



# Ontology-driven Visual Browsing of Historical Industrial Archives

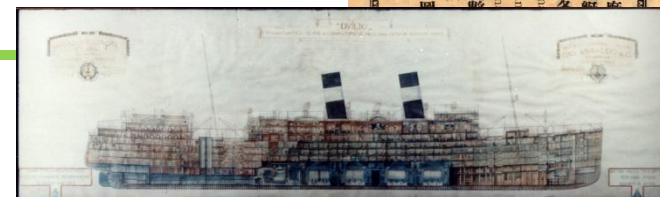


Monica DeMartino  
Marina Monti  
Simone Pastorino  
Chiara Rosati  
CNR IMATI



# 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

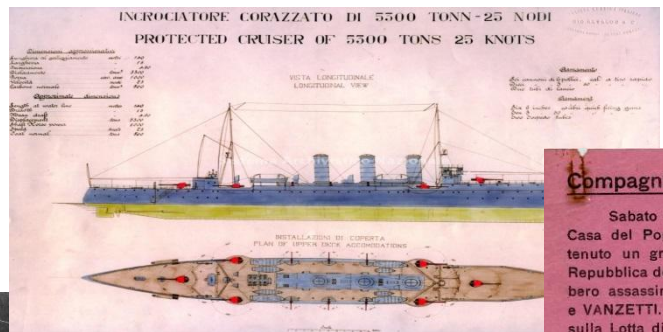
- Founded in 1980 as Archivio Storico Ansaldo, in 2000 became Ansaldo Foundation and in 2007 Ansaldo Foundation - Finmeccanica Group

### PARTNERSHIP





# Photo



# Paper



**Compagni Lavoratori!**

Sabato sera alle ore 20,30 nell'atrio della Casa del Popolo (già Palazzo d'Istruzione) sarà tenuto un grande Comizio di Protesta contro la Repubblica del Dollaro che il 1° Novembre ha osato assassinare i due nostri Compagni e VANZETTI. - Farleranno oratori di tutti sulla Lotta di Classe.

Partito Comunista d'Italia. - Partito Socialista Federato. Gioventù Comunista - Unione Anarchica. Camera del Lavoro Confed. - Camera del Lavoro.

Per entrare al Comizio occorre presentarsi muniti di una delle Organizzazioni.

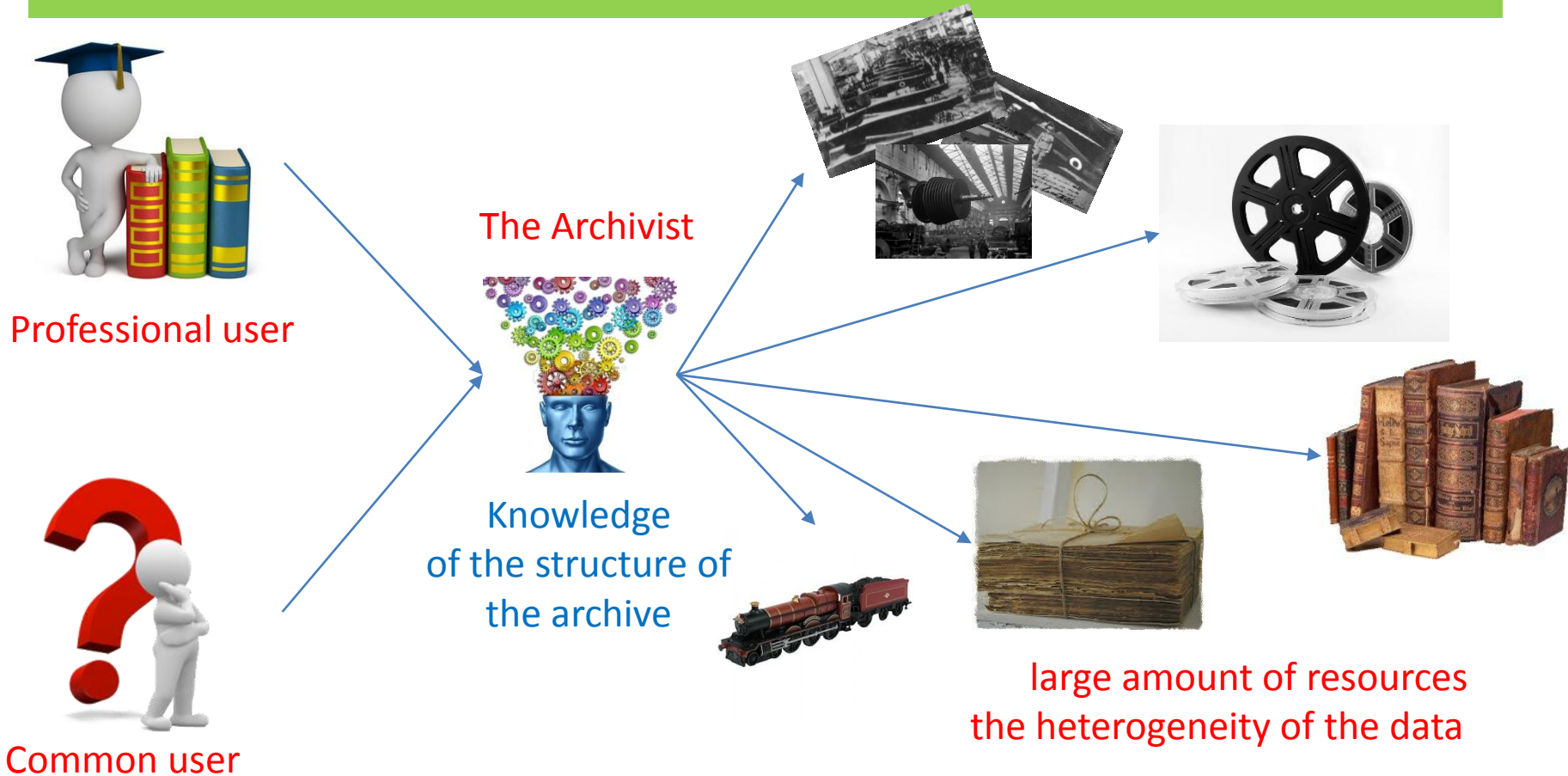


# Book



# Movie

# Industrial Artifacts





# 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 description → 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

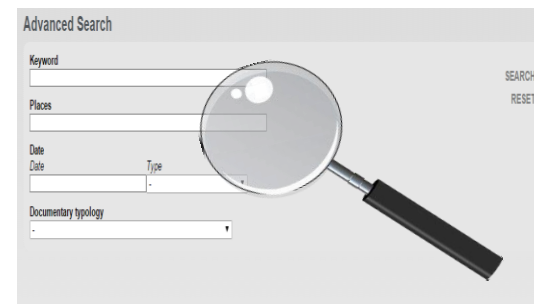
for archive semantic exploration through descriptive concepts of the domain.



# 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.



Advanced Search

Keyword

Places

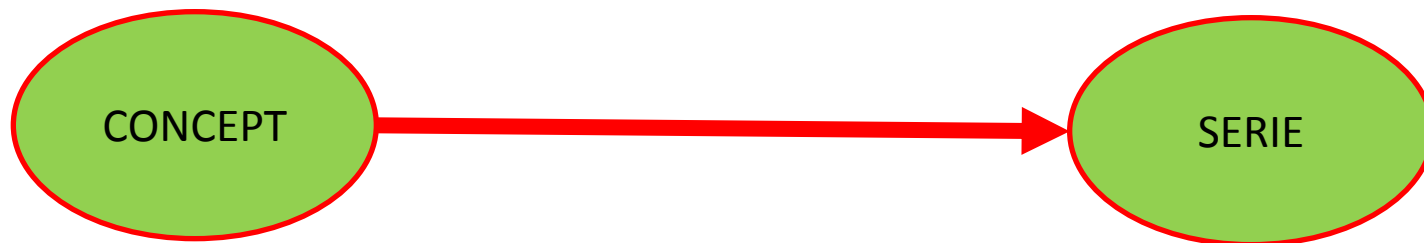
Date  Type

Documentary typology

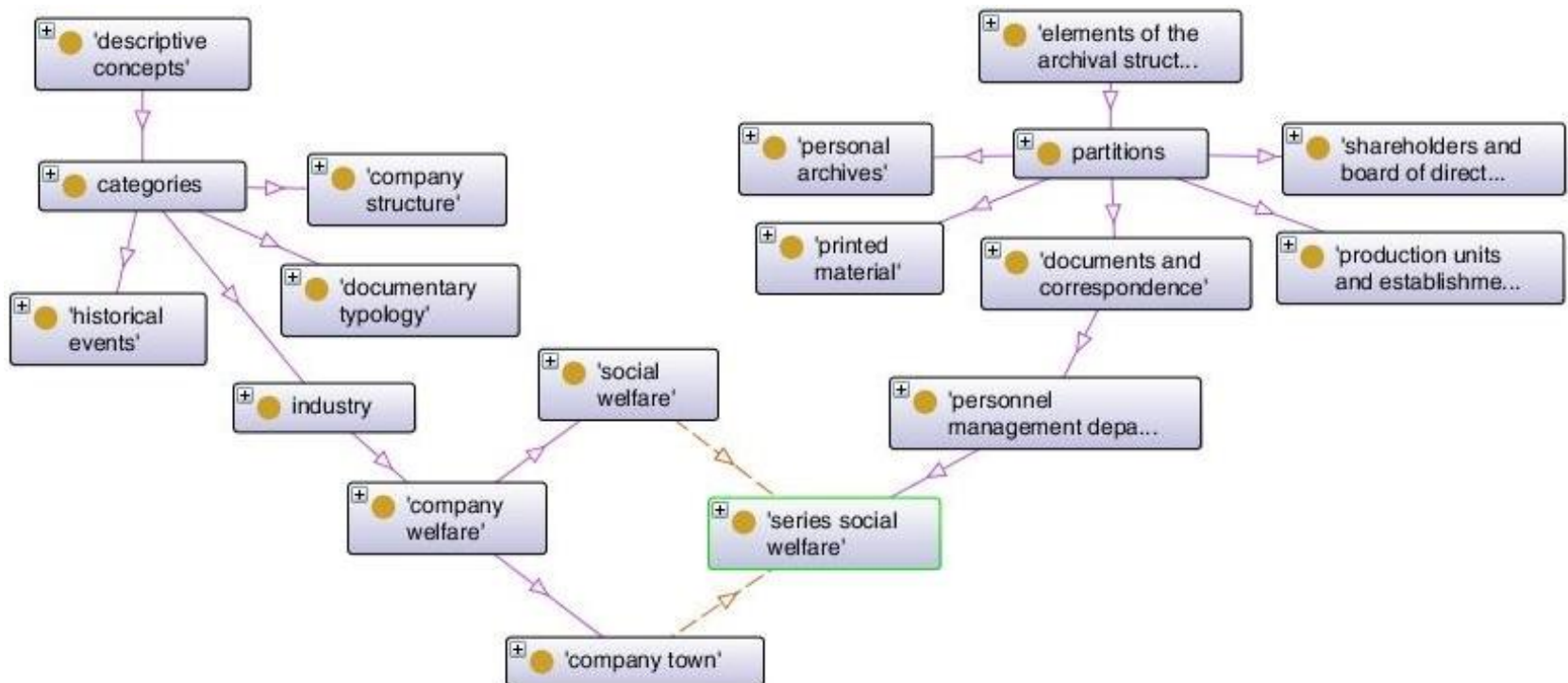
SEARCH RESET

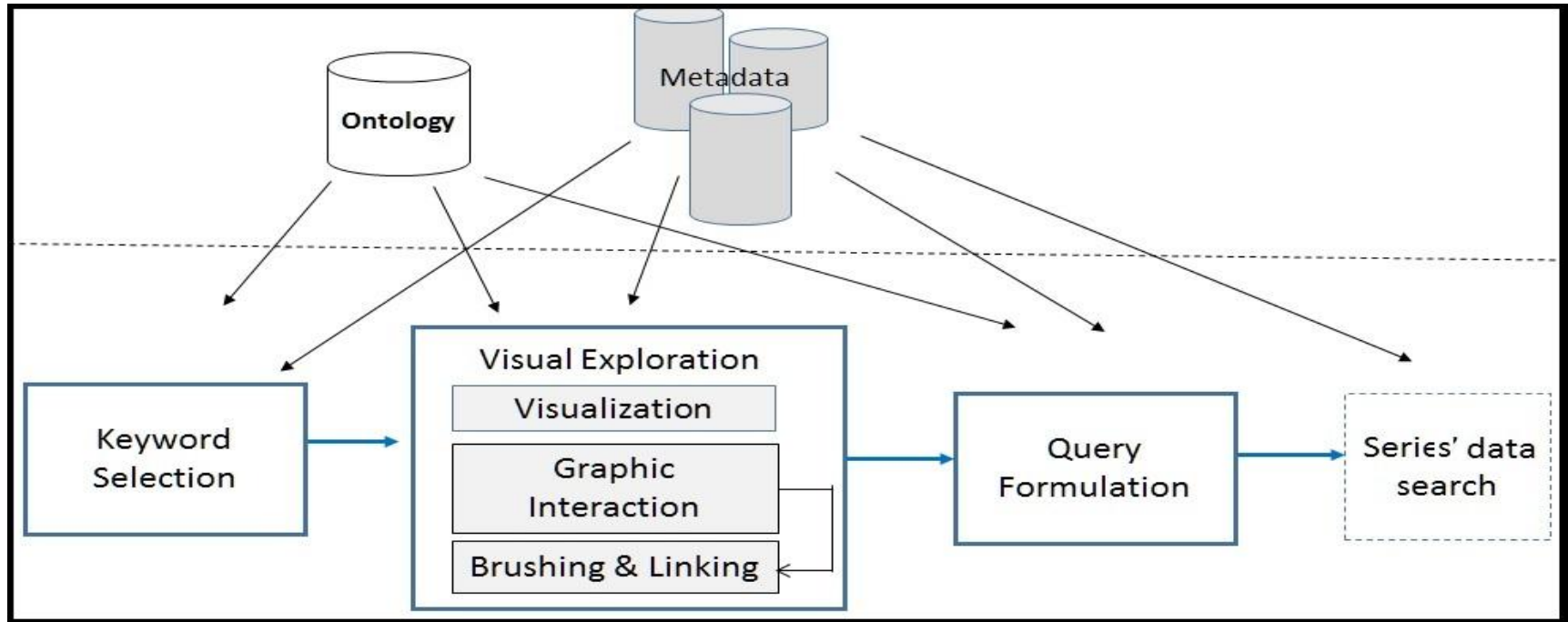
- Define a **knowledge model** to support the user exploration of the archive content, by identifying **categories of descriptive concepts** for industrial archives and their **relationship with different categories of archive series**.
- Formalization of this knowledge model in an ontology, edited in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



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





- 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

Currently **two visualizations** are available:

Visual Exploration

**airplanes**

Series in the archive

raccolta manufatti d'epoca e cimeli	archivio ansaldo fondo bombrini	archivio ansaldo fondo pelizzari	archivio ansaldo fondo rouger	archivio ansaldo fondo sari	archivio perrone serie f.m. perrone. gestione dell'ansaldo e vicende familiari
archivio perrone serie m. e p. perrone. gestione dell'ansaldo	archivio perrone materiale a stampa	raccolta disegni tecnici industriali	archivio ansaldo collezione opuscoli	archivio ansaldo collezioni pubblicazioni tecniche	archivio perrone serie m. e p. perrone. gestione dell'ansaldo e vicende familiari
fondo fotografico ansaldo	fondo fotografico perrone				

SEARCH IN SELECTED SERIES

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

The **list visualization of the series** linked to the specific concept



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

Identified Resource

Advanced Search in:

Raccolta Memorie d'epoca e Corsali

Raccolta Disegni Tecnici Industriali

Collezione Opuscoli

Collezione pubblicazioni tecniche

Mano e Pio Perrone Centro dell'Ansaldo e vicine ditte

Fondo Fotografico Ansaldo

Fondo Fotografico Perrone

Keyword

Places

Date  Type

Documentary typology

SEARCH RESET

Objects from the Digital Archive

Jump to page:  GO

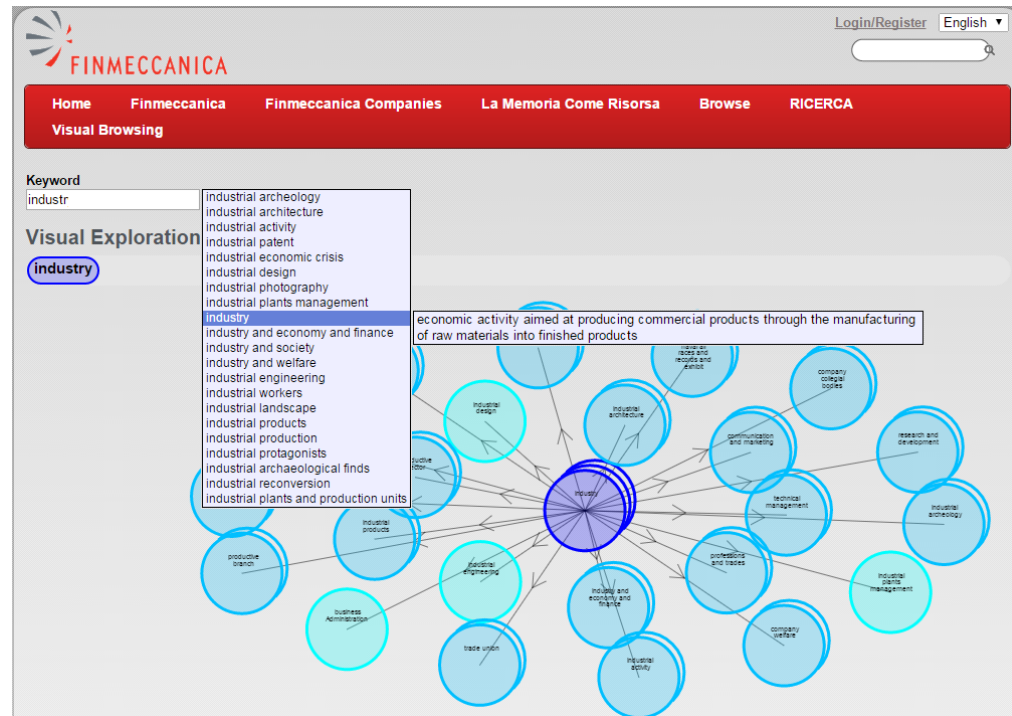
Your search found 31 results.

« Previous page 1/2 Next »

OPTION S FILTER SEARCH

Aereo Ansaldo -1 IT.AHFA.SRAC.DT1.1	Aereo Ansaldo SVA IT.AHFA.SRAC.MEC.1	Aereo Ansaldo SVA IT.AHFA.SRAC.DT1.2	Aereo Ansaldo SVA, Aeroplano Cristoforo Colombo IT.AHFA.SRAC.MEC.2	Aeroplano Ansaldo SVA 1917, Porio Anico Gerona IT.AHFA.SRAC.MEC.3	Aeroplano SVA, con dedica di Antonio Locatelli IT.AHFA.SDOC.FP.S4.2
Aeroplano SVA, Gabriele D'Annunzio, raid su Vite... IT.AHFA.SDOC.FP.S4.1	Aeroplano SVA, raid Roma-Tokio, Ferrarini e Ca... IT.AHFA.SFOT.FA.1	Ansaldo San Giorgio, equipaggiamento elettrico IT.AHFA.SDOC.FA.CPT.2	Carolina postale aeroplano Ansaldo, Gio Ansaldo... IT.AHFA.SDOC.FP.S4.3	Copertina di opuscolo, produzioni aeronautiche... IT.AHFA.SDOC.FA.CPT.1	Famiglia Perrone - Biografia IT.AHFA.SIMG.FP.SDBPH
Famiglia Perrone - Occupazione IT.AHFA.SIMG.FP.SDOCC	Illustrazione italiana, il volo della Saverise IT.AHFA.SDOC.FA.CO.4	Introduzione Archivistica - Perrone IT.AHFA.SIMG.FP.SDDE.SCR	Locomotiva a vapore di George Saffrisson IT.AHFA.SRAC.DT1.4	Los aeroplanos Ansaldo, Porspiano con dedica... IT.AHFA.SDOC.FA.CPT.4	Opuscolo aereo Ansaldo SVA IT.AHFA.SDOC.FA.CO.5

- 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

The screenshot shows the FINMECCANICA website interface. The navigation bar includes Home, Finmeccanica, Finmeccanica Companies, La Memoria Come Risorsa, Browse, and RICERCA. The Visual Browsing section shows a search bar with 'industry' entered and a 'Visual Exploration' section with a button for 'industry'. The main area displays a network graph with a central 'industry' node and various surrounding nodes like 'productive branch', 'industrial design', and 'industrial engineering'.



This screenshot shows the same FINMECCANICA website interface after a user has selected 'productive branch'. The 'Visual Exploration' section now has buttons for both 'industry' and 'productive branch', with 'productive branch' being the active selection. The network graph is filtered to show only nodes and edges related to 'productive branch', such as 'mechanics productive branch', 'industrial design', and 'industrial photography'.

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.

Visual Exploration

industry productive branch mechanics productive branch aeronautic productive sector airplanes

heterogeneous documentation about the activities carried out within the technical research and the industrial production (photographs of product and production process phases), scientific and commercial communication (photographs of specialized magazines and advertising material), celebrations and events of the company life and recreational employees club

company and plant photographic laboratory

technical drawings

industrial artifacts collection

airplanes

research department

publications and company magazines and advertising material

technical and designing documentation

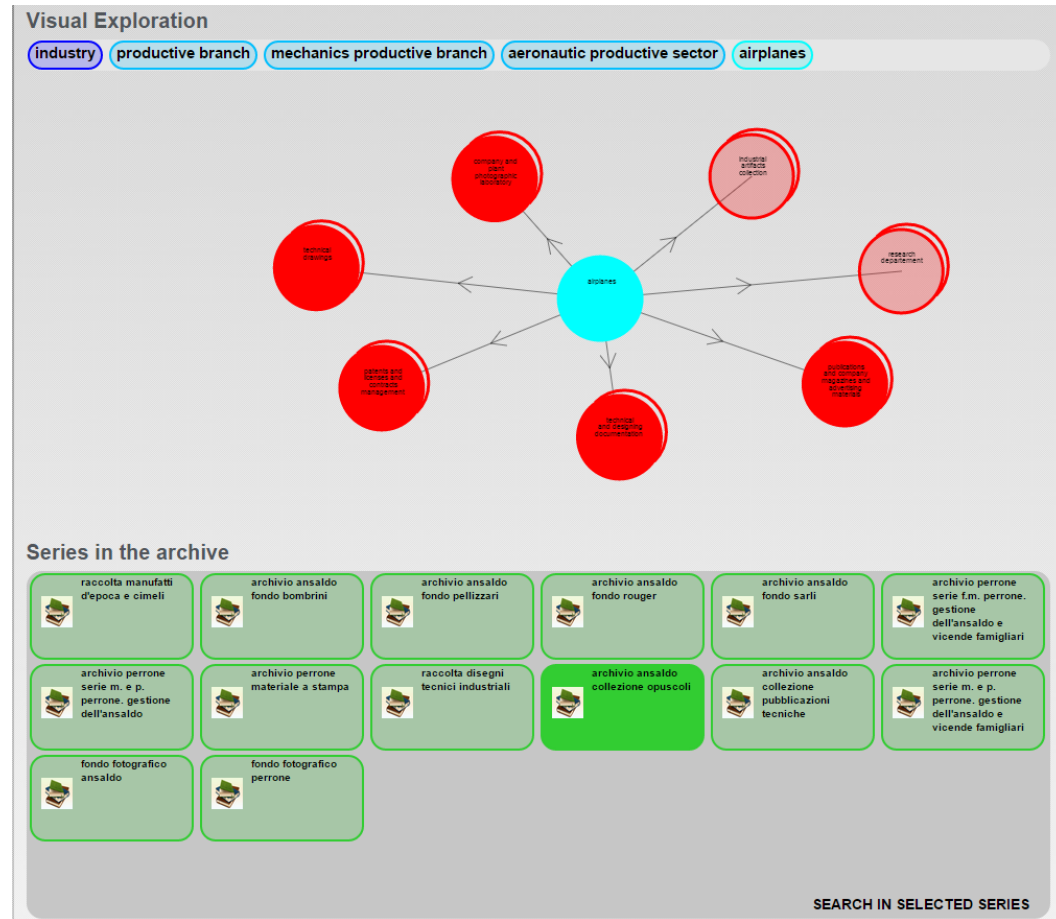
patents and technical consultants management

Series in the archive

raccolta manufatti d'epoca e cimeli	archivio ansaldo fondo bombrini	archivio ansaldo fondo pellizzari	archivio ansaldo fondo rouger	archivio ansaldo fondo sarli	archivio perrone serie f.m. perrone. gestione dell'ansaldo e vicende famigliari
archivio perrone serie m. e p. perrone. gestione dell'ansaldo	archivio perrone materiale a stampa	raccolta disegni tecnici industriali	archivio ansaldo collezione opuscoli	archivio ansaldo collezione pubblicazioni tecniche	archivio perrone serie m. e p. perrone. gestione dell'ansaldo e vicende famigliari
fondo fotografico ansaldo	fondo fotografico perrone				

SEARCH IN SELECTED SERIES

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.

Series in the archive

 raccolta manufatti d'epoca e cimeli	 archivio ansaldo fondo bombrini	 archivio ansaldo fondo pellizzari	 archivio ansaldo fondo rouger	 archivio ansaldo fondo sarli	 archivio perrone serie f.m. perrone. gestione dell'ansaldo e vicende famigliari
 archivio perrone serie m. e p. perrone. gestione dell'ansaldo	 archivio perrone materiale a stampa	 raccolta disegni tecnici industriali	 archivio ansaldo collezione opuscoli	 archivio ansaldo collezione pubblicazioni tecniche	 archivio perrone serie m. e p. perrone. gestione dell'ansaldo e vicende famigliari
 fondo fotografico ansaldo	 fondo fotografico perrone				

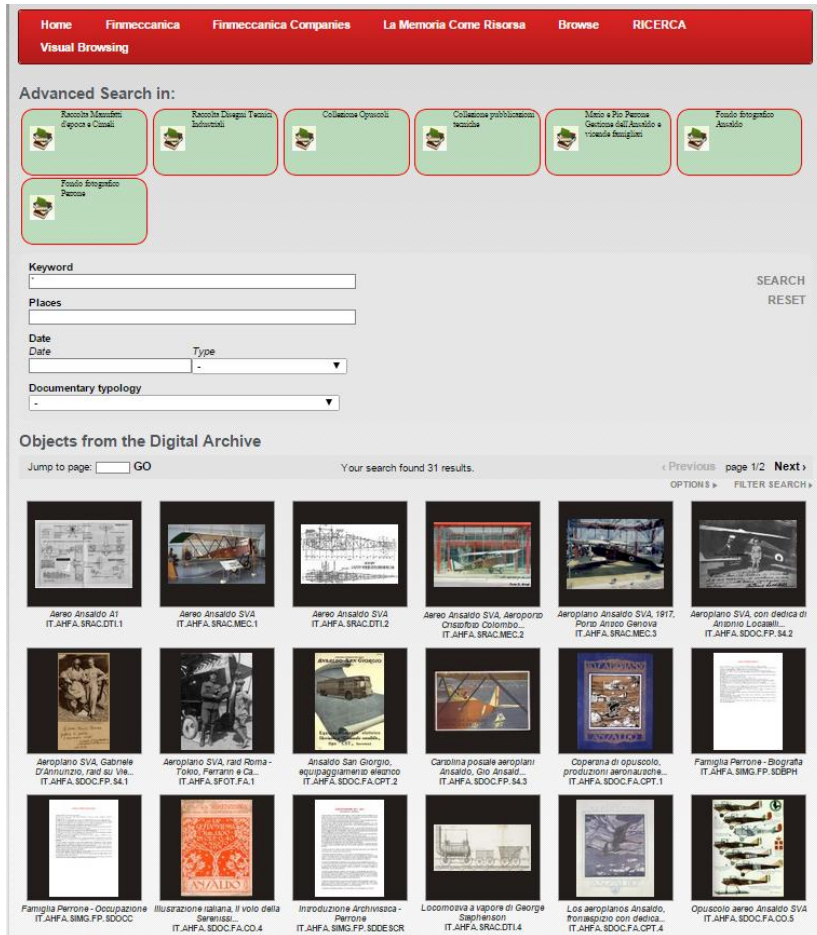
SEARCH IN SELECTED SERIES

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.

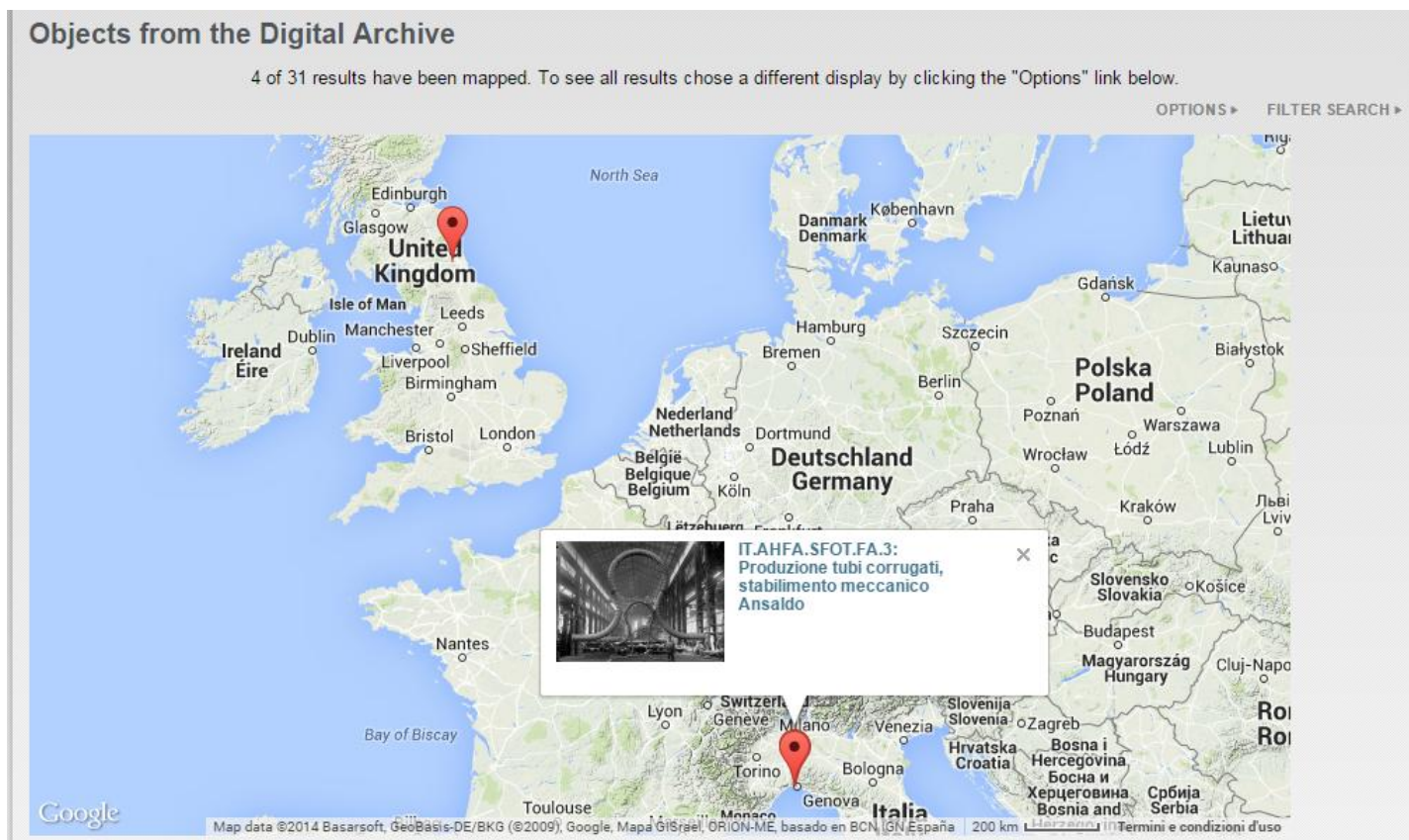


The screenshot shows the R.I.C.E.R.C.A. website search interface. At the top, there is a navigation bar with links: Home, Finmeccanica, Finmeccanica Companies, La Memoria Come Ricorsa, Browse, and RICERCA. Below this is a 'Visual Browsing' section with 'Advanced Search in:' and several filter buttons: Raccolta Memorie d'epoca e Cronici, Raccolta Disegni Tecnici Industriali, Collezione Opuscoli, Collezione pubblicazioni tecniche, Museo e Pio Perrone Sezione dell'Ansaldo e vascella fotografate, and Fondo fotografico Ansaldo. Below the filters is a search form with fields for Keyword, Places, Date (with a Type dropdown), and Documentary typology. There are 'SEARCH' and 'RESET' buttons. Below the search form, it says 'Objects from the Digital Archive' and 'Your search found 31 results.' There are navigation links for 'Previous', 'page 1/2', and 'Next'. Below this is a grid of 18 search results, each with a thumbnail image and a title. The results include: Aereo Ansaldo A1, Aereo Ansaldo SVA, Aereo Ansaldo SVA, Aereo Ansaldo SVA, Aeroplano Ansaldo SVA, Aeroplano SVA, Aeroplano SVA, Aeroplano SVA, Aeroplano SVA, Aeroplano SVA, Aeroplano SVA, Aeroplano SVA, Aeroplano SVA, Aeroplano SVA, Aeroplano SVA, Aeroplano SVA, Aeroplano SVA, and Aeroplano SVA.

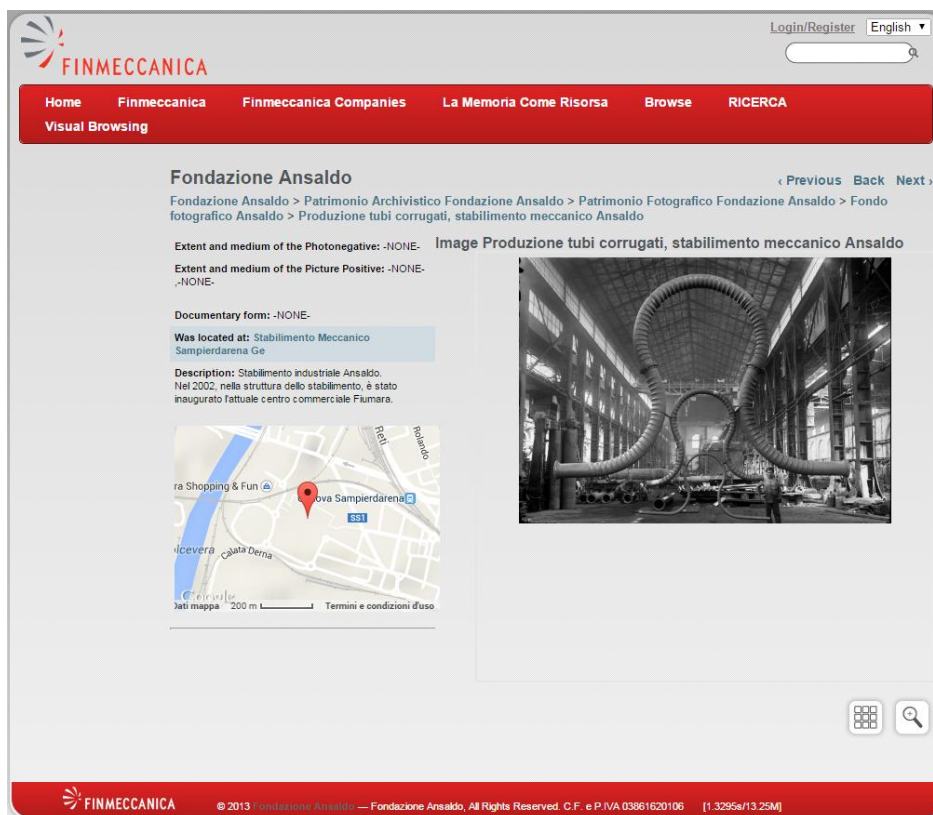
# Illustrative Example

## Geographic based search

Search results can be mapped



## The page dedicated to the resource



The screenshot displays the FINMECCANICA website interface. At the top, there is a navigation bar with links for Home, Finmeccanica, Finmeccanica Companies, La Memoria Come Risorsa, Browse, and RICERCA. Below this is a search bar and a language selector set to English. The main content area features the title 'Fondazione Ansaldo' and a breadcrumb trail: 'Fondazione Ansaldo > Patrimonio Archivistico Fondazione Ansaldo > Patrimonio Fotografico Fondazione Ansaldo > Fondo fotografico Ansaldo > Produzione tubi corrugati, stabilimento meccanico Ansaldo'. The page includes metadata such as 'Extent and medium of the Photonegative: -NONE-', 'Extent and medium of the Picture Positive: -NONE-', and 'Documentary form: -NONE-'. A map shows the location of the 'Stabilimento Meccanico Sampierdarena Ge'. A large photograph depicts the industrial interior of the factory, showing large corrugated metal pipes. The footer contains the FINMECCANICA logo, copyright information for 2013, and contact details for Fondazione Ansaldo.

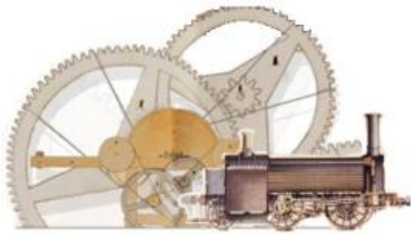
## □ The Semantic Framework

- 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



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

EUROMED 2014

03 - 08 November 2014