Unione europea
Fondo sociale europeo

Regione Liguria

Ontology-driven Visual Browsing of Historical Industrial Archives

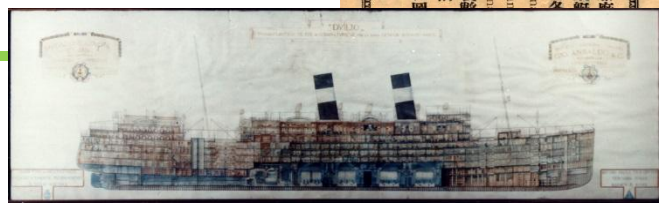
Monica DeMartino
Marina Monti
Simone Pastorino
Chiara Rosati
CNR IMATI

imati

EUROMED 2014
03 - 08 November 2014

Summary

- Project R.I.C.E.R.C.A. overview
  - Business Archives overview
  - Resource access issue
- Browsing and search Approach
  - The Semantic framework
- Outcomes and use case example



Project R.I.C.E.R.C.A.

R.I.C.E.R.C.A. (2013-2015)

"Risorse Industriali Culturali: Estrazione e Raccolta dei Contenuti per Accessibilità", Italian project funded by European Commission

- Project Aim: identify new strategies in the management and access to business historical archive.
- Project Application: The Italian business industrial archive of Ansaldo Foundation.
- Objective:
  - Make resources available and easily accessible — *Heterogeneous users Different targets*
  - To define the Archive management model — *Resource preservation Archival documents description Archive enhancement*
  - To design a web platform for archives exploration — *Semantic information Interactive navigation tools*

Business Archives

What is an historical business archive?

- A hierarchical structured collection of heterogeneous materials produced by a company during its activity. The stored resources are grouped in series.
- It is important for the development of a country
- The possibility to easily access to them remain a challenge in most countries.

Example of the most important Italian business archive

- The Industrial archive of Fondazione Ansaldo

Example of Industrial Business Archive
The Industrial Archive of Ansaldo Foundat

MISSION

- Preserve and enhance the Historical Business Archive produced by industrial-economic subjects of Genova and Liguria in XIX-XX century

ARCHIVAL HERITAGE

- Over 15 km. of archival documentation, 400.000 original photos, 5000 original movies, industrial artefacts collections, Library specialized in industrial-economic issues, an original photographic laboratory

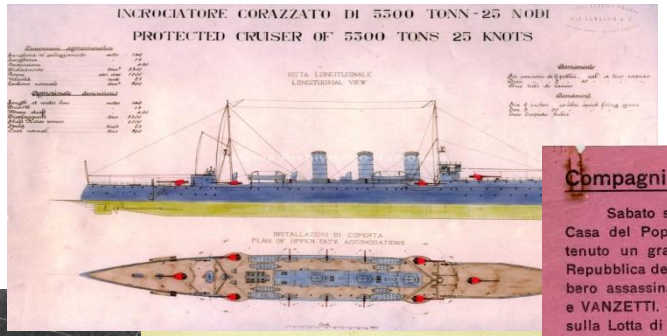
HISTORY

- Founded in 1980 as Archivio Storico Ansaldo, in 2000 became Ansaldo Foundation and in 2007 Ansaldo Foundation - Finmeccanica Group

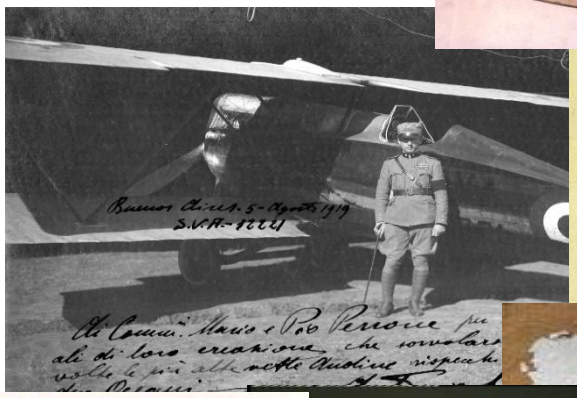
PARTNERSHIP

FondazioneAnsaldo

Photo



Paper



Book

Movie

Industrial Artifacts

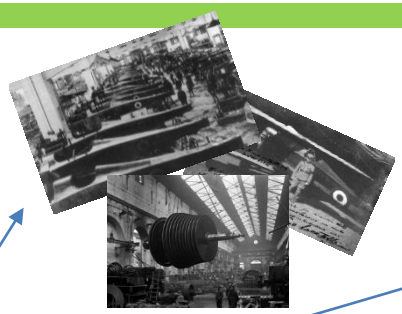
Business Archives
Access issue

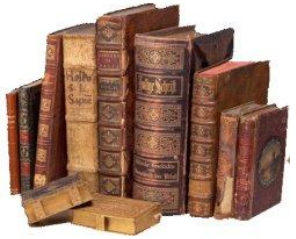
Professional user

Common user

The Archivist

Knowledge
of the structure of
the archive





large amount of resources
the heterogeneity of the data

Business Archives
Access issues

❑ Issue to search in a business historical archive?

- To deal with large amount of resources stored and to the heterogeneity of the data.
  - Needs of an uniform resource descriprion → Metadata

- Needs of knowledge of the structure of the archive,
- The online archive access there is no the archivist's support
  - Need of the archivist knowledge formalization

- User have not competencies for browsing an archive and search content
  - Needs of advanced tool to easy the archive browsing and search

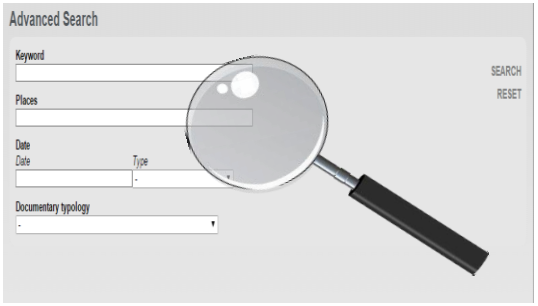
Semantic Framework

for archive semantic exploration through descriptive concepts of the domain.

The semantic framework

The framework components:

- A knowledge model formalized in an ontology, which includes categories of descriptive concepts and series
- A visual browsing exploration module, which exploits the ontology to facilitate the archive data browsing and searching
- A keyword search module based on metadata.

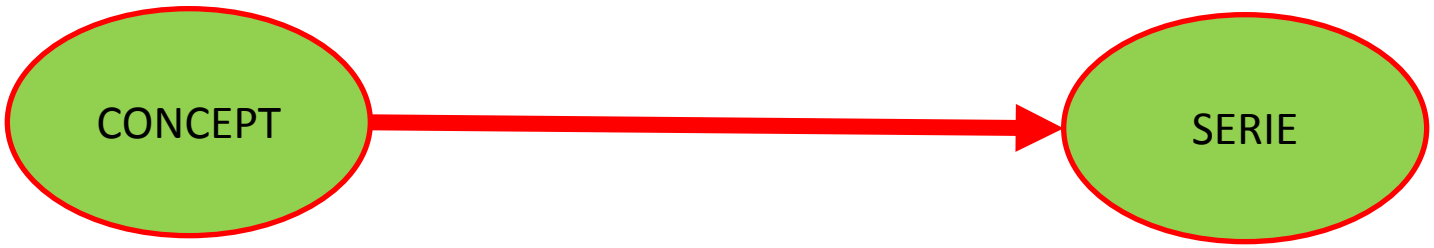


The semantic framework:

- Define a knowledge model to support the user exploration of the archive content, by identifying categories of descriptive concepts for industrial archives and their relationship with different categories of archive series.

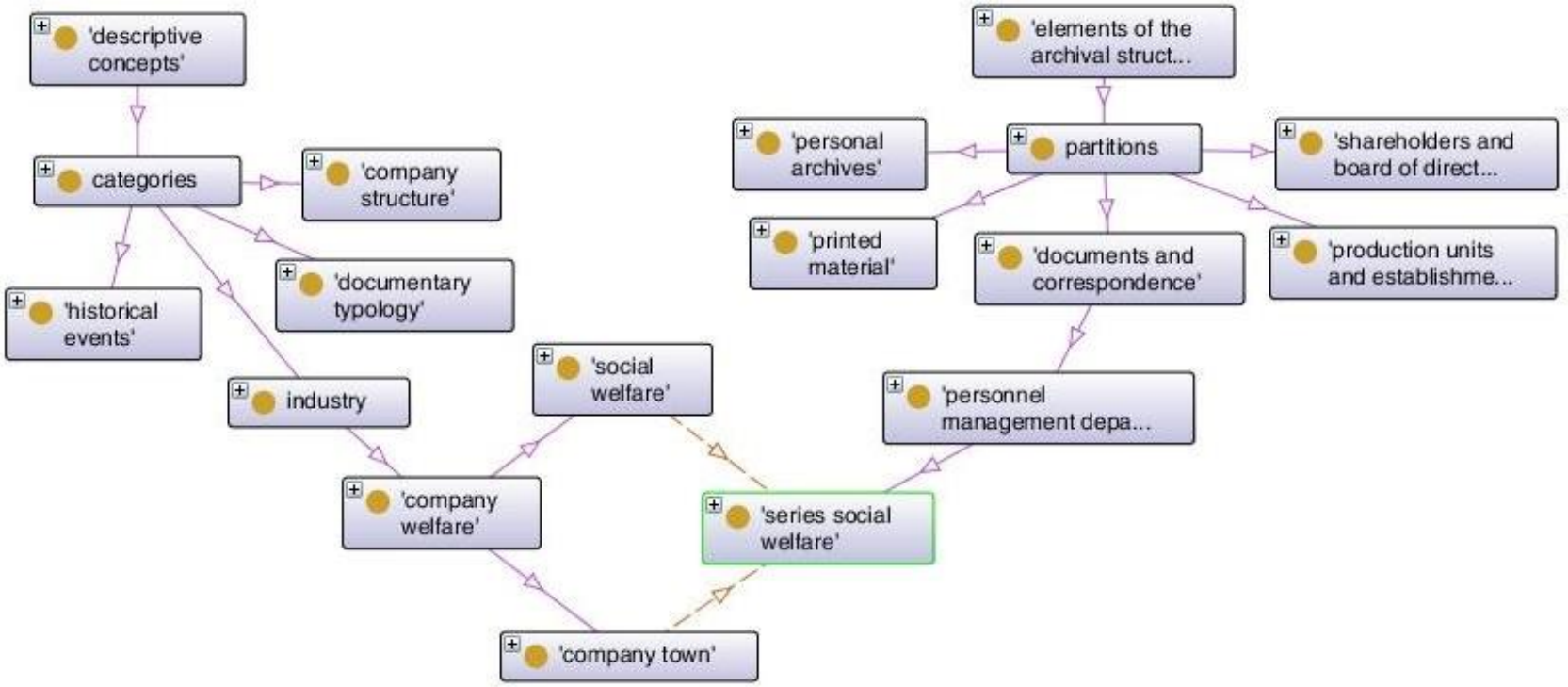
- Formalization of this knowledge model in an ontology, edited in Protégé 4.3, according to the OWL 2.0 language.

The idea is to connect categories of concept to the series that may contain information about them

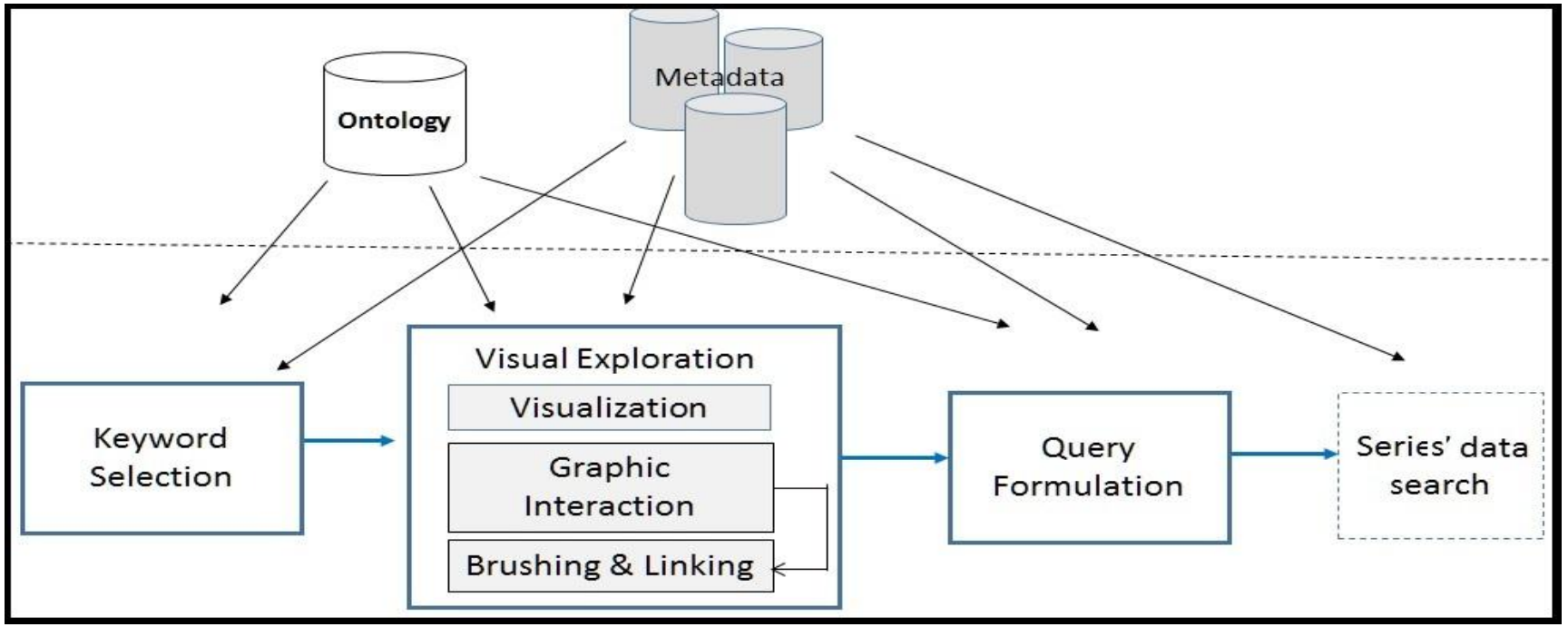


The semantic framework:

Connection between concepts, that represents search topic, and the series of Ansaldo Foundation where user can find materials about the topic required



The semantic framework:



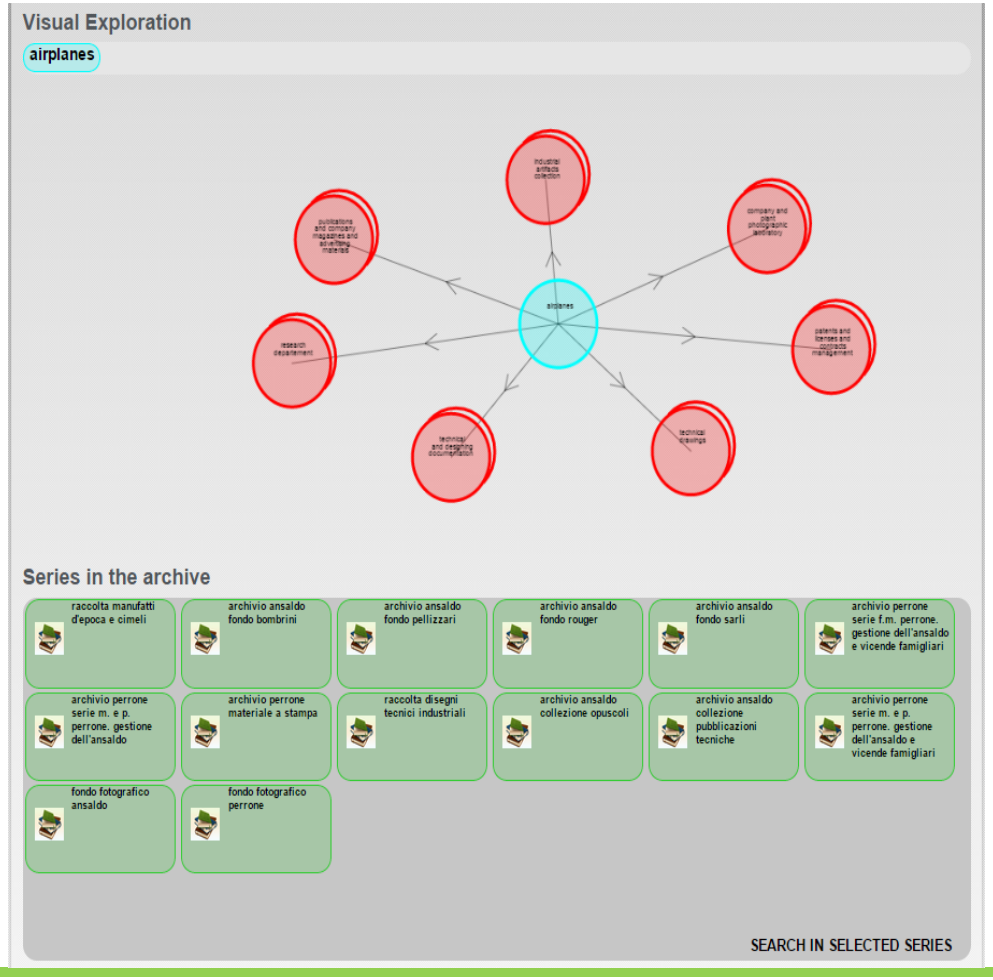
- User select a keyword (for a concept) from an autocomplete list
  - Visua exploration asks the user to refine the concepts (if necessary), and connect the concept to the series of Ansaldo Foundation
    - User search data in the selected series through metadata

The semantic framework:

Currently two visualizations are available:

The ontology graph visualization of the keyword/concepts referring to the series content

The list visualization of the series linked to the specific concept



The ontology graph visualization allow the user to better specify the topic of interest.

The list visualization of the series give the user an overview of the series where he can find materials about the chosen topic.

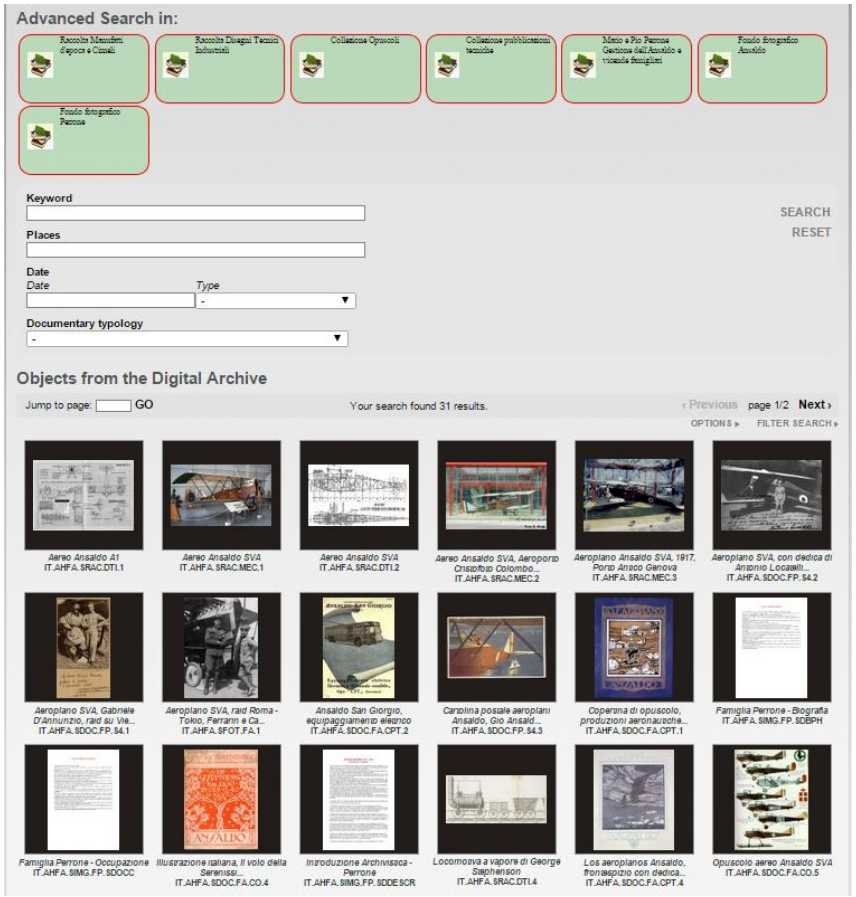
The semantic framework:

search restricted to the selected series

Search is based on metadata by

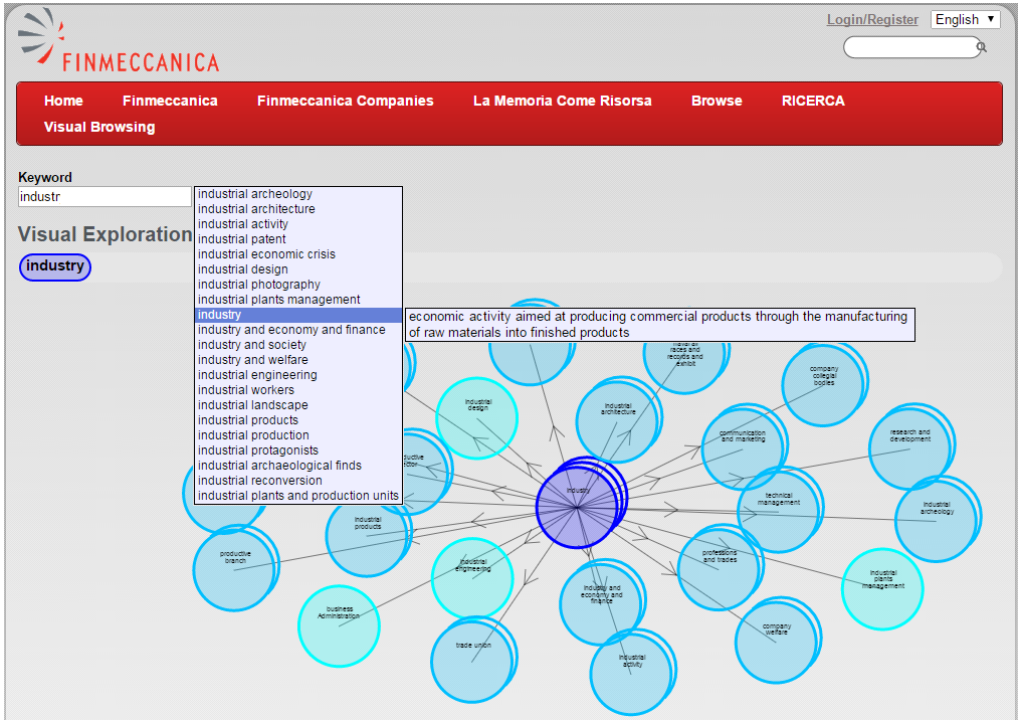
- the resource title
- the description of the resources
- the geographical position
- the associated dates
- the documentary typology

Identifyed Resource



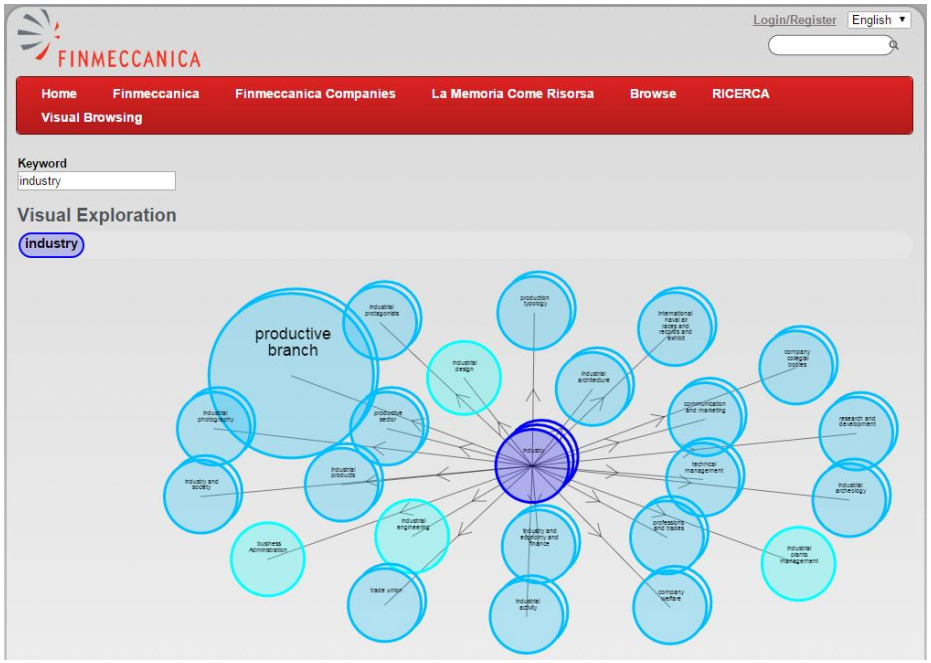
Illustrative Example Browsing

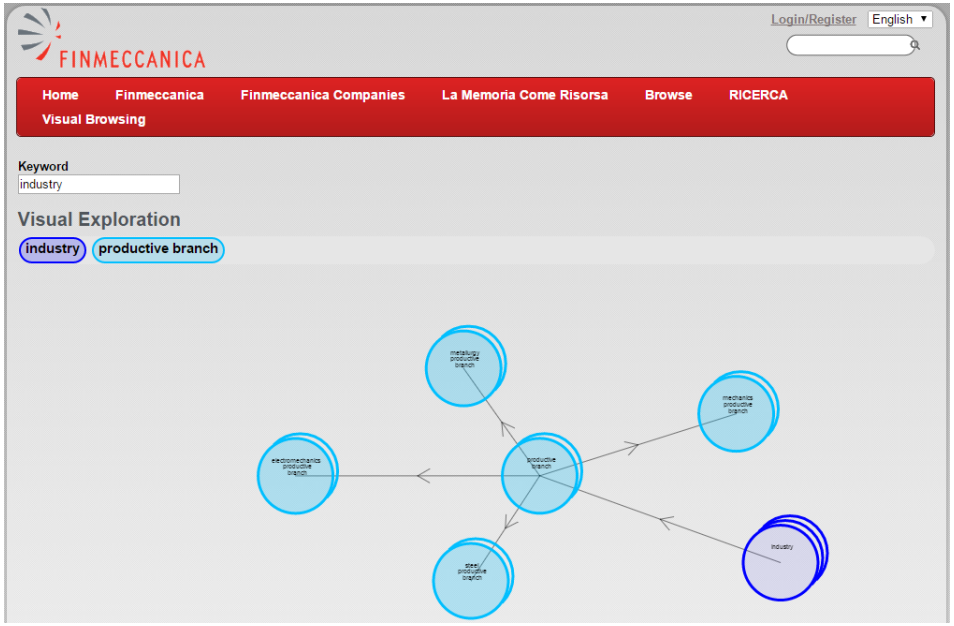
- User type a keyword in the search field
- choose a topic from the autocomplete suggestions.
- A brief description of the selection help during the choice.
- If the concept is too much generic, the system display a graph with more specific related concepts.
- User has to specify the topic of his search.



Illustrative Example Browsing

User specify the concept selecting from the graph nodes

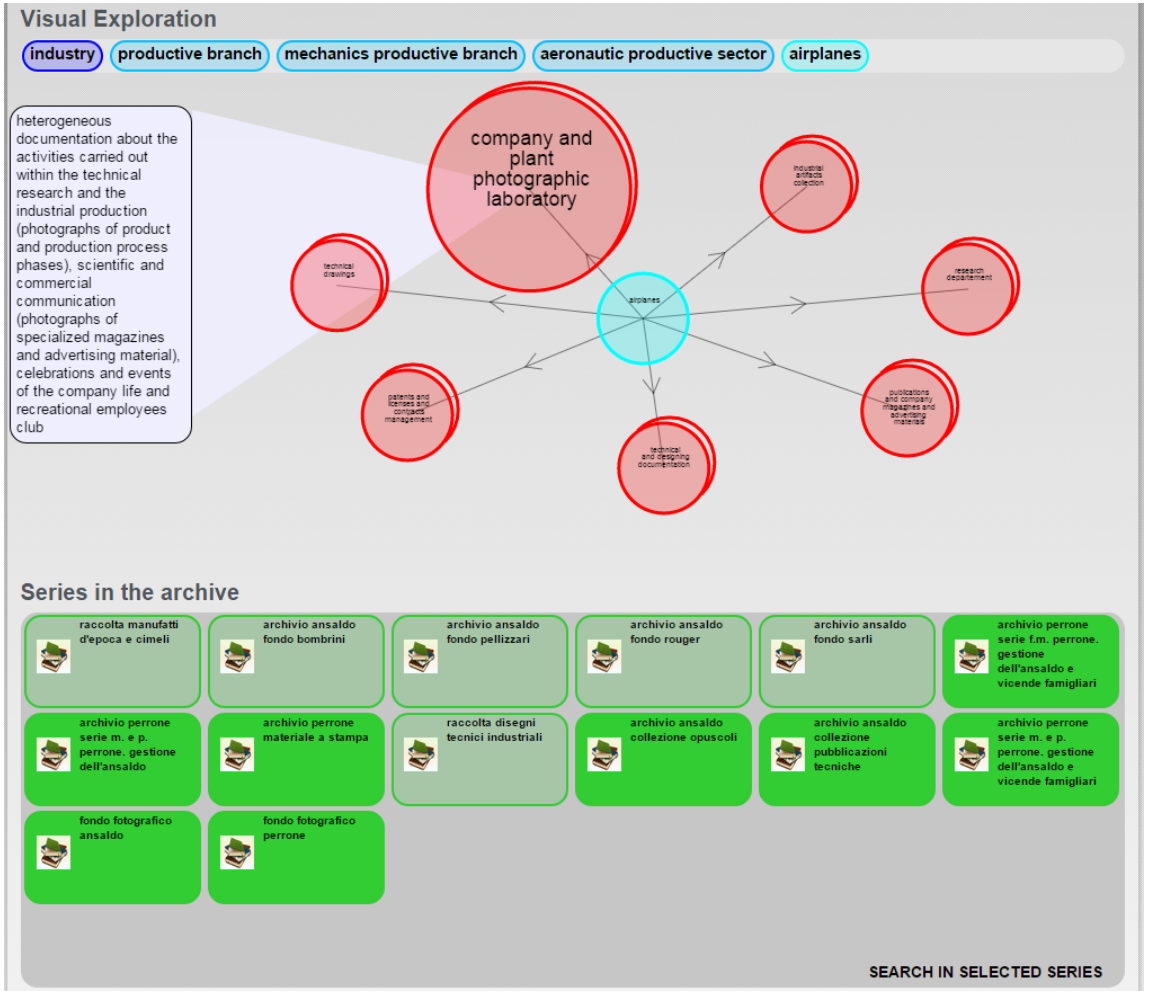




Illustrative Example Browsing

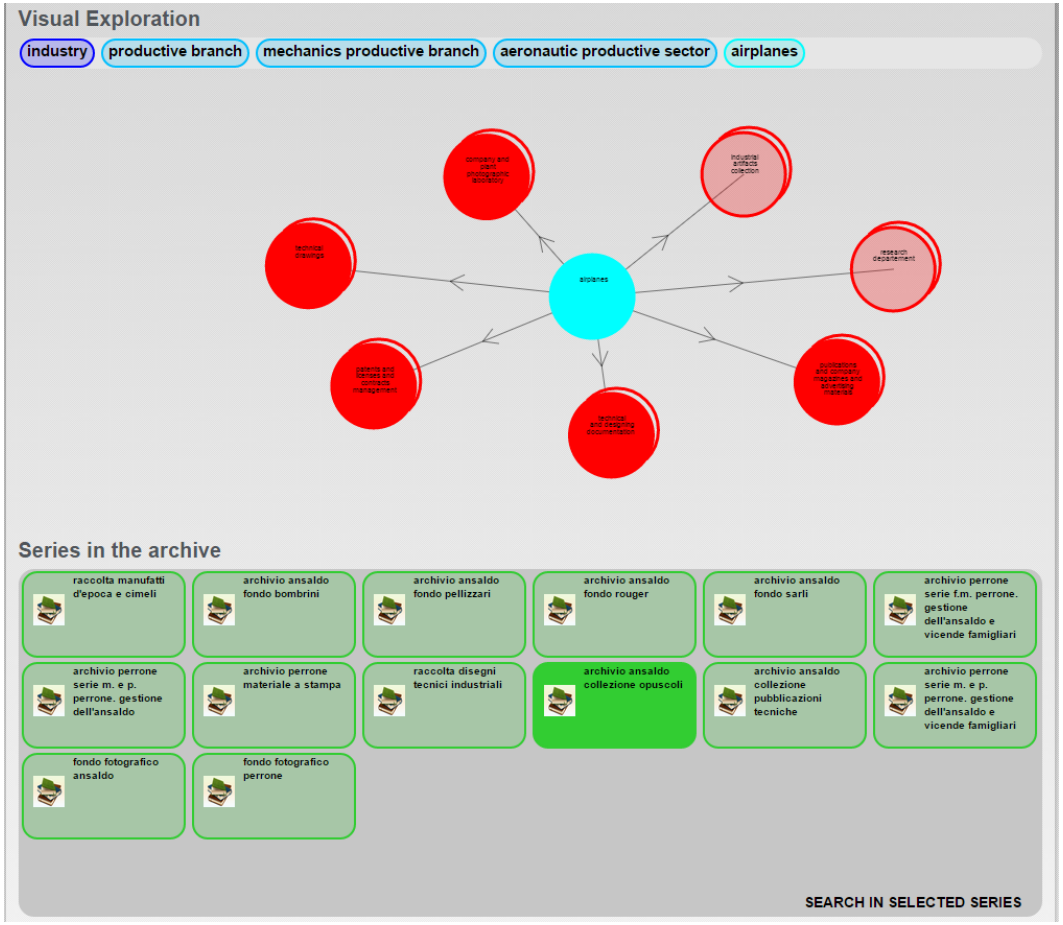
Browsing the graph user find a quite specific concept like “airplanes”.

- Red nodes indicates the classes of series that can be reached,
- Green rectangles are the series of the archive that can be reached from the highlighted node.
- While highlighting a none, a description appear on the left.



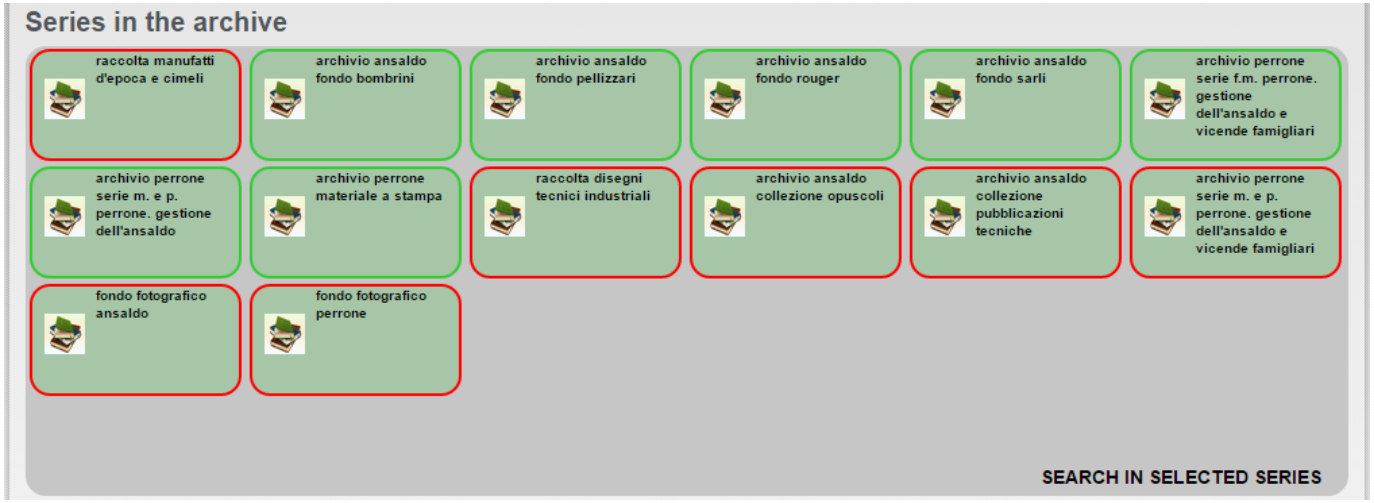
Illustrative Example Bowsing

If the user move the mouse over one or more series in the lower view, the related series categories are highlighted in the graph view.

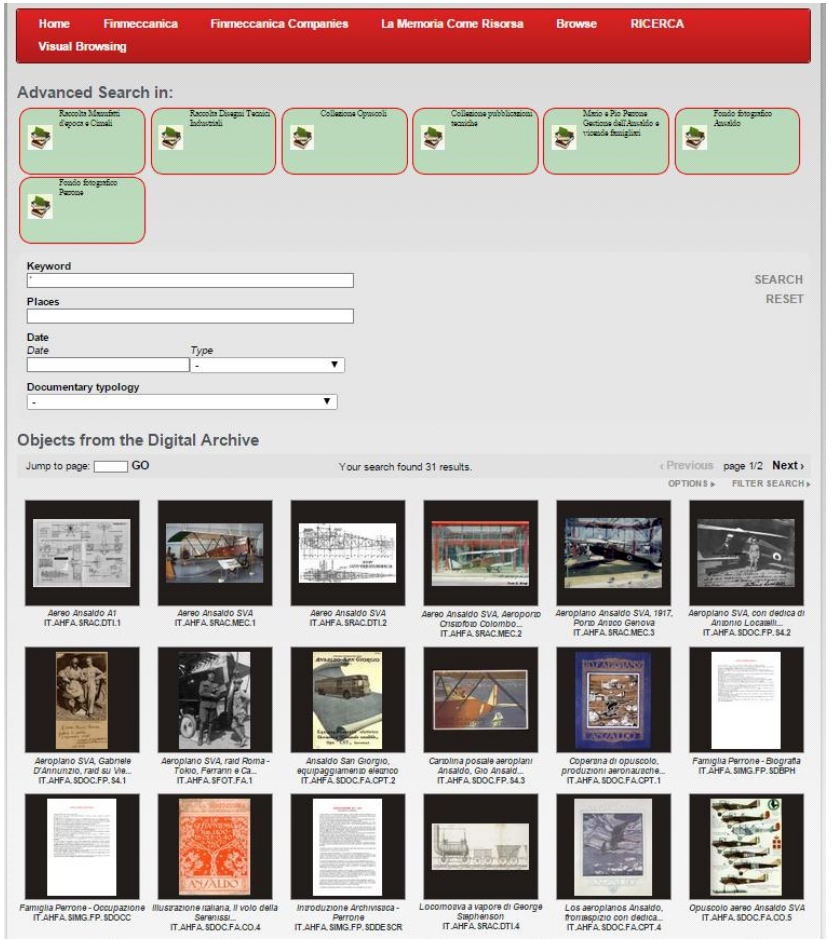


Illustrative Example Browsing

- User select (and deselect) any number of the proposed series by clicking, and perform a search restricted to them.
- The selected ones have red border.
- Once selected the series, user can click on “SEARCH IN SELECTED SERIES” and open a page for the advanced search.



Illustrative Example Search



At the top, the series in which the user is searching

In the middle the field for metadata search.
It is possible search for keyword, place, date and documentary typology.

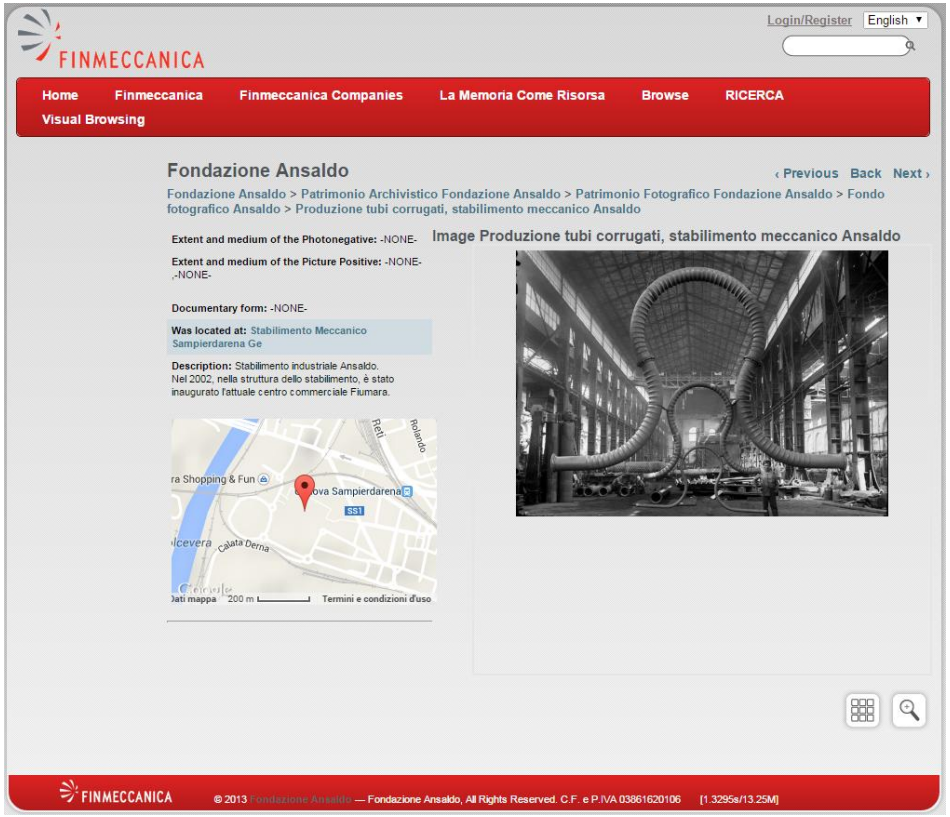
At the bottom, the found resources.

Illustrative Example
Geographic based search

Search results can be mapped

Illustrative Example search

The page dedicated to the resource



Conclusions

The Semantic Framework

- Uniform description (metadata) of the Industrial Archives of Ansaldo Foundation to easily find resources of interest
- Visualization to facilitate the user in the exploration of the archive
- Knowledge model to exploit part of the archivist's knowledge from the web,
- **http://ricerca.ge.imati.cnr.it**

Open issue and on going activity

- More efforts are needed in the definition of common terminology to describe the archive
- Design and identification of a path of resources according with a selected topic.

R.I.C.E.R.C.A

THANKS!

http://ricerca.ge.imati.cnr.it

Contact person

CNR IMATI

demartino@ge.imati.cnr.it

monti@ge.imati.cnr.it

pastorino@ge.imati.cnr.it

rosati@ge.imati.cnr.it

FondazioneAnsaldo

imati

softeco sismat information technology

All above pictures reproduced courtesy of Fondazione Ansaldo. Copyright © All rights reserved