A semantic collaborative system for the management of translational research projects

Matteo Gabetta, Giuseppe Milani, Cristiana Larizza, Valentina Favalli, Eloisa Arbustini, Riccardo Bellazzi

Fondazione IRCCS Policlinico San Matteo

INHERITANCE PROJECT

BIO-MEDICAL INFORMATICS "Mario Stefanelli"

UNIVERSITÀ DI PAVIA

Outline

- The INHERITANCE project
- Biomedical Informatics Tools
- Semantic Wiki
  - Technologies
  - Organizational Data Management
  - Scientific Data Management
    - NLP
    - Literature Mining
- Conclusions

The INHERITANCE project

Cardiomyopathies:

"*primary myocardial disorders of unknown cause*"

4 main subtypes**:**

- Hypertrophic (HCM)
- Dilated (DCM)
- Restrictive (RCM)
- Arrhythmogenic Right Ventricular (ARVC)

The INHERITANCE project

Cardiomyopathies:

"*primary myocardial disorders of unknown cause*"

4 main subtypes**:**

- Hypertrophic (HCM)
- **Dilated (DCM)**
- Restrictive (RCM)
- Arrhythmogenic Right Ventricular (ARVC)

The INHERITANCE project

Dilated Cardiomyopathy:

*"[…] myocardial disorder characterized by the presence of left ventricular dilatation and systolic impairment, in the absence of abnormal loading conditions (e.g. hypertension, valve disease) or coronary artery disease sufficient to cause global systolic dysfunction." **

* Elliott P, et al. *Classification of the cardiomyopathies: a position statement from the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases*. Eur Heart J. 2008; 29: 270–276.

The INHERITANCE project

Dilated Cardiomyopathy:

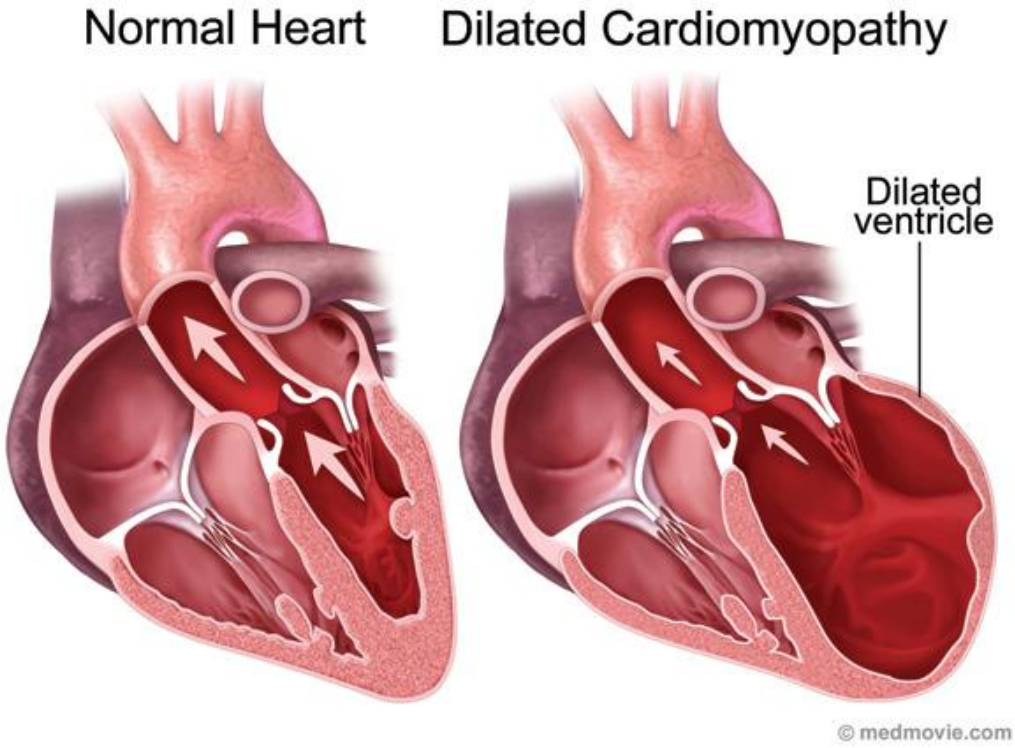
*"[…] myocardial disorder characterized by the presence of left ventricular dilatation and systolic impairment, in the absence of abnormal loading conditions (e.g. hypertension, valve disease) or coronary artery disease sufficient to cause global systolic dysfunction." **

- **20 disease-causing genes** (to date)

* Elliott P, et al. *Classification of the cardiomyopathies: a position statement from the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases*. Eur Heart J. 2008; 29: 270–276.

The INHERITANCE project

Dilated Cardiomyopathy:



* Elliott P, et al. *Classification of the cardiomyopathies: a position statement from the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases*. Eur Heart J. 2008; 29: 270–276.

The INHERITANCE project

INtegrated **HE**art **R**esearch **I**n Tr**AN**slational genetics of **C**ardiomyopathies in **E**urope

- 3-year health research project
- European Commission Funding Program 7
- 11 European centers

The INHERITANCE project

INtegrated **HE**art **R**esearch **I**n Tr**AN**slational genetics of **C**ardiomyopathies in **E**urope

Translational strategy:

Disease-specific features

(red flags)

⟷

Biological features

(genetic or metabolic pathways)

The INHERITANCE project

INtegrated **HE**art **R**esearch **I**n Tr**AN**slational genetics of **C**ardiomyopathies in **E**urope

6 research areas:

- Clinical Cardiogenetics
- -omics
- Animal Studies
- Structural Studies
- Treatments
- Biomedical Informatics

The INHERITANCE project

INtegrated **HE**art **R**esearch **I**n Tr**AN**slational genetics of **C**ardiomyopathies in **E**urope

6 research areas:

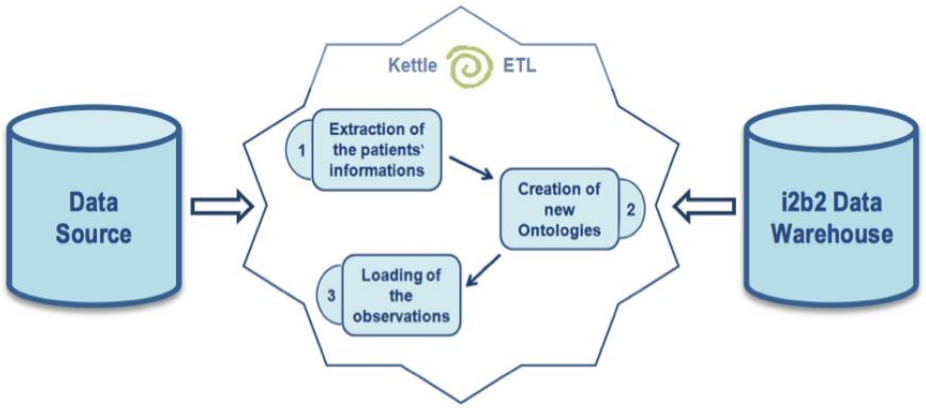
- Clinical Cardiogenetics
- -omics
- Animal Studies
- Structural Studies
- Treatments
- **Biomedical Informatics**

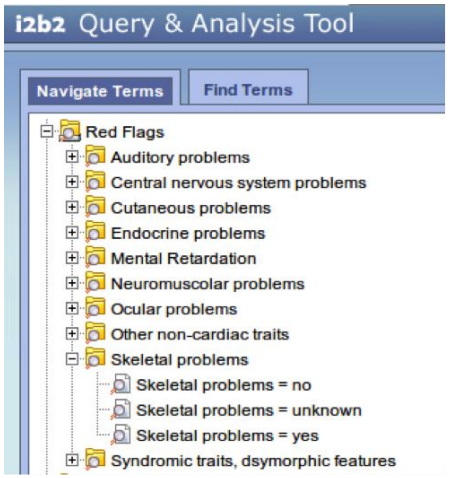
Biomedical Informatics Tools

- Data Warehouse
- Automated Literature Analysis
- Case-Based Reasoning
- Literature-Based Gene Prioritization
- Semantic Wiki

Biomedical Informatics Tools

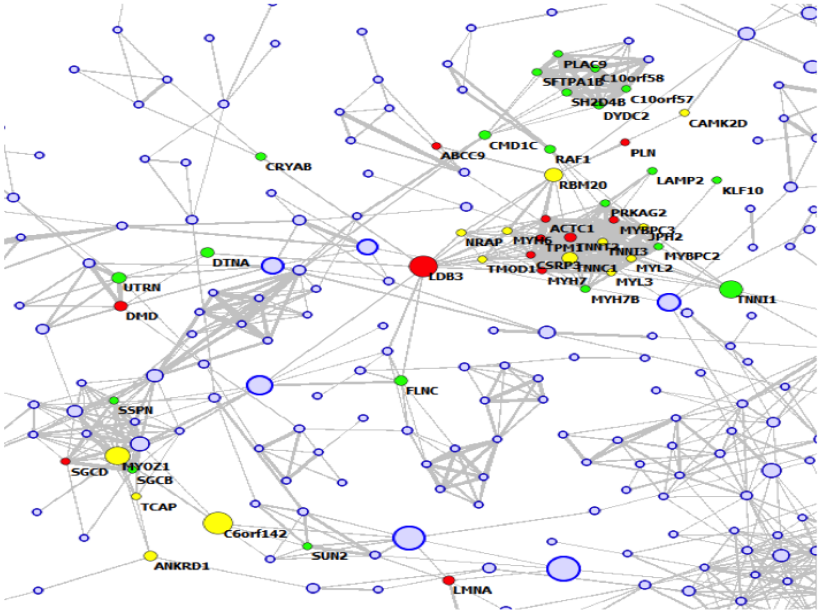
- **Data Warehouse**
- Automated Literature Analysis
- Case-Based Reasoning
- Literature-Based Gene Prioritization
- Semantic Wiki





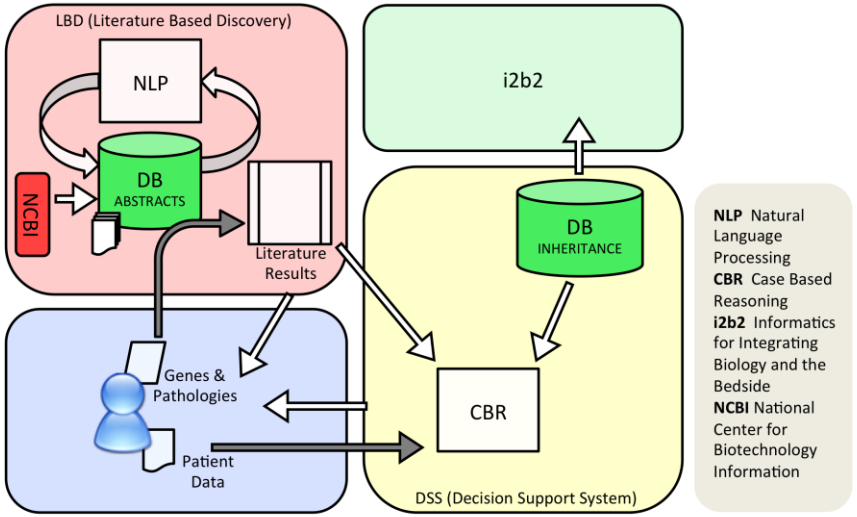
Biomedical Informatics Tools

- Data Warehouse
- **Automated Literature Analysis**
- Case-Based Reasoning
- Literature-Based Gene Prioritization
- Semantic Wiki



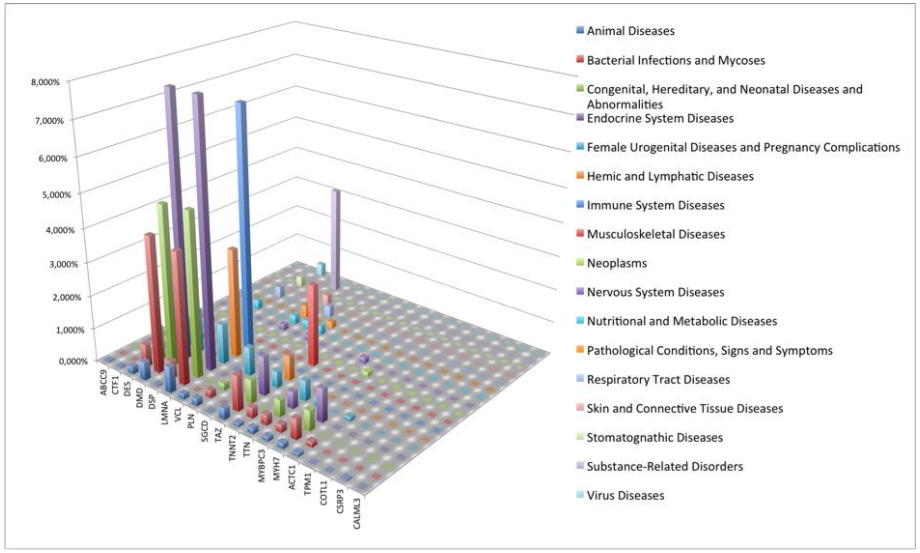
Biomedical Informatics Tools

- Data Warehouse
- Automated Literature Analysis
- **Case-Based Reasoning**
- Literature-Based Gene Prioritization
- Semantic Wiki



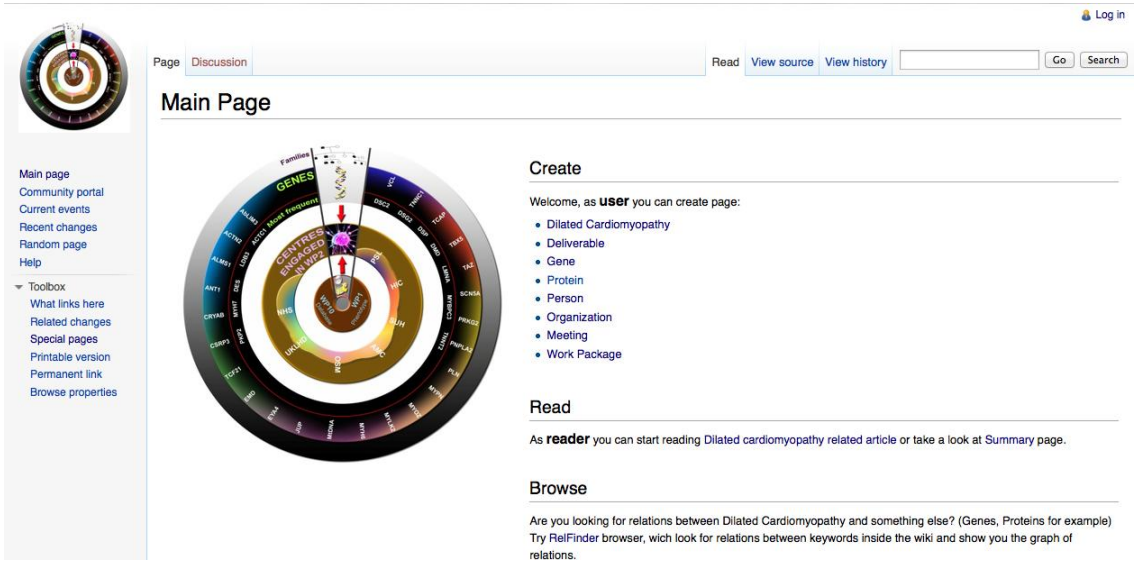
Biomedical Informatics Tools

- Data Warehouse
- Automated Literature Analysis
- Case-Based Reasoning
- **Literature-Based Gene Prioritization**
- Semantic Wiki



Biomedical Informatics Tools

- Data Warehouse
- Automated Literature Analysis
- Case-Based Reasoning
- Literature-Based Gene Prioritization
- **Semantic Wiki**



Semantic Wiki

- Track project activities
- Share ideas
- Share data
- Exchange information between investigators
- Manage scientific research products

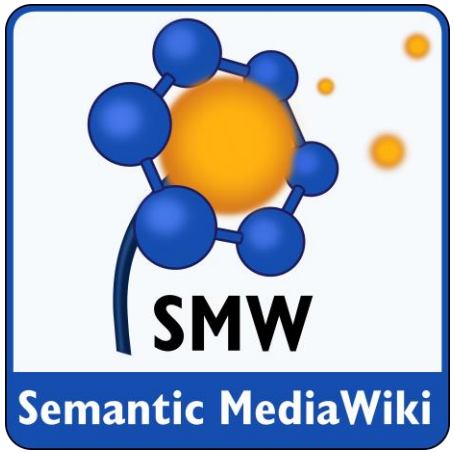
ORGANIZATIONAL ASPECTS

+

SCIENTIFIC KNOWLEDGE

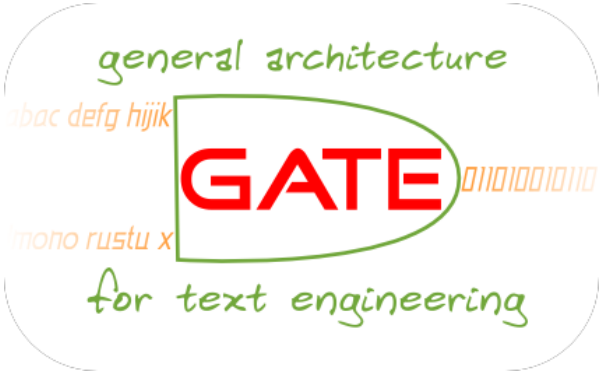
Semantic Wiki

- Free web-based wiki software
- Wikimedia Foundation / Wikipedia
- Extensibility



- MediaWiki extension
- Semantic data
- Semantic search
- Data export (e.g. RDF)

Semantic Wiki



- Open-source framework for NLP
- Libraries of Text Mining tools
- API's for tools development

RelFinder

- Querying tool
- Graphical relation browser

NCBI

Entrez Utilities Web Service

- Pubmed access
- Web service + APIs
- SOAP protocol

Semantic MediaWiki

Building blocks:

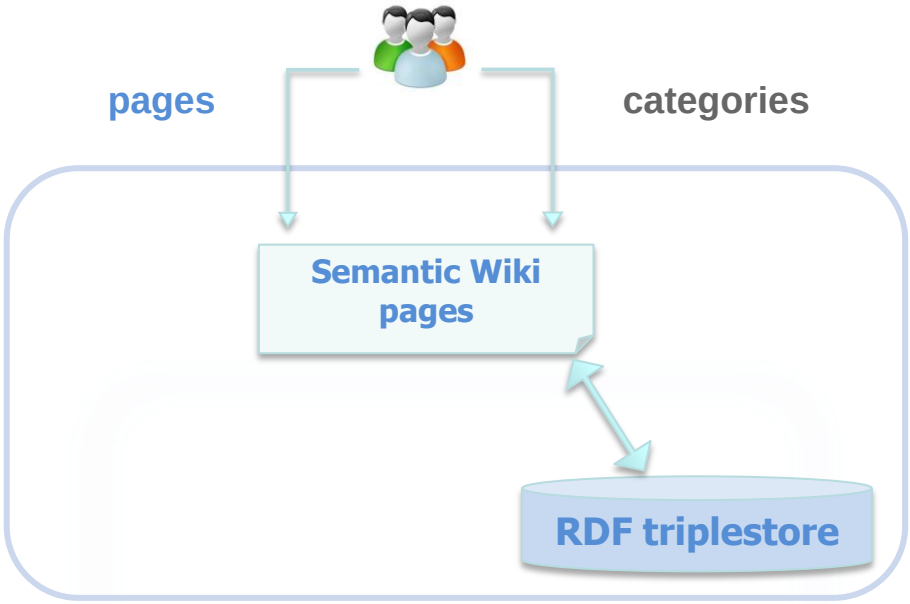
- **Categories** → data model in the Wiki
- **Templates**
- **Forms**

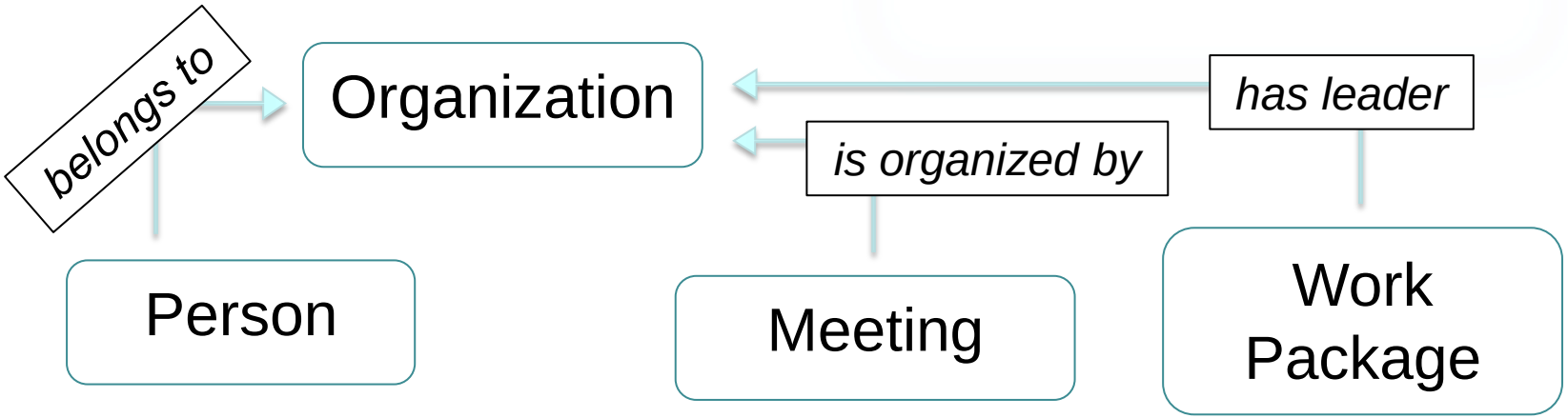
→ define content of Categories

Organizational Aspects

Categories:

- Person
- Organization
- Meeting
- Work Package





Organizational Aspects

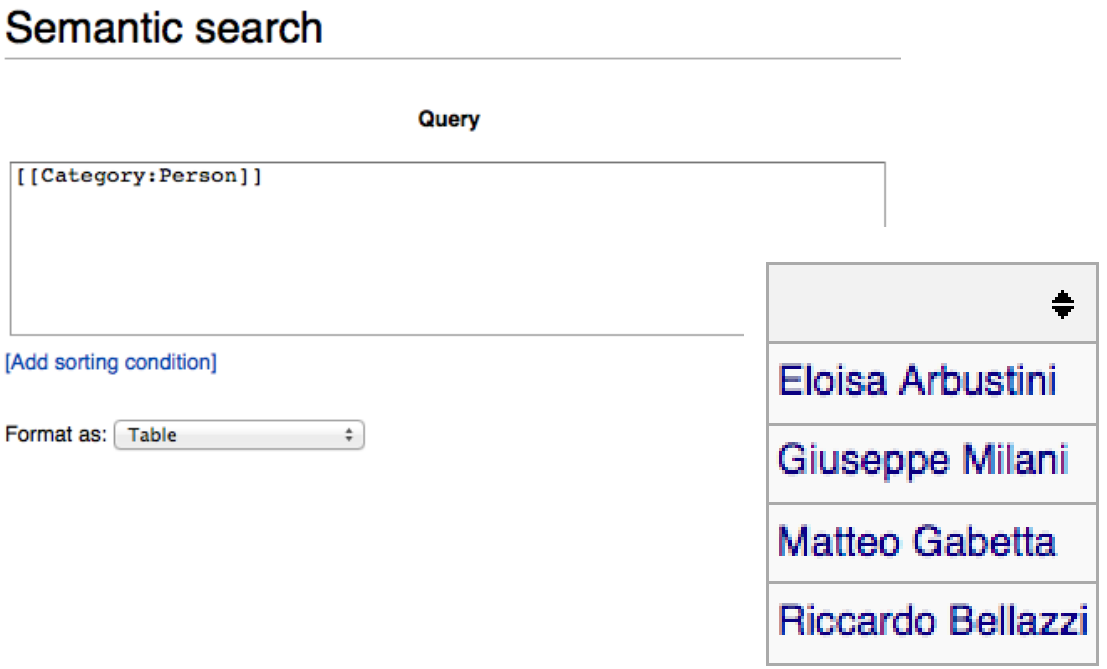
Queries:

- Built-in tool (inline queries)
- RDF export ➔ SPARQL
- RelFinder

Organizational Aspects

Queries:

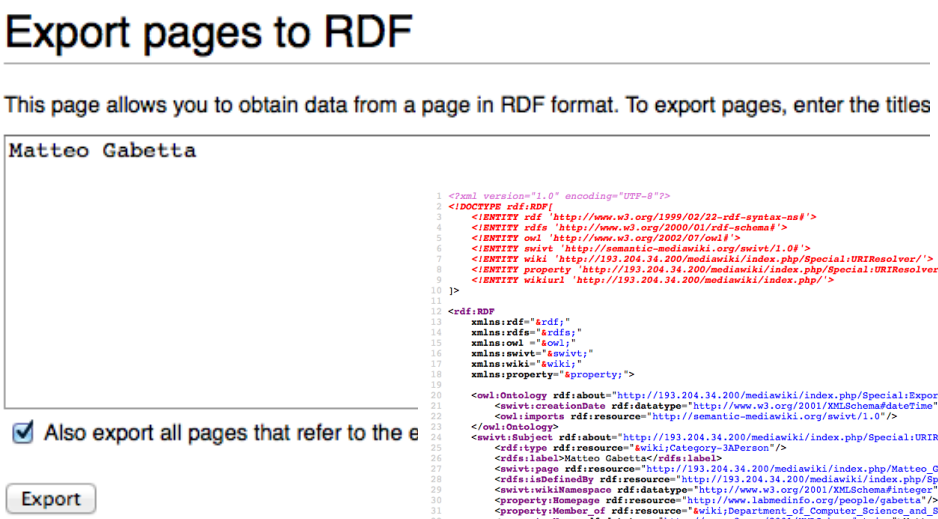
- **Built-in tool (inline queries)**
- RDF export ➔ SPARQL
- RelFinder



Organizational Aspects

Queries:

- Built-in tool (inline queries)
- **RDF export ➔ SPARQL**
- RelFinder

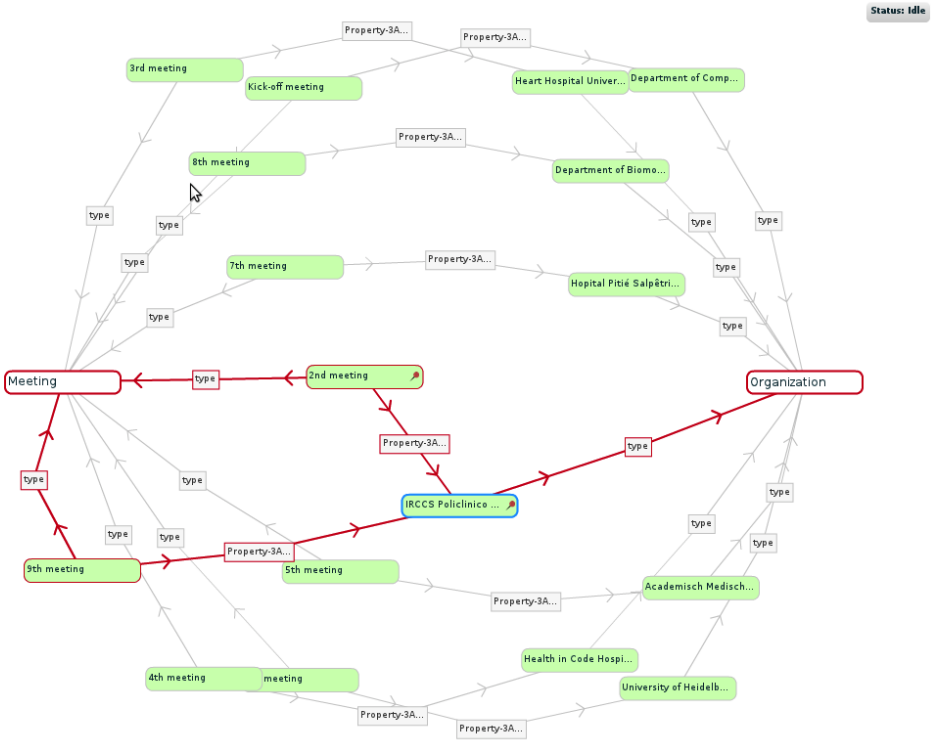



```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE rdf:RDF[
    <!ENTITY rdf 'http://www.w3.org/1999/02/22-rdf-syntax-ns#'>
    <!ENTITY rdfs 'http://www.w3.org/2000/01/rdf-schema#'>
    <!ENTITY owl 'http://www.w3.org/2002/07/owl#'>
    <!ENTITY swivt 'http://semantic-mediawiki.org/swivt/1.0#'>
    <!ENTITY wiki 'http://193.204.34.200/mediawiki/index.php/Special:URIResolver/'>
    <!ENTITY property 'http://193.204.34.200/mediawiki/index.php/Special:URIResolver/Property-3A'>
    <!ENTITY wikiurl 'http://193.204.34.200/mediawiki/index.php/'>
]>

<rdf:RDF
    xmlns:rdf="&rdf;"
    xmlns:rdfs="&rdfs;"
    xmlns:owl="&owl;"
    xmlns:swivt="&swivt;"
    xmlns:wiki="&wiki;"
    xmlns:property="&property;">

    <owl:Ontology rdf:about="http://193.204.34.200/mediawiki/index.php/Special:ExportRDF/Matteo+Gabetta">
        <swivt:creationDate rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2012-11-13T17:29:43+01:00</swivt:creationDate>
        <owl:imports rdf:resource="http://semantic-mediawiki.org/swivt/1.0"/>
    </owl:Ontology>
    <swivt:Subject rdf:about="http://193.204.34.200/mediawiki/index.php/Special:URIResolver/Matteo_Gabetta">
        <rdf:type rdf:resource="&wiki;Category-3APerson"/>
        <rdfs:label>Matteo Gabetta</rdfs:label>
        <swivt:page rdf:resource="http://193.204.34.200/mediawiki/index.php/Matteo_Gabetta"/>
        <rdfs:isDefinedBy rdf:resource="http://193.204.34.200/mediawiki/index.php/Special:ExportRDF/Matteo_Gabetta"/>
        <swivt:wikiNamespace rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">0</swivt:wikiNamespace>
        <property:Homepage rdf:resource="http://www.labmedinfo.org/people/gabetta"/>
        <property:Member_of rdf:resource="&wiki;Department_of_Computer_Science_and_System._University_of_Pavia"/>
        <property:Name rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Matteo Gabetta</property:Name>
        <swivt:wikiPageModificationDate rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2012-01-24T18:55:46Z</swivt:wikiPageModificationDate>
        <swivt:wikiPageSortKey rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Matteo Gabetta</swivt:wikiPageSortKey>
    </swivt:Subject>
    <owl:DatatypeProperty rdf:about="http://semantic-mediawiki.org/swivt/1.0#creationDate" />
    <owl:Class rdf:about="http://193.204.34.200/mediawiki/index.php/Special:URIResolver/Category-3APerson" />
    <owl:ObjectProperty rdf:about="http://semantic-mediawiki.org/swivt/1.0#page" />
    <owl:DatatypeProperty rdf:about="http://semantic-mediawiki.org/swivt/1.0#wikiNamespace" />
    <owl:ObjectProperty rdf:about="http://193.204.34.200/mediawiki/index.php/Special:URIResolver/Property-3AHomepage" />
    <owl:ObjectProperty rdf:about="http://193.204.34.200/mediawiki/index.php/Special:URIResolver/Property-3AMember_of" />
    <owl:DatatypeProperty rdf:about="http://193.204.34.200/mediawiki/index.php/Special:URIResolver/Property-3AName" />
    <owl:DatatypeProperty rdf:about="http://semantic-mediawiki.org/swivt/1.0#wikiPageModificationDate" />
    <owl:DatatypeProperty rdf:about="http://semantic-mediawiki.org/swivt/1.0#wikiPageSortKey" />
    <!-- Created by Semantic MediaWiki, http://semantic-mediawiki.org/ -->
</rdf:RDF>
```

Organizational Aspects

Queries:

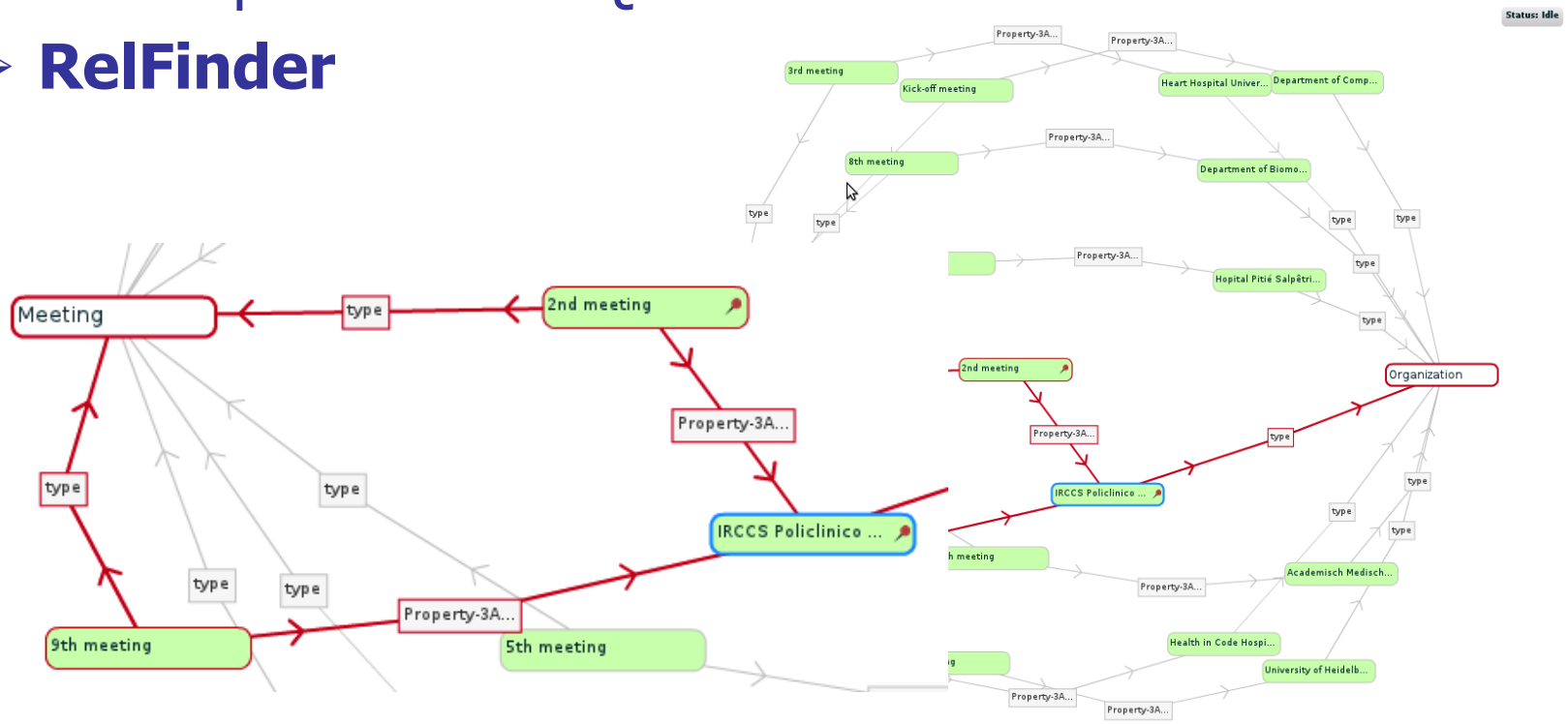
- Built-in tool (inline queries)
- RDF export ➔ SPARQL
- **RelFinder**



Organizational Aspects

Queries:

- Built-in tool (inline queries)
- RDF export ➔ SPARQL
- **RelFinder**



Organizational Aspects

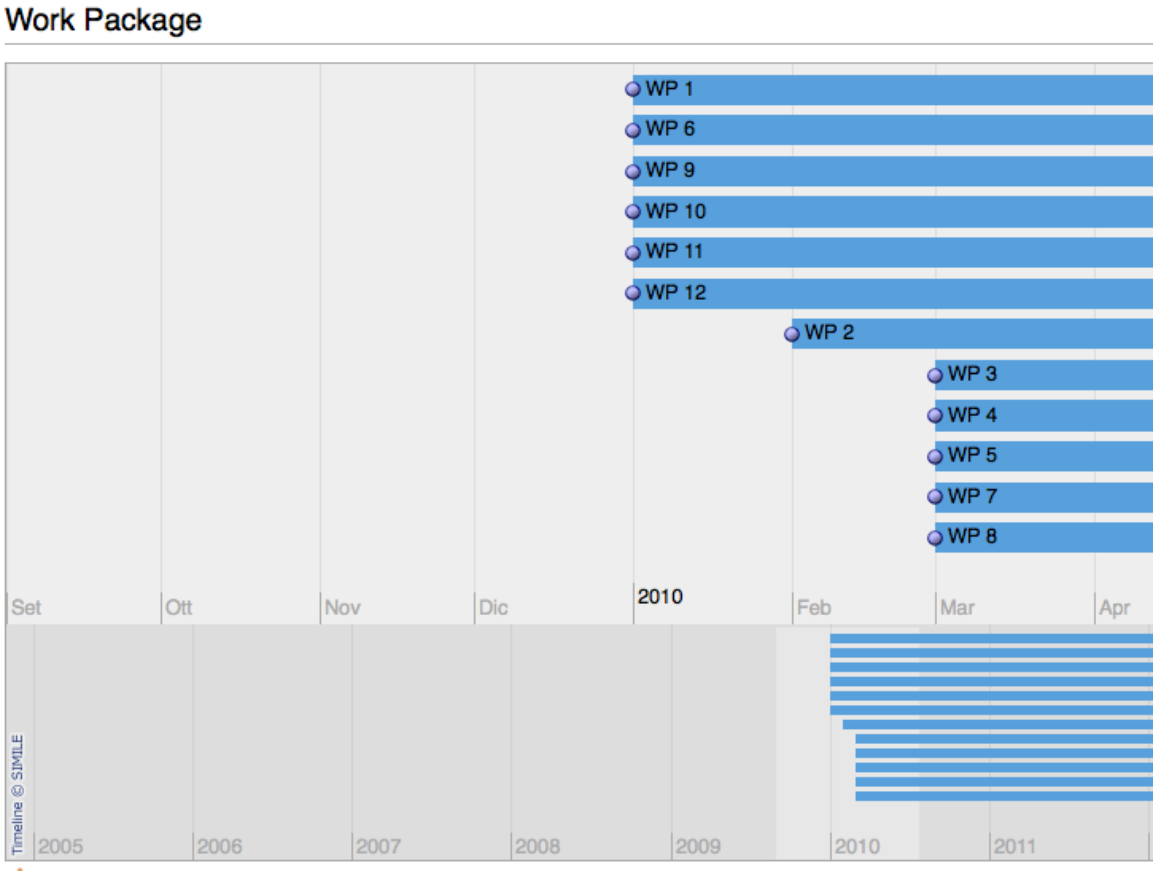
➢ Summary Page

Organization

Organization	Partner number
IRCCS Policlinico San Matteo	1
Heart Hospital University College London	2
University Hospital of Umeà	3
University of Heidelberg	4
Hopital Pitié Salpêtrière	5
Health in Code Hospital Marítimo de Oza As Xubias	6
Sanger Building Biochemistry Department	7
Skejby University Hospital	8
Academisch Medisch Centrum	9
Department of Computer Science and System. University of Pavia	10
Department of Biomolecular and Biotechnology Science Milan	11

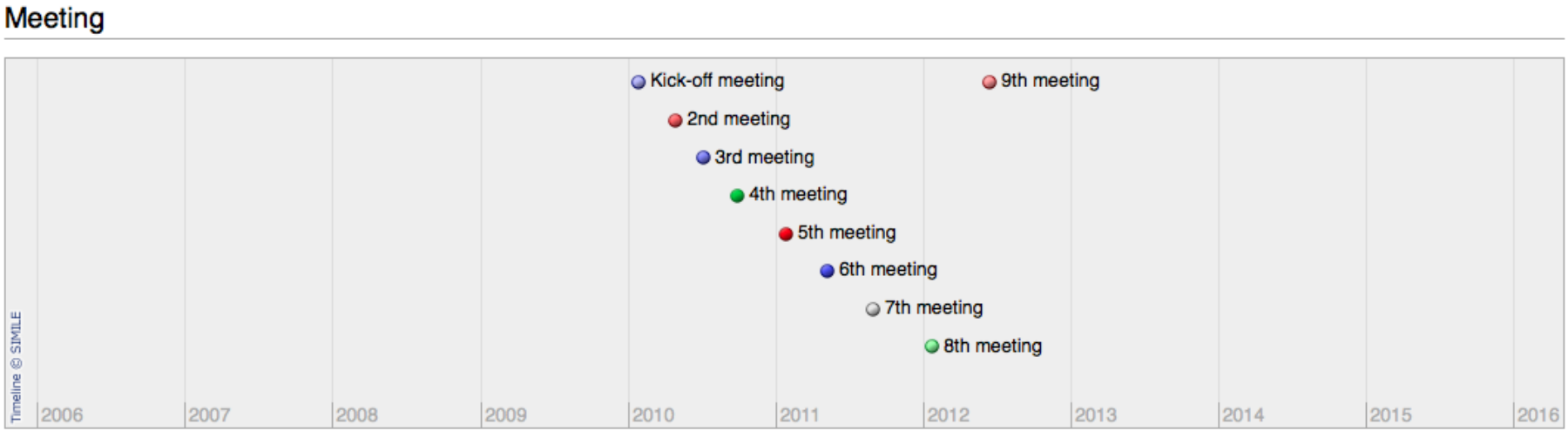
Organizational Aspects

➢ Summary Page



Organizational Aspects

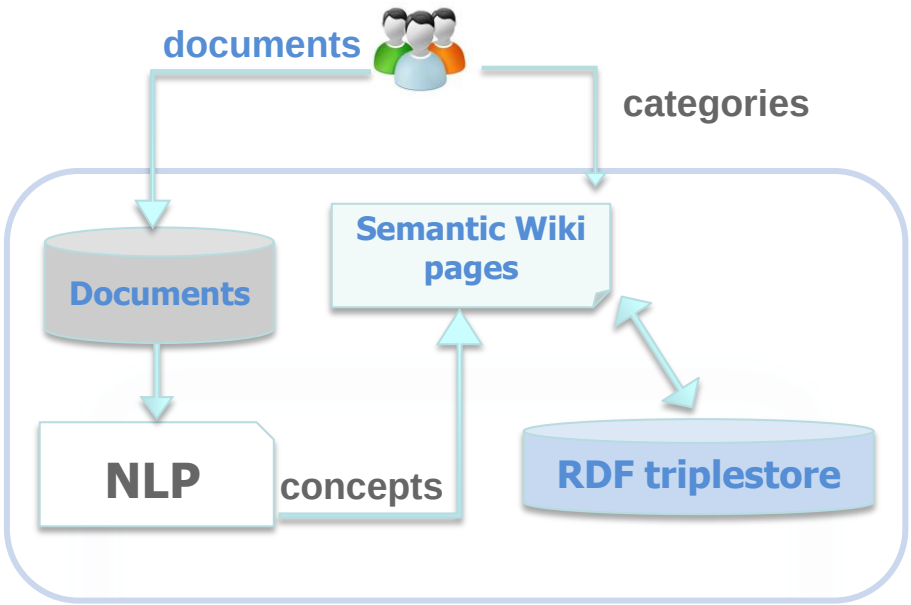
➢ Summary Page

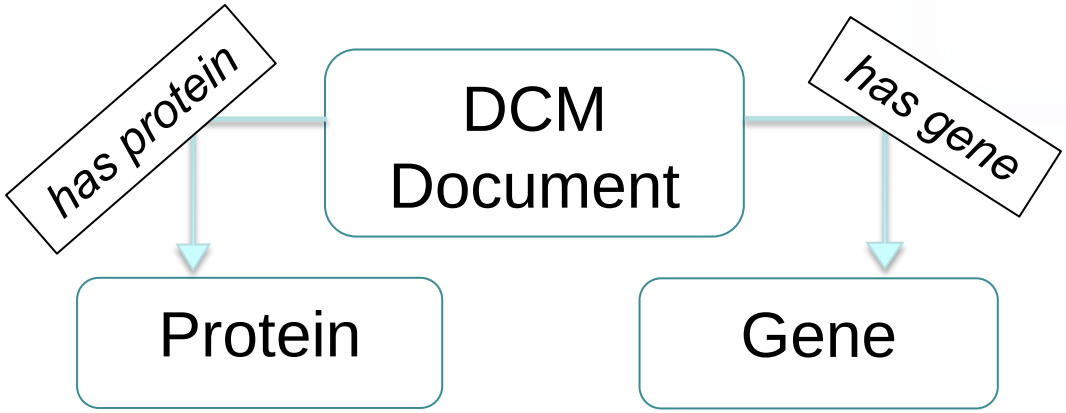


Scientific Knowledge

Categories:

- Gene
- Protein
- Dilated Cardiomyopathy Document

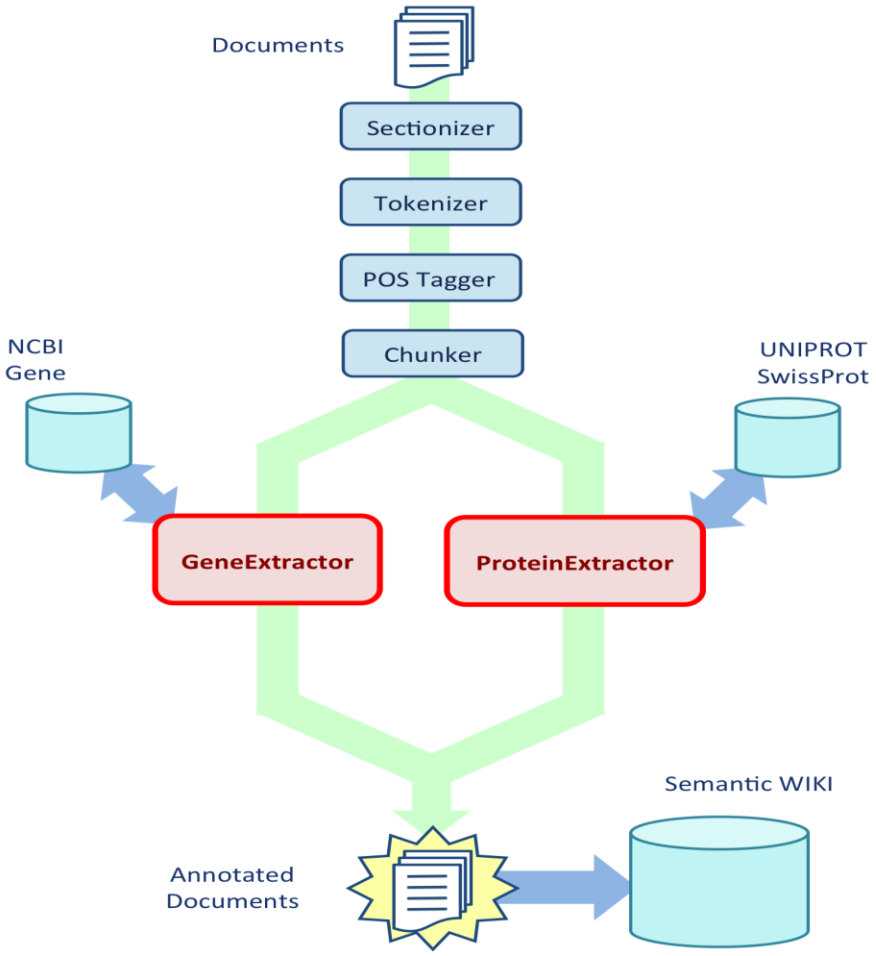




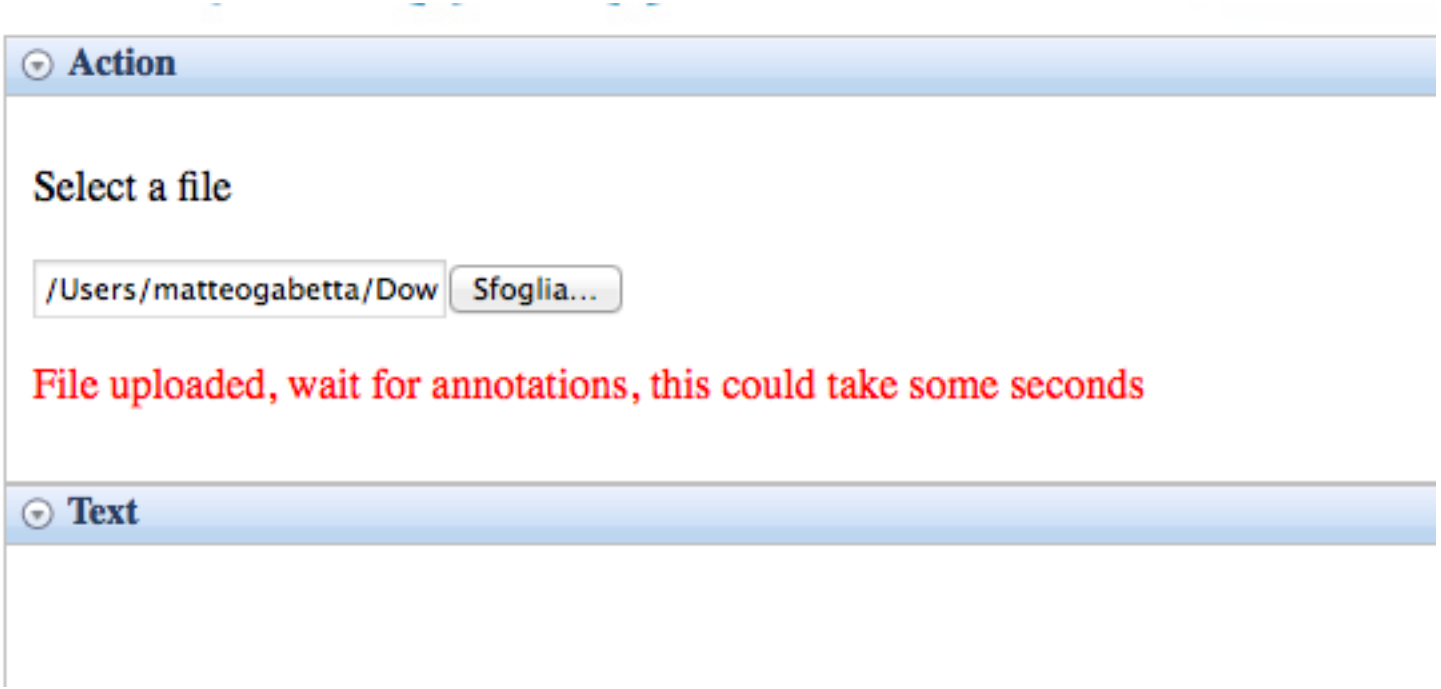
Natural Language Processing

- GATE
- accessed via servlet
- .txt, .rtf, MS Word
- API plugins + purposely developed plugins
- GeneExtractor (NCBI Gene)
- ProteinExtractor (Uniprot / Swiss-Prot)

Natural Language Processing



Natural Language Processing



Natural Language Processing

Action

Select Wiki Title Page

CREATE

Text

Post-natal myogenic and adipogenic developmental Defects and metabolic impairment upon loss of A-type lamins
A-type lamins are a major component of the nuclear lamina. Mutations in the [[LMNA]] gene, which encodes the A-type lamins A and C, cause a set of phenotypically diverse diseases collectively called laminopathies. While adult [[LMNA]] null mice show various symptoms typically associated with laminopathies, the effect of loss of lamin A/C on early postnatal development is poorly understood. Here we developed a novel [[LMNA]] null mouse (LMNAGT-/-) based on genetrap technology and analyzed its early post-natal development. We detect [[LMNA]] transcripts in heart, the outflow tract, dorsal aorta, liver and somites during early embryonic development. Loss of A-type lamins results in severe growth retardation and developmental defects of the heart, including impaired myocyte hypertrophy, skeletal muscle hypotrophy, decreased amounts of subcutaneous adipose tissue and impaired ex vivo adipogenic differentiation. These defects cause death at 2 to 3 weeks post partum associated with muscle weakness and metabolic complications, but without the occurrence of dilated cardiomyopathy or an obvious progeroid phenotype. Our results indicate that defective early postnatal development critically contributes to the disease phenotypes in adult laminopathies.

Natural Language Processing

Action

Select Wiki Title Page

Art2-abs CREATE

Text

Post-natal myogenic and adipogenic developmental Defects and metabolic impairment upon loss of A-type lamins

A-type lamins are a major component of the nuclear lamina. Mutations in the [[LMNA]] gene, which encodes the A-type lamins A and C, cause a set of phenotypically diverse diseases collectively called laminopathies. While adult [[LMNA]] null mice show various symptoms typically associated with laminopathies, the effect of loss of lamin A/C on early postnatal development is poorly understood. Here we developed a novel [[LMNA]] null mouse (LMNAGT-/-) based on genetrap technology and analyzed its early post-natal development. We detect [[LMNA]] transcripts in heart, the outflow tract, dorsal aorta, liver and somites during early embryonic development. Loss of A-type lamins results in severe growth retardation and developmental defects of the heart, including impaired myocyte hypertrophy, skeletal muscle hypotrophy, decreased amounts of subcutaneous adipose tissue and impaired ex vivo adipogenic differentiation. These defects cause death at 2 to 3 weeks post partum associated with muscle weakness and metabolic complications, but without the occurrence of dilated cardiomyopathy or an obvious progeroid phenotype. Our results indicate that defective early postnatal development critically contributes to the disease phenotypes in adult laminopathies.

Natural Language Processing

Art2-abs

Has Author(s)	
Has Gene(s)	LMNA, LMNA, LMNA, LMNA
Has Protein(s)	

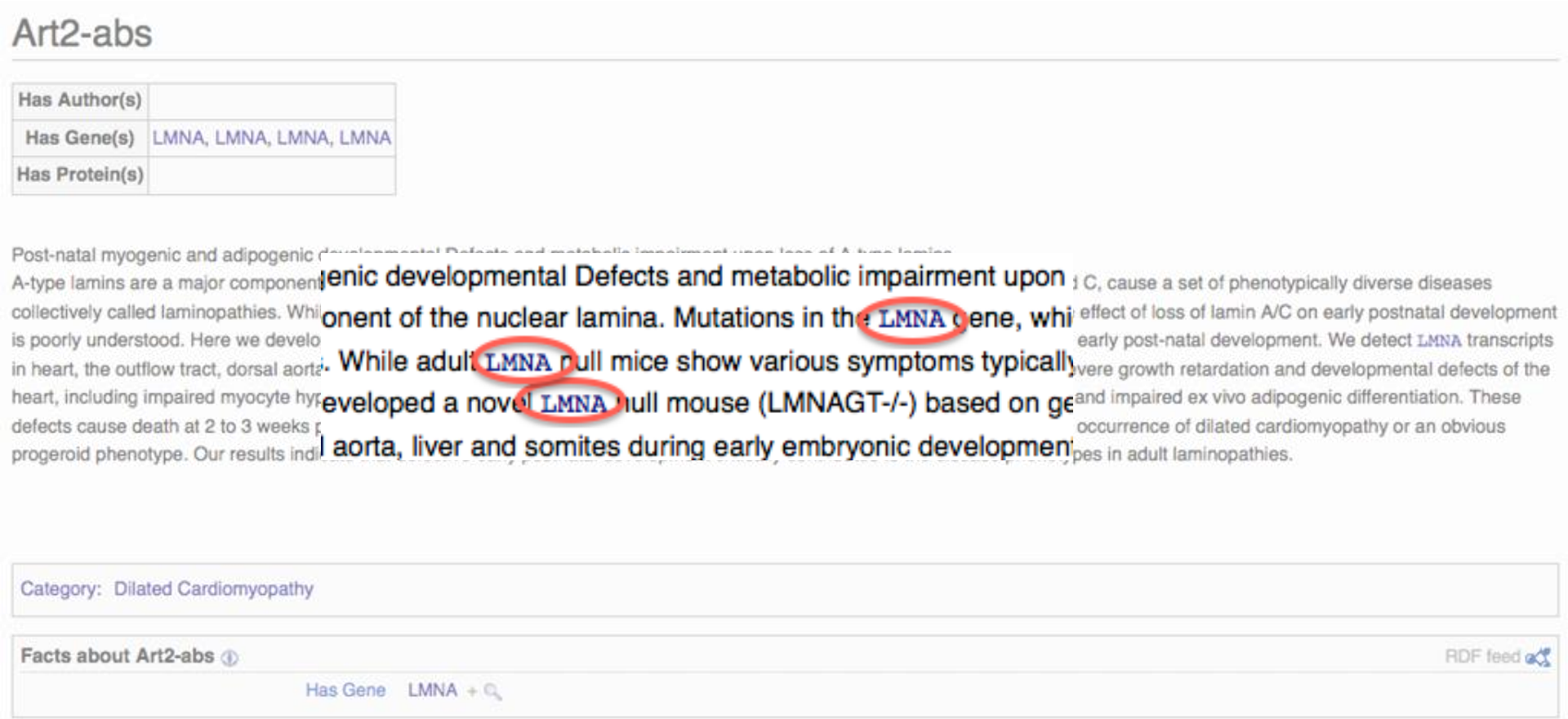
Post-natal myogenic and adipogenic developmental Defects and metabolic impairment upon loss of A-type lamins
A-type lamins are a major component of the nuclear lamina. Mutations in the LMNA gene, which encodes the A-type lamins A and C, cause a set of phenotypically diverse diseases collectively called laminopathies. While adult LMNA null mice show various symptoms typically associated with laminopathies, the effect of loss of lamin A/C on early postnatal development is poorly understood. Here we developed a novel LMNA null mouse (LMNAGT-/-) based on genetrap technology and analyzed its early post-natal development. We detect LMNA transcripts in heart, the outflow tract, dorsal aorta, liver and somites during early embryonic development. Loss of A-type lamins results in severe growth retardation and developmental defects of the heart, including impaired myocyte hypertrophy, skeletal muscle hypotrophy, decreased amounts of subcutaneous adipose tissue and impaired ex vivo adipogenic differentiation. These defects cause death at 2 to 3 weeks post partum associated with muscle weakness and metabolic complications, but without the occurrence of dilated cardiomyopathy or an obvious progeroid phenotype. Our results indicate that defective early postnatal development critically contributes to the disease phenotypes in adult laminopathies.

Category: Dilated Cardiomyopathy

Facts about Art2-abs RDF feed

Has Gene LMNA

Natural Language Processing



Relevant Literature

- NCBI E-utilities
- for Genes and Proteins pages
- 5 most recent articles in Pubmed
- Gene/Protein + "Dilated Cardiomyopathy" (or synonyms)
- retrieved "on the fly"
- link to Pubmed

Relevant Literature

LMNA

Name	LMNA
Symbol	LMNA
Ncbi	http://www.ncbi.nlm.nih.gov/gene/4000

Nicola Carboni, Claudia Sardu, Eleonora Cocco, Giovanni Marrosu, Rosa C Manzi, Vincenzo Nissardi, Franco Isola, Anna Mateddu, Elisabetta Solla, Maria A Maioli, Valentina Oppo, Rachele Piras, Giancarlo Coghe, Carlo Lai, Maria G Marrosu
Cardiac involvement in patients with lamin A/C gene mutations: a cohort observation.
Muscle Nerve: 2012, 46(2);187-92
[PubMed:22806367] [WorldCat.org] [DOI] (I p)

Fresnida J Ramos, Steven C Chen, Michael G Garelick, Dao-Fu Dai, Chen-Yu Liao, Katherine H Schreiber, Vivian L MacKay, Elroy H An, Randy Strong, Warren C Ladiges, Peter S Rabinovitch, Matt Kaeberlein, Brian K Kennedy
Rapamycin reverses elevated mTORC1 signaling in lamin A/C-deficient mice, rescues cardiac and skeletal muscle function, and extends survival.
Sci Transl Med: 2012, 4(144);144ra103
[PubMed:22837538] [WorldCat.org] [DOI] (I p)

Jason C Choi, Antoine Muchir, Wei Wu, Shinichi Iwata, Shunichi Homma, John P Morrow, Howard J Worman
Temsirolimus activates autophagy and ameliorates cardiomyopathy caused by lamin A/C gene mutation.
Sci Transl Med: 2012, 4(144);144ra102
[PubMed:22837537] [WorldCat.org] [DOI] (I p)

Antoine Muchir, Wei Wu, Jason C Choi, Shinichi Iwata, John Morrow, Shunichi Homma, Howard J Worman
Abnormal p38α mitogen-activated protein kinase signaling in dilated cardiomyopathy caused by lamin A/C gene mutation.
Hum. Mol. Genet.: 2012, 21(19);4325-33
[PubMed:22773734] [WorldCat.org] [DOI] (I p)

Relevant Literature

PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed | Search | Advanced | Help

Display Settings: Abstract | Send to:

Full Text Online WILEY ONLINE LIBRARY

Muscle Nerve. 2012 Aug;46(2):187-92. doi: 10.1002/mus.23294.

Cardiac involvement in patients with lamin A/C gene mutations: a cohort observation.

Carboni N, Sardu C, Cocco E, Marrosu G, Manzi RC, Nissardi V, Isola F, Mateddu A, Solla E, Maioli MA, Oppo V, Piras R, Coghe G, Lai C, Marrosu MG.

Neuromuscular Unit, Multiple Sclerosis Center, Department of Cardiovascular and Neurological Sciences, University of Cagliari, Cagliari, Italy. nikola.carboni@tiscali.it

Abstract

INTRODUCTION: LMNA gene mutations are associated with cardiac and skeletal muscle alterations.

METHODS: A cohort of 21 mutated individuals was assessed with clinical and instrumental investigations over the years.

RESULTS: The median observation period was 6 years. Cardiac compromise was detected in 16 patients. Bradyarrhythmias were the most frequent manifestations, followed by supraventricular arrhythmias. Two individuals suffered from nonsustained and 1 from sustained ventricular tachyarrhythmias. Dilated cardiomyopathy was detected in 3 patients. Evaluation of the frequencies of the clinical expressions showed a high probability of suffering from analogue heart compromise in study subjects bearing the same LMNA gene mutation.

CONCLUSIONS: Cardiac involvement represents a very common phenotypic expression of LMNA gene mutation. Subjects sharing common genetic background seem to suffer from analogue pattern of cardiac manifestation.

PMID: 22806367 [PubMed - indexed for MEDLINE]

⊕ MeSH Terms, Substances

⊕ LinkOut - more resources

Save items

Add to Favorites

Related citations in PubMed

A novel lamin A/C mutation in a family with dilated cardiomyopathy, promi [Am Heart J. 2002]

High yield of LMNA mutations in patients with dilated cardiomyopathy and/or [Am Heart J. 2007]

Lamin A/C haploinsufficiency causes dilated cardiomyopathy and ap [J Mol Cell Cardiol. 2008]

Review Lamin A/C and cardiac diseases. [Curr Opin Cardiol. 2006]

Review Heart involvement in lamin A/C related diseases. [Arch Mal Coeur Vaiss. 2006]

See reviews...

See all...

Related information

Related Citations

In conclusion…

- Collaborative Wiki System + Semantic features
- Organizational + Scientific data management
- NLP
- Literature retrieval
- Different query strategies

Future Developments

- Improve scientific knowledge management
  - New Text Mining pipelines ➔ New concepts
  - Link to new databases
- Evaluate usage of INHERITANCE partners
- Integration with other systems

A semantic collaborative system for the management of translational research projects

THANKS FOR YOUR ATTENTION !

Fondazione IRCCS Policlinico San Matteo

INHERITANCE PROJECT

BIO-MEDICAL INFORMATICS "Mario Stefanelli"

UNIVERSITÀ DI PAVIA