation systems at national/regional/local level cooperating in a networked/hierarchical organisation
- Until late in the 1980s
  - The public online access to bibliographic data was not available, only traditional card catalogues were in use

SBN at the University of Padua

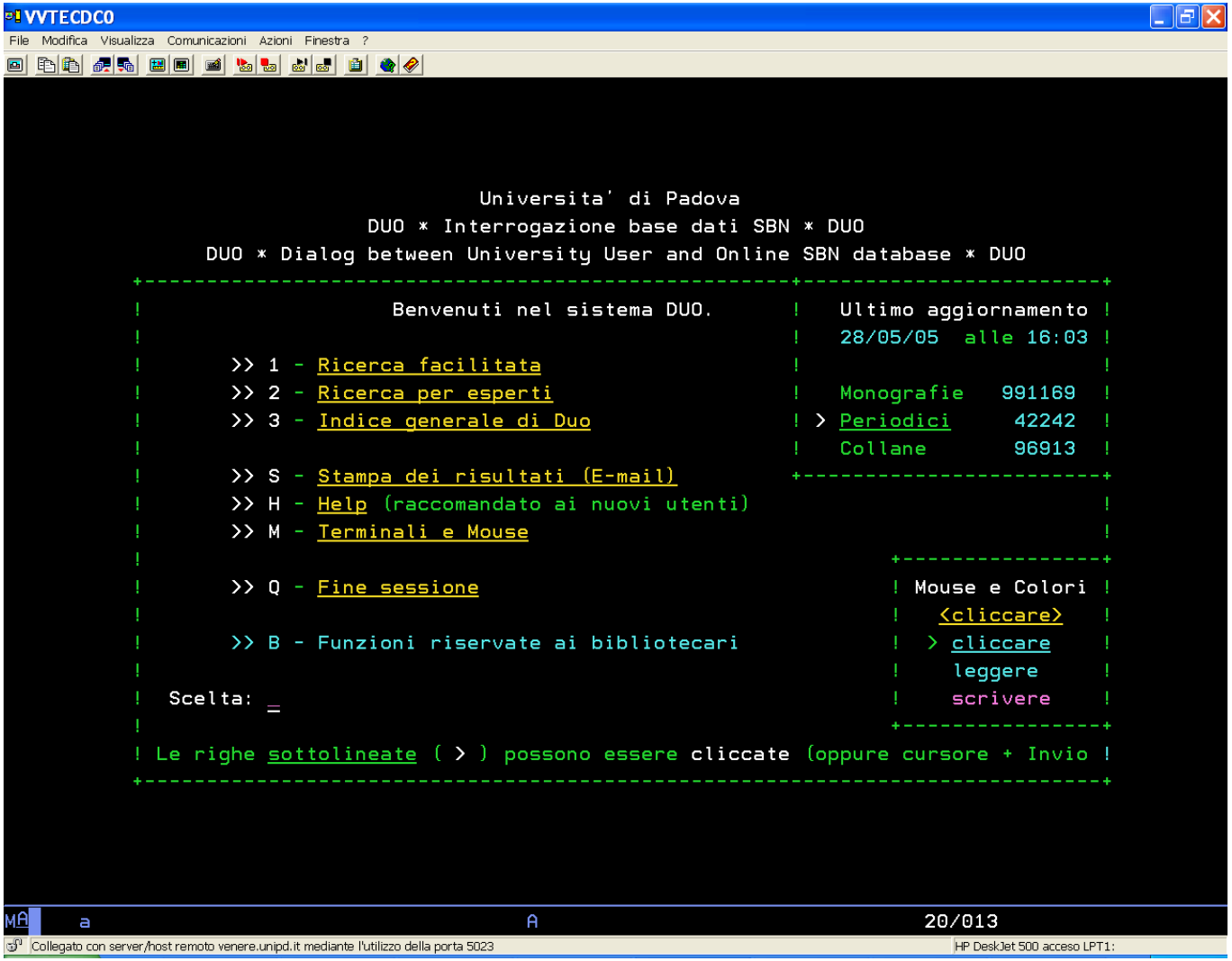
- The University of Padua became a node of the SBN project in the late 1980s and we were considering in giving a direct and innovative access to the catalogue to final and professional users
- We launched a project for an OPAC: the Project Team members were computer scientists, professional librarians, and users of the university libraries

The OPAC was Named DUO

DUO as Dialogue between University User and Online SBN Database

- OPAC search functions on structured catalogue data like title, author…
- All structured catalogue data related to a library item were transformed into unstructured data and were combined together to form a virtual textual document for each library item
- Innovation: All virtual documents formed a document collection that was indexed and fully searchable with textual queries
  - Prototype available to users in June 1991 on the Internet through the “OPAC” public login using Telnet
  - Possibly the first OPAC openly accessible on the Internet free of charge - the Web did not exist at that time

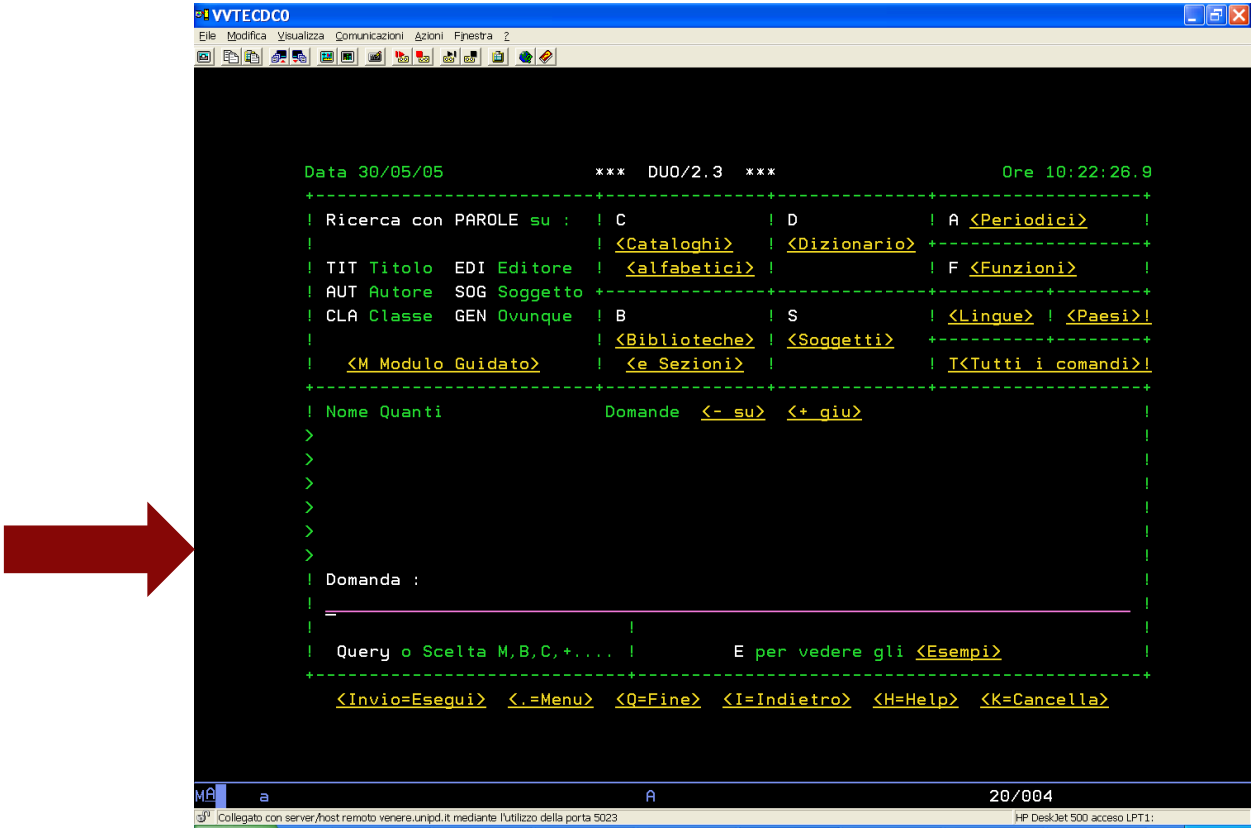
The OPAC DUO Interface (in Italian)



Maristella Agosti

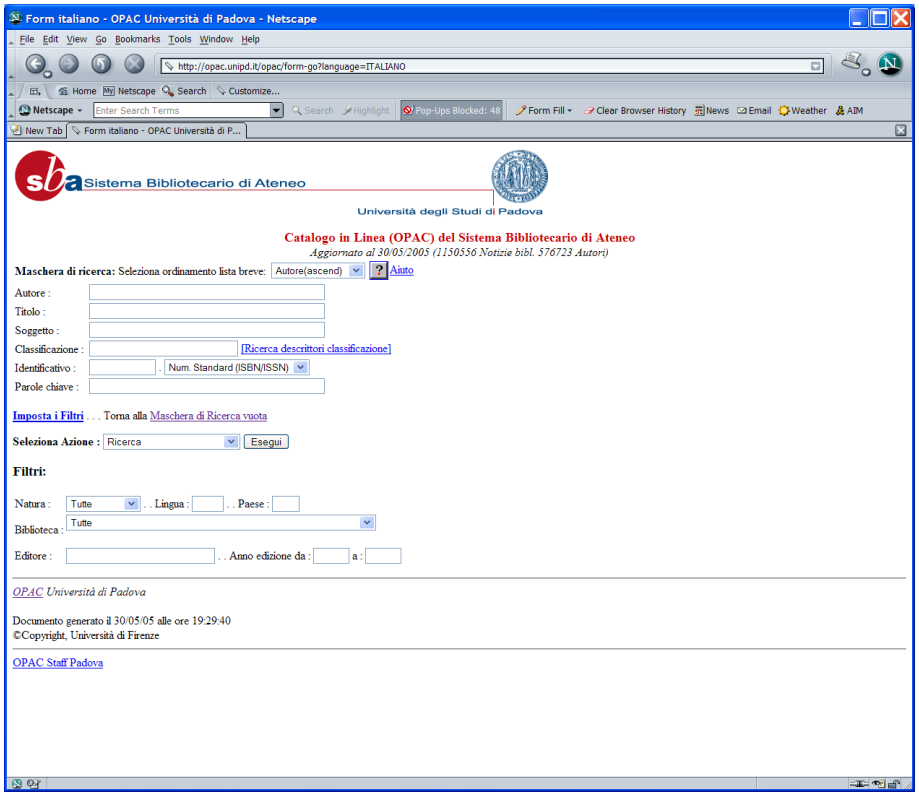
A Text Box for Free Query Input

- Text box for free query input was innovative
- IR functions similar to the ones made available years later through search engines



We studied Okapi – probably the first system with a text box for free query input

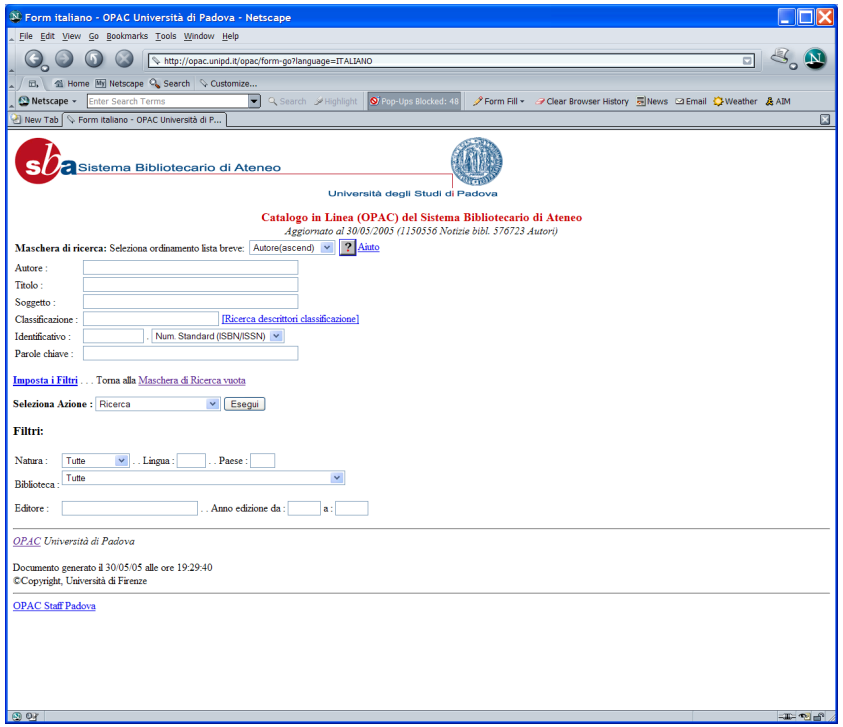
“Evolution” of DUO: Access to the Catalogue through the Web



The Web browser: a fascinating graphic interface

Maristella Agosti

“Evolution” of DUO: Access to the Catalogue through the Web

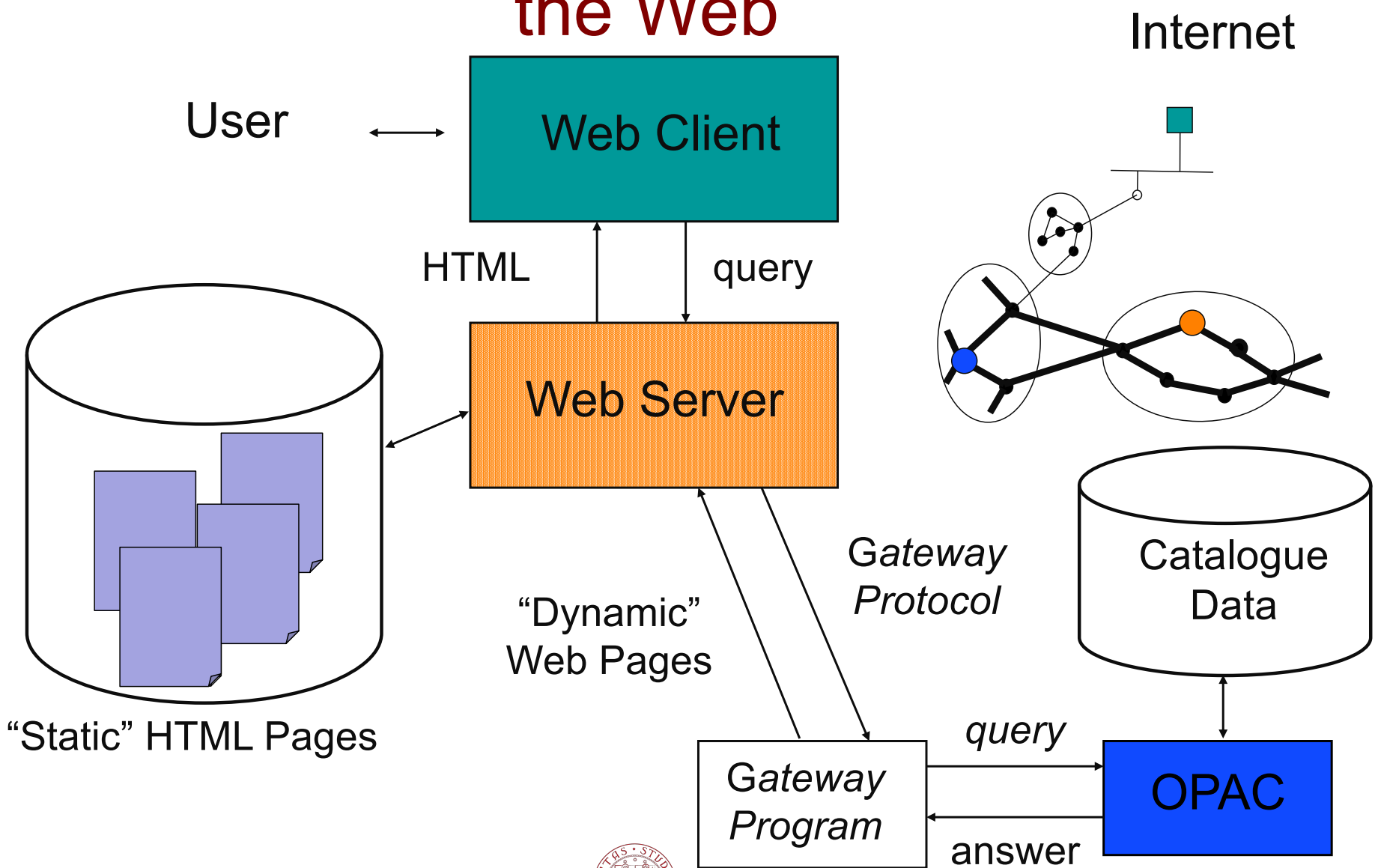


The time was not ripe for Web applications:

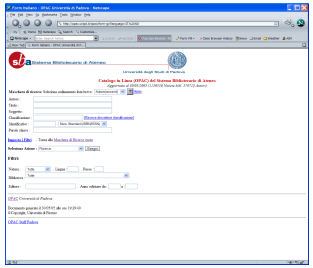
IR functions were lost, mostly because of a lack of cooperation and understanding between the Digital Humanist and the Computer Scientist

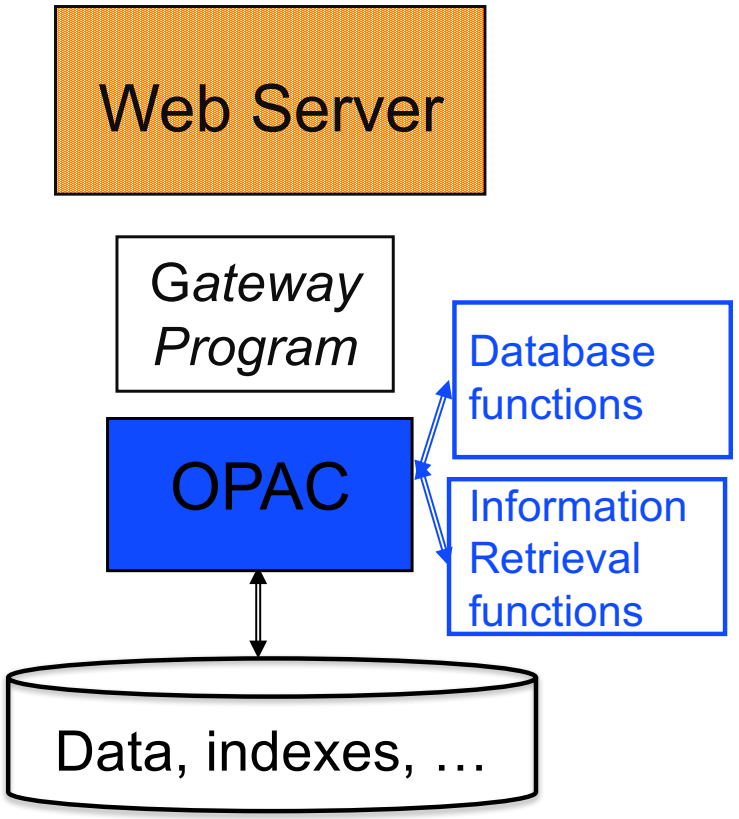
A posteriori analysis of the architecture and skills

Architecture: Access to an OPAC through the Web

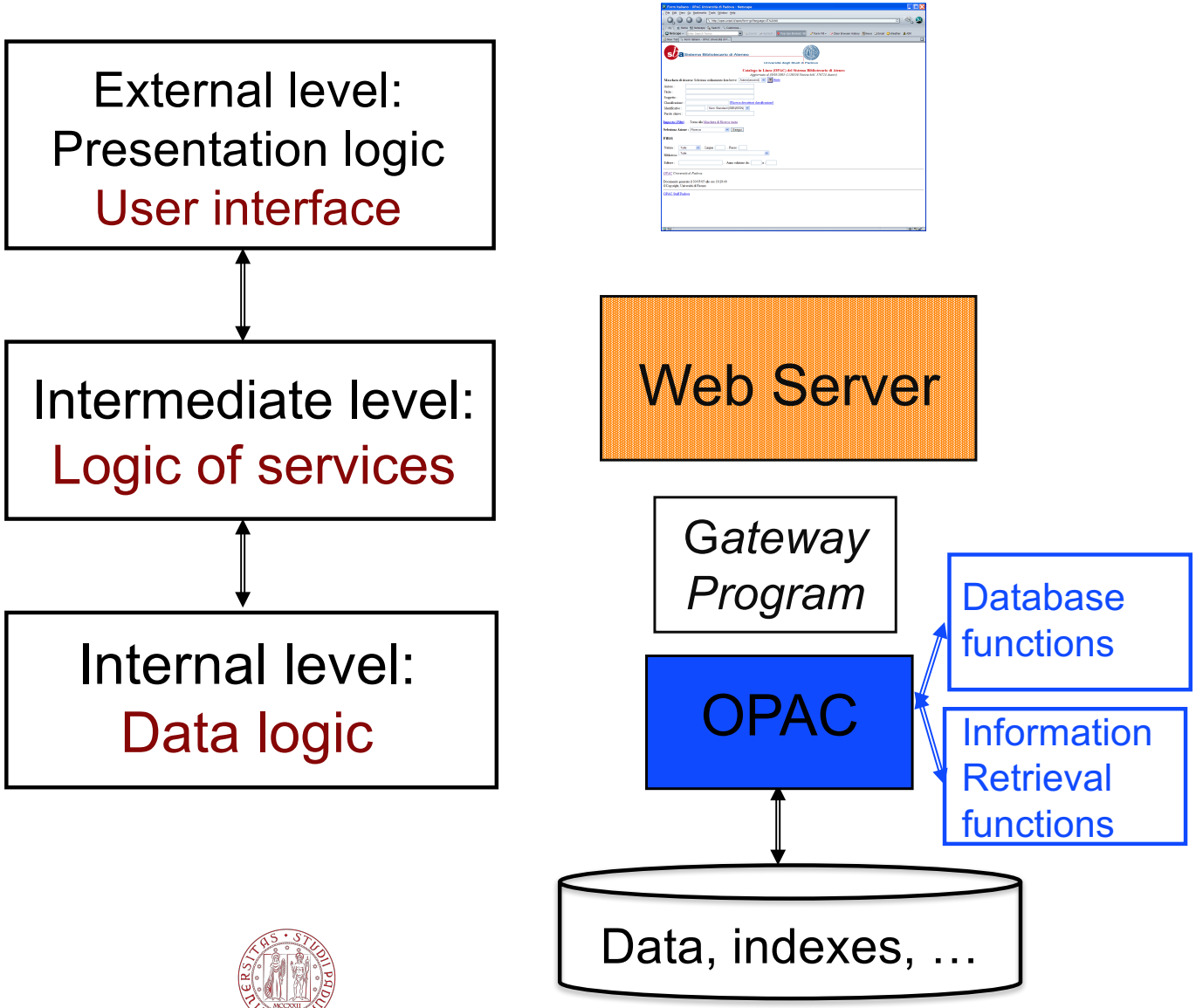


A posteriori analysis of the architecture





A posteriori analysis of the architecture



A posteriori analysis of the necessary skills to build such an application

Information need representation through a query

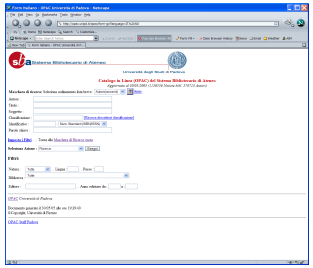
Digital representation of the document content

Content representation and management through different and sometimes new models of information and data management

External level:
Presentation logic
User interface

Intermediate level:
Logic of services

Internal level:
Data logic



Web Server

Gateway Program

OPAC

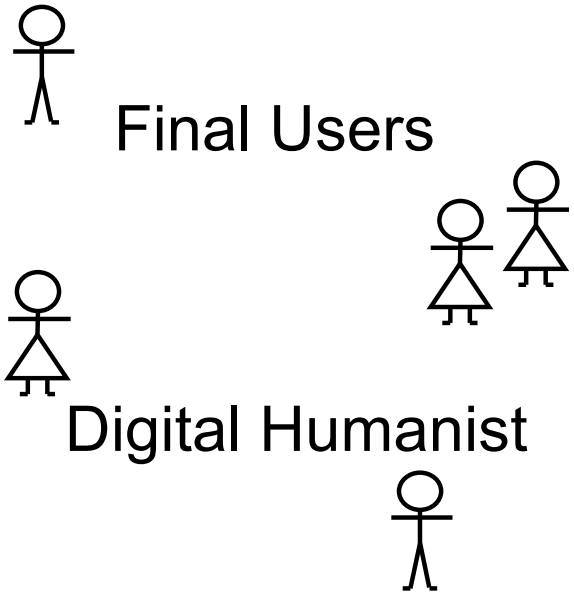
Database functions

Information Retrieval functions

Data, indexes, …

A posteriori analysis of the main actors

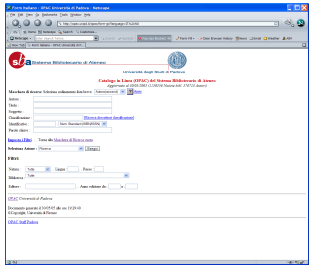
Main actors:



External level:
Presentation logic
User interface

Intermediate level:
Logic of services

Internal level:
Data logic



Web Server

Gateway Program

OPAC

Database functions

Information Retrieval functions

Data, indexes, …

The birth of the digital library area

The birth of the digital library area

- The Library Automation community realises the lack of computer science and engineering knowledge
- The area of Digital Library starts in those years as a new scientific area
  - In USA - Digital Libraries Initiatives (DLI-1 and DLI-2) of the National Science Foundation (from late 1993)
  - In Europe - A group of projects supported by the European Commission under the 4th, 5th and 6th Framework Programme named DELOS Working Group 1996-99, first DELOS Network of Excellence 2000-2003, and DELOS Network of Excellence for Digital Libraries 2004-2007
- It is an area of confluence: library and archive science, library automation, database management, information retrieval, human-computer interaction, information visualization, Web, …

Why we need to devise new models and systems

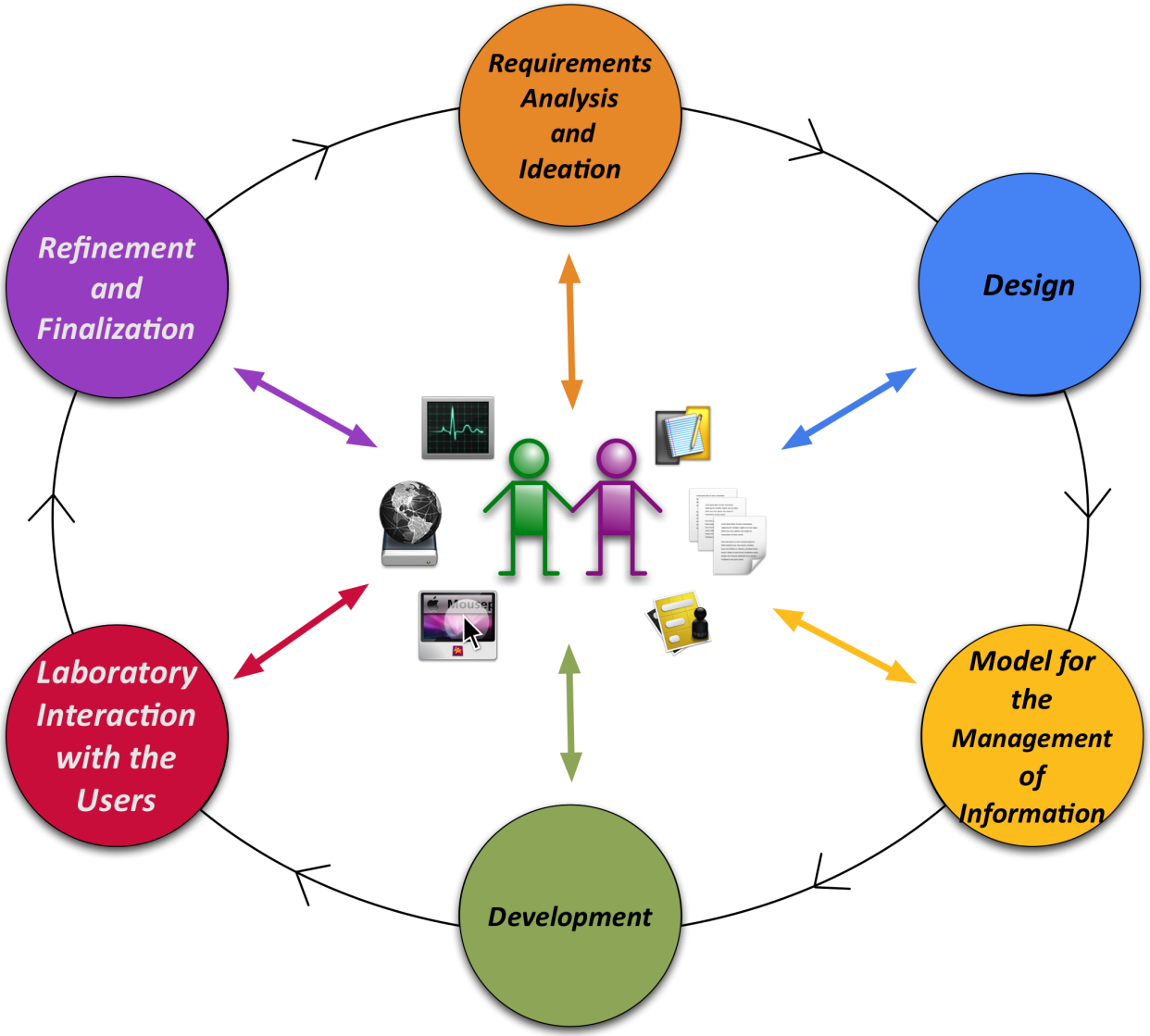
We may wonder why we need to devise new models and systems

- Our answer would be
  - because we want to consider aspects of reality that are different and more complex than those that were previously addressed
- As the aspects of reality that we want to address and manage become increasingly diversified and complex
  - we need to devise new methods and systems capable of dealing with and managing them.

What is behind a new aspect of digital cultural heritage?

- There is a model that is the basis for designing the necessary innovative information management system
- The richness of the digital cultural heritage domain requires that the process of envisaging the model is conducted by experts of computer and information science and of the specific domain of interest (history of art, history, …) together with final users

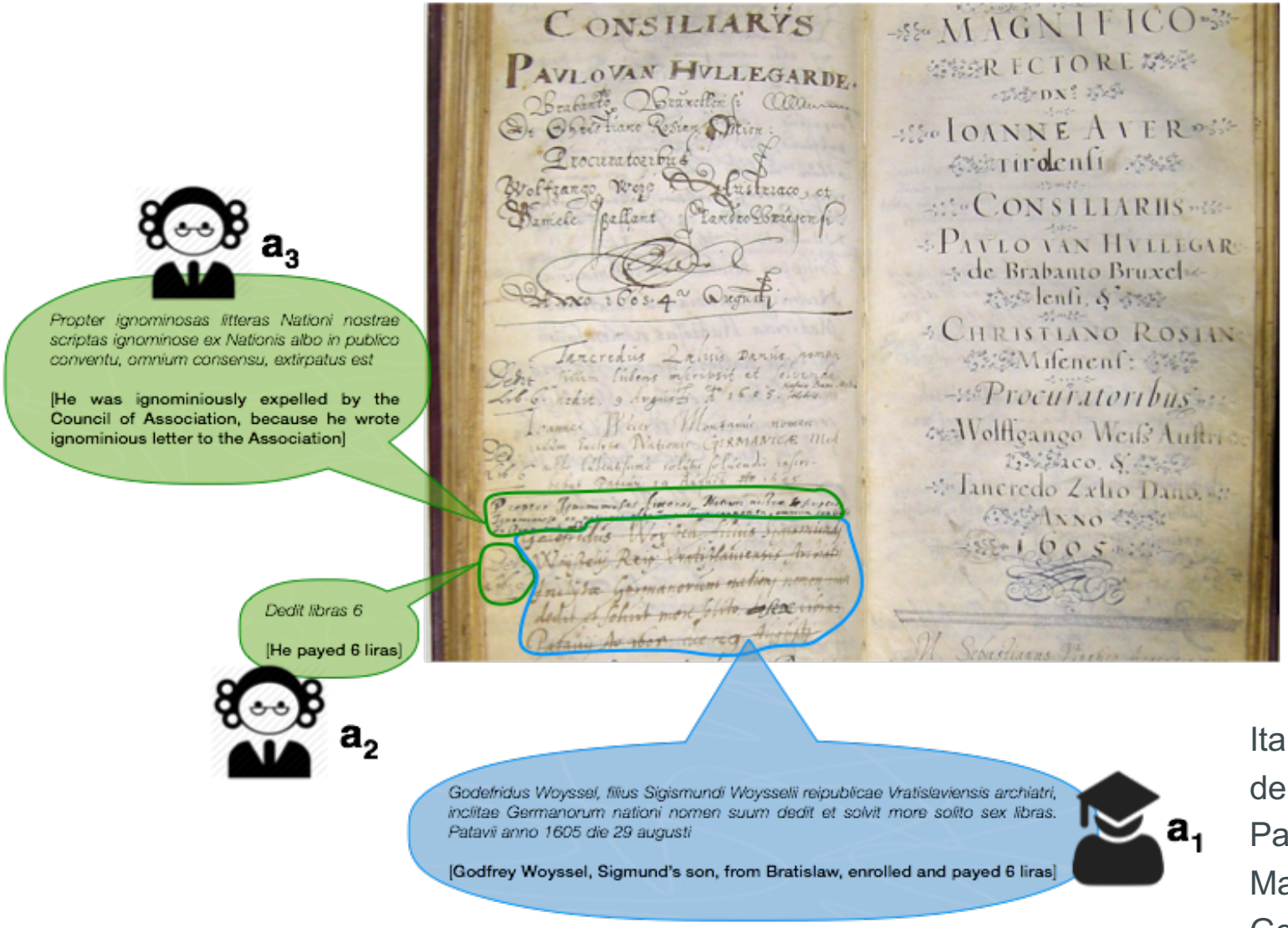
Process of Envisaging New Models



Maristella Agosti

Case study
Representation and management of digital annotations

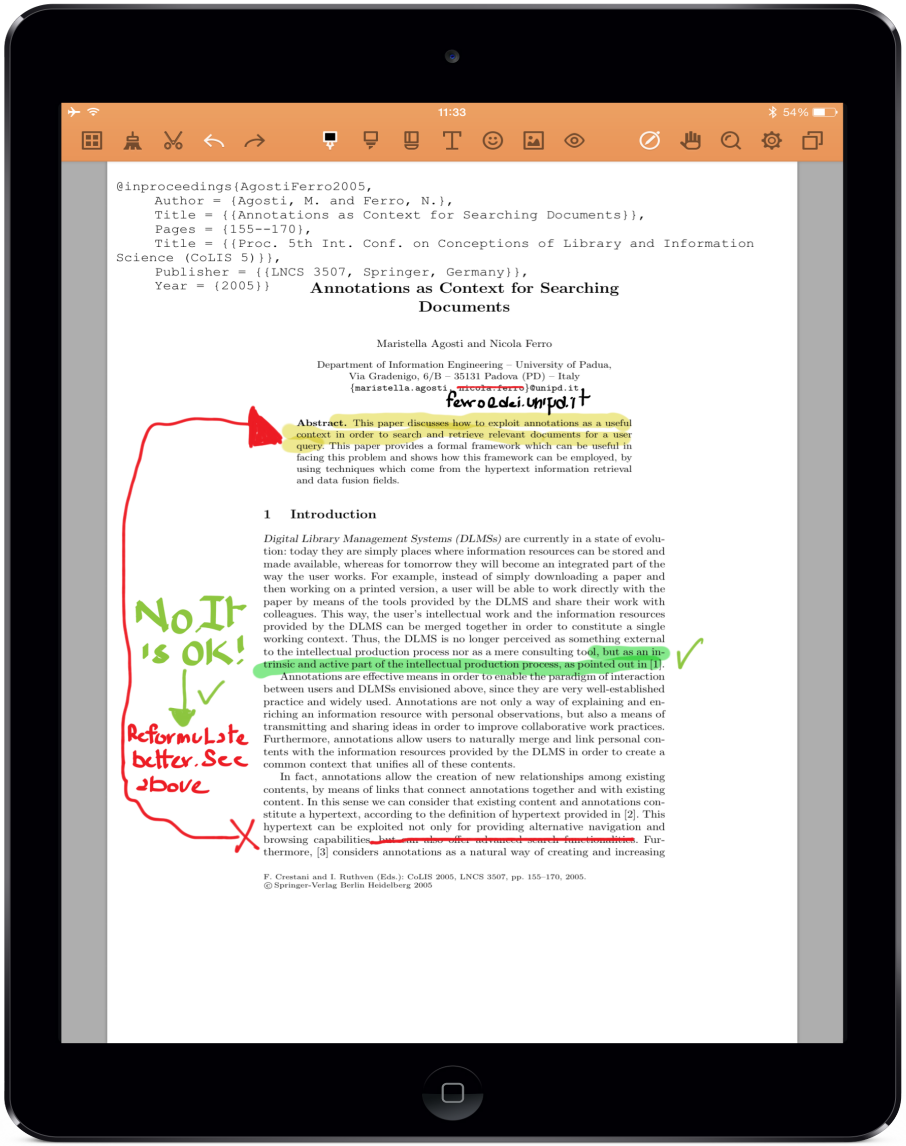
Historical Annotations: Padua University



Italia, Padova, Archivio dell'Università di Padova, Archivio antico, Matricula Nationis Germanicae artistarum, reg. 465, c. 69v

Key Issues - also on Today Tablets

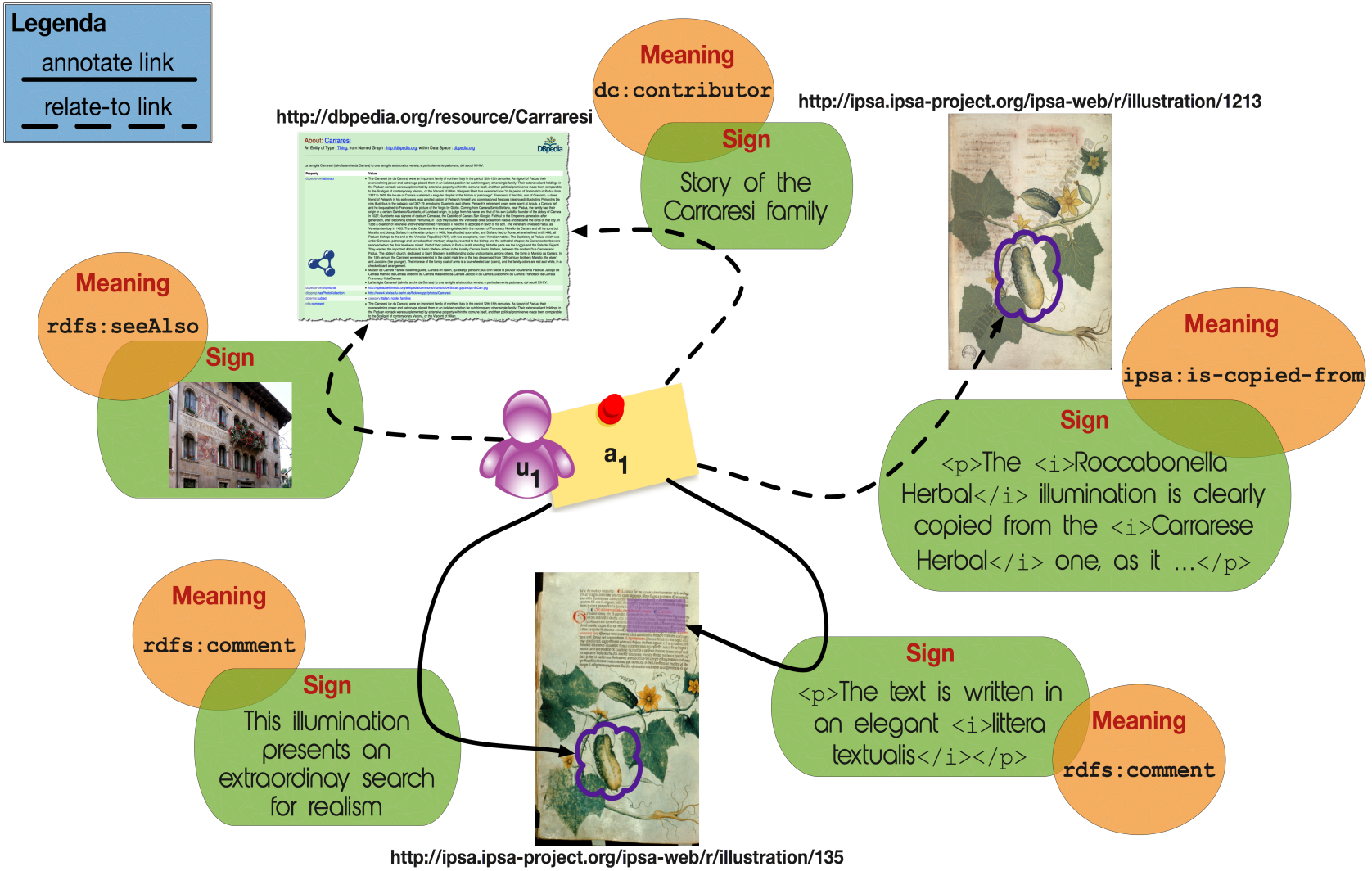
- Annotations are embedded in the annotated document
- Annotations semantics is not explicit or hard-coded
- Annotations are not related one to the other
- All the annotations have the same scope
- Annotations are not searchable



What to Expect from Annotations?

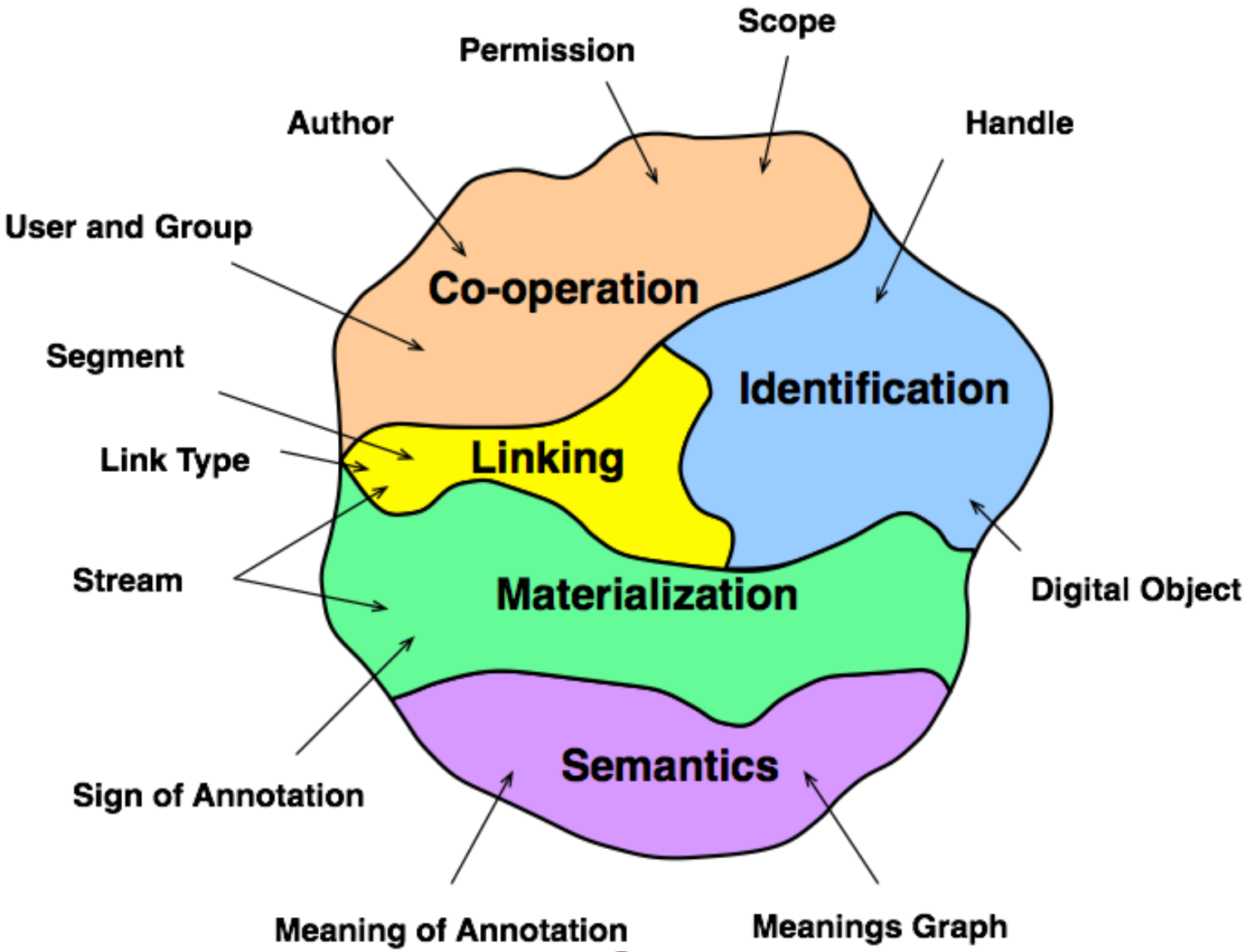
- A collaborative tool for user generated content
- Open, distributed and interoperable among different systems (the Web, digital libraries, digital archives, …)
- Able to engage research communities, foster their research work and transfer knowledge to students and the general public

User Annotations: the FAST Model

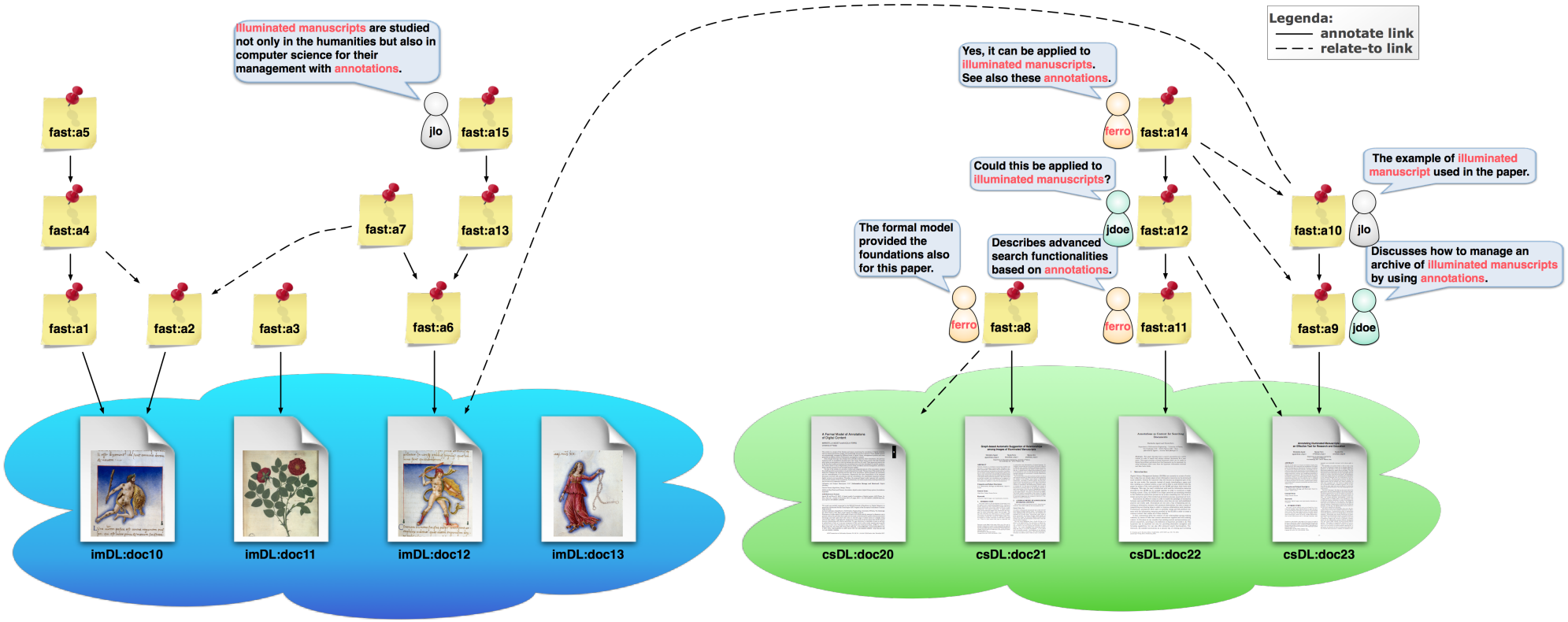


Maristella Agosti

Annotation Model Map



The Document-Annotation Hypertext

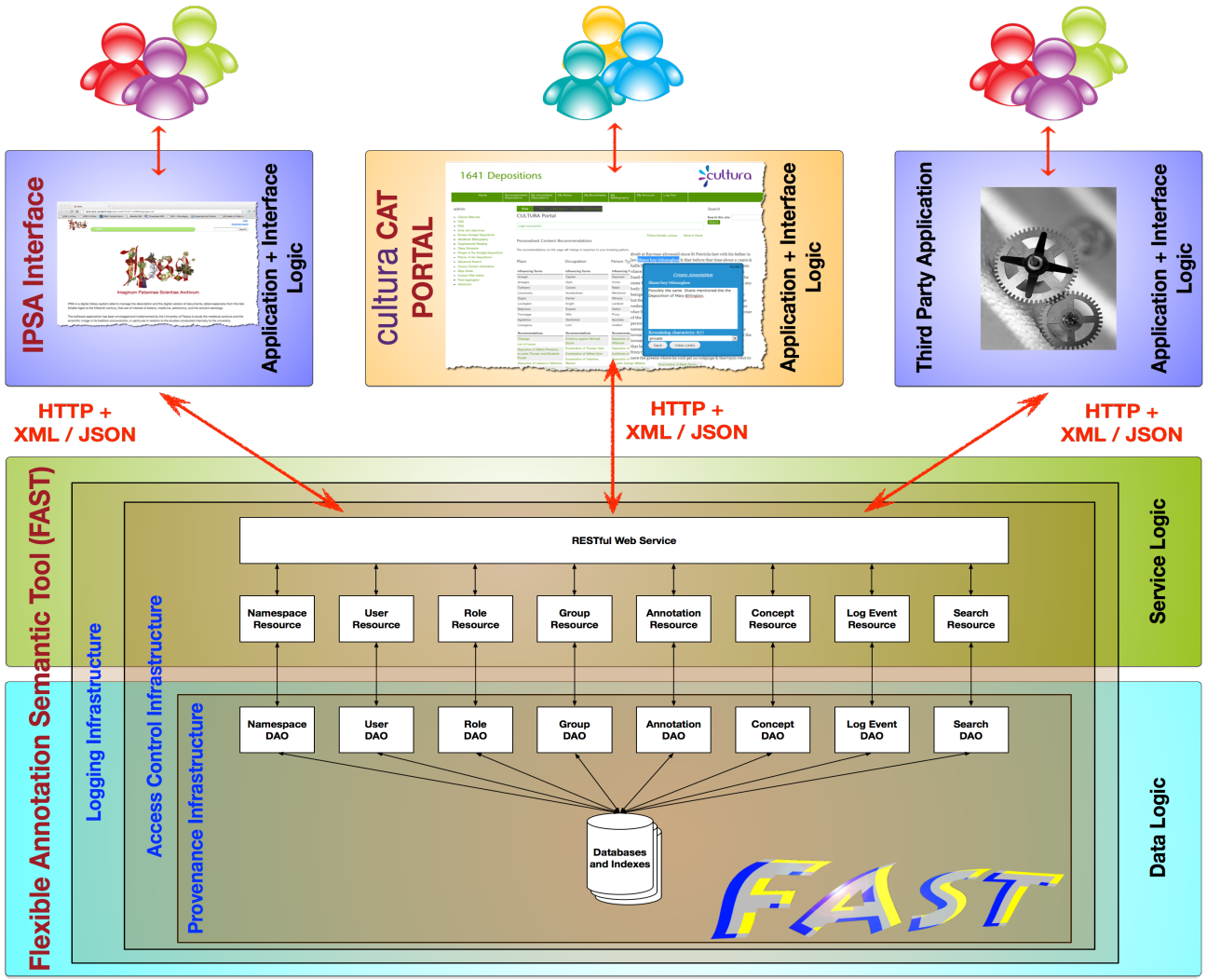


imDL: Illuminated Manuscripts Digital Library

csDL: Computer Science Digital Library

- Search by using annotations: exact match, best match, and navigation of the hypertext

FAST (Flexible Annotation Semantic Tool): a Tool to Innovate



Maristella Agosti

Considerations on the Annotations Effort

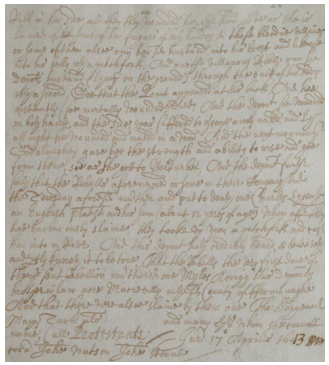
- Modelling, managing and searching annotations is a challenging research problem
  - 5 years to achieve a comprehensive formal model
  - 2 more years to achieve search over/by annotations
  - impact on the field - see W3C Open Annotation Collaboration - OAC, and only for Web annotations
- Developing a fully fledged annotation service is a demanding activity
  - 7 years to develop the FAST service and integrate it into several digital library systems in effective use

Case study
User Engagement in Digital Humanities

Digital Humanities: which artefacts to engage the user?

The EU STREP CULTURA project has used

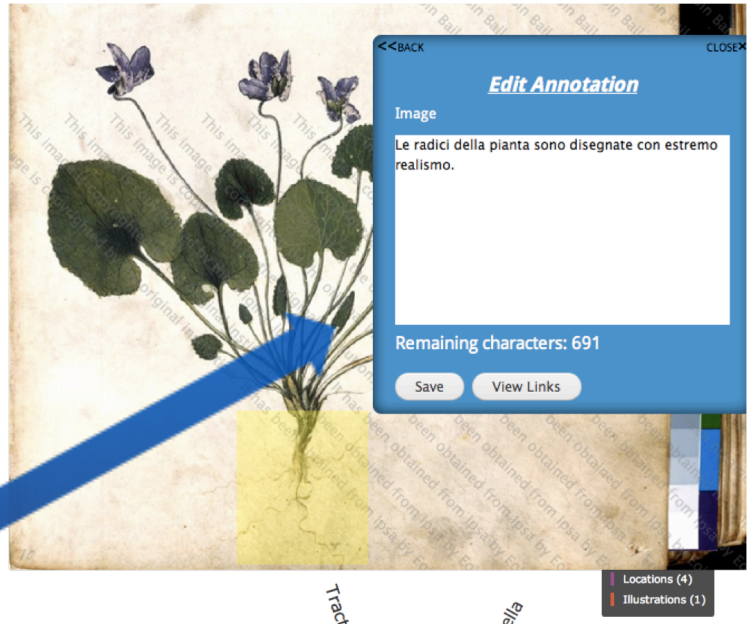
- two contrasting digital humanities artefact collections involving both textual artefacts and image-based artefact collections:
  - the 1641 Depositions
  - the *Imaginum Patavinae Scientiae Archivum* (IPSA) collection

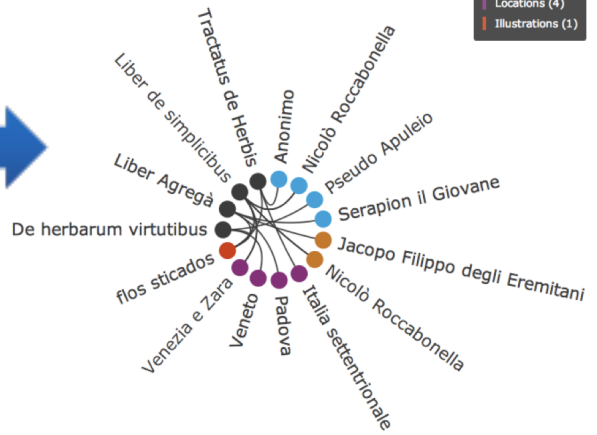


User Engagement through Multiple Services

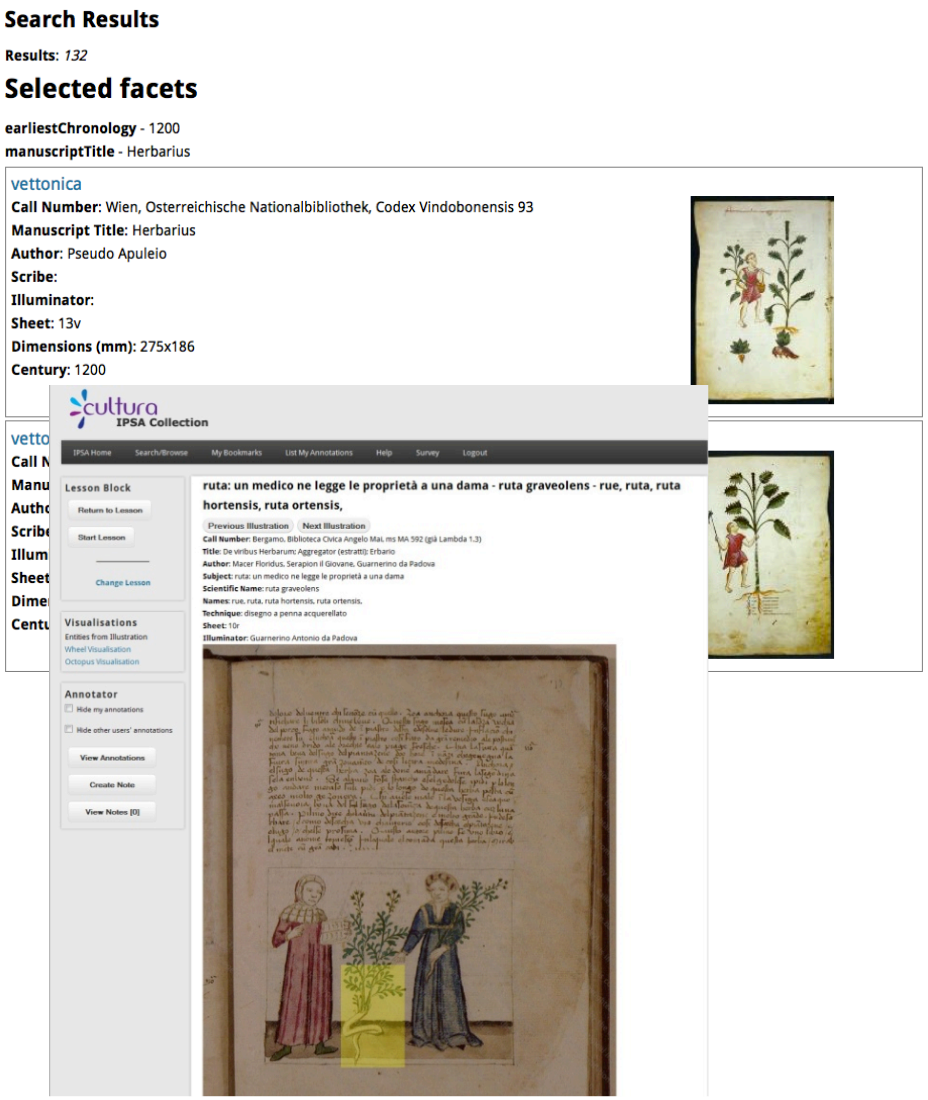
The VRE engages the users with multiple services, all accessed via the portal, that include:

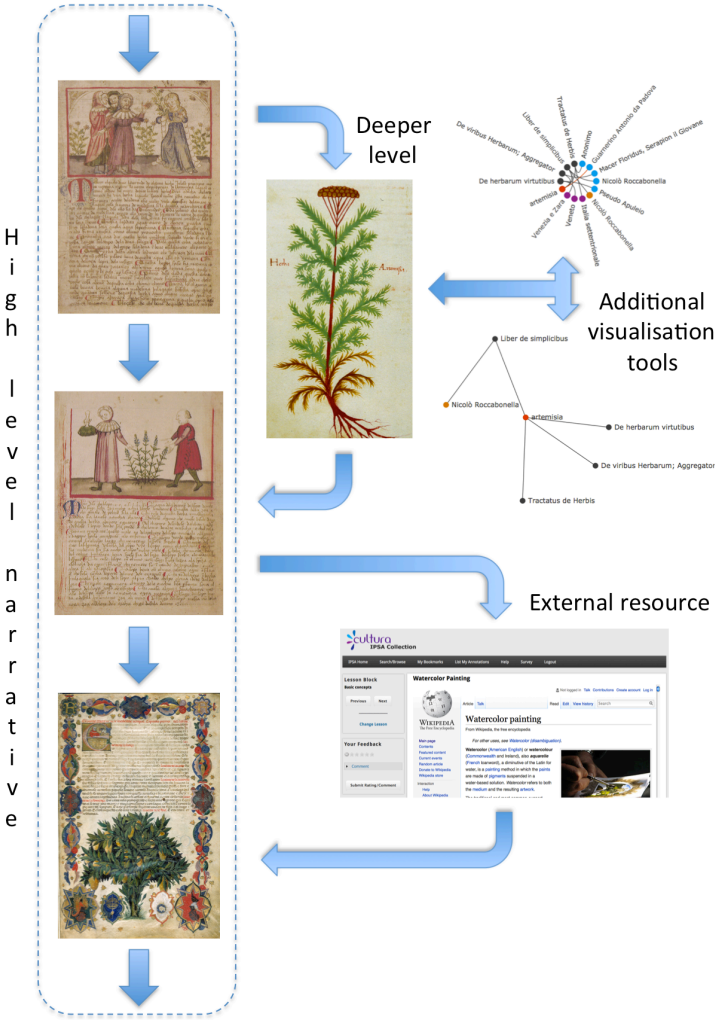
- Personalised search tools
- Faceted search tools
- Annotators
- Social network exploration tools
- Recommenders





User Engagement through Narratives





Maristella Agosti

The Confluence in Digital Humanities

It has been shown that the computer scientist, the digital humanist, and the final user efforts produce valuable models and applications when each of them brings his/her expertise taking into account the context and the expertise of the others

If this happens the result is similar to what is taking place at the confluence of the three rivers Inn (left), Danube (center), and Ilz (right) in Passau where the initial distinct colours merge together resulting in a new and augmented colour

Thank you for your attention

Questions?

References

- M. Agosti, M. Masotti, A.M. Moressa. An Online Public Access Catalogue (OPAC) for University Library End-users Using TRS: project and prototype. *Proc. Software AG's European Users' Conference*, Hamburg, Germany, 1990, Vol.1, Paper N.52
- M. Agosti, M. Masotti. Design of an OPAC database to permit different subject searching accesses in a multi-disciplines universities library catalogue database. In: N.J. Belkin, P. Ingwersen, A.M. Pejtersen (Eds.). *Proc. ACM SIGIR 1992,* Copenhagen, Denmark, ACM, 1992, 245-255
- S. Robertson, On the history of evaluation in IR. *Journal of Information Science,* 2008, 34(4):439-456
- S. Walker. Improving subject access painlessly: recent work on the Okapi online catalogue projects. *Program*, 1988, 22(1):21-31
- M. Agosti, R. Colotti, G. Gradenigo. A two-level hypertext retrieval model for legal data. In: A Bookstein, Y. Chiaramella, G. Salton, V.V. Raghavan (Eds). *Proc. ACM SIGIR 1991*, Chicago, USA, 316-325
- M. Agosti, G. Bonfiglio-Dosio, N. Ferro. A Historical and Contemporary Study on Annotations to Derive Key Features for Systems Design. *International Journal on Digital Libraries*, 2007, 8(1):1-19
- M. Agosti, N. Ferro. A Formal Model of Annotations of Digital Content. *ACM Transactions on Information Systems (TOIS)*, 2008, 26(1):3:1-3:57
- M. Agosti, O. Conlan, N. Ferro, C. Hampson, G. Munnelly. Interacting with Digital Cultural Heritage Collections via Annotations: The CULTURA Approach. In S. Marinai, K. Marriot (Eds). *Proc. 13th ACM DocEng*, ACM Press, New York, USA, 2013, 13-22
- O. Conlan, A. Staikopoulos, C. Hampson, S. Lawless, I. O'Keeffe. The narrative approach to personalisation. *The New Review of Hypermedia and Multimedia*, 2013, 19(2):132-157
- M. Agosti, N. Orio, C. Ponchia. Promoting user engagement with digital cultural heritage collections. *International Journal on Digital Libraries,* 2018, https://doi.org/10.1007/s00799-018-0245-y