



ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO DI FISICA E ASTRONOMIA - DIFA

AIM Live Meeting

CHIMeRA

**Complex Human Interactions in Medical Records
and Atlases**

Nico Curti^{1,2}, Carlo Mengucci^{2,3}, Enrico Giampieri¹,
Gastone Castellani¹, **Daniel Remondini²**

¹*Department of Experimental, Diagnostic and Specialty Medicine, UNIBO*

²*Department of Physics and Astronomy, UNIBO*

³*Department of Agricultural and Food Sciences, UNIBO*

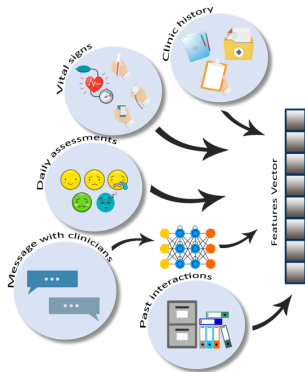
February 2, 2020



Aim of Chimera

Build a network-of-networks with different types of biomedical/chemical/genomical information to enhance interpretability of experimental analyses for:

- ▶ advanced diagnostics
- ▶ novel drug targeting
- ▶ patient stratification
- ▶ multi-omics annotation



This section displays the project's partners and system architecture. At the top are logos for INFN, Filippetti (Engineering, IT & Consulting), Azienda Ospedaliera Sant'Andrea, exprivia, and bimind (We shape technology). Below is a system architecture diagram with a legend: INFN (yellow), Exprivia (green), bimind (light blue), and Filippetti (dark blue). The architecture includes:

- Sentiment Analysis**: A central component connected to 'APP Paziente' and 'APP Operatore'.
- APP Paziente**: Includes 'Componente Device Manager' and 'Sistema AI Analisi e intercomando'.
- APP Operatore**: Includes 'Componente integrazione SO & Moduli applicativi' and 'SIS - Sistema informativo Operatore'.

 At the bottom are logos for the European Union (Unione europea), the University of Bologna, the Regione Lazio, and the Fesr 2014-2020 POR project.

Disease Ontology for FiloBlu

Disease-Symptom Network (from Italian language websites)



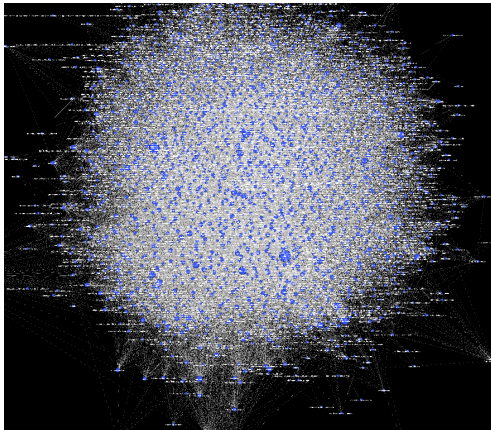
- 1381 nodes;
- 14002 links;



- 1200 nodes;
- 16035 links;

Symptoms Network:

- nodes: 2285;
- links: 29557;





Web sites with
HTML pages &
Ajax

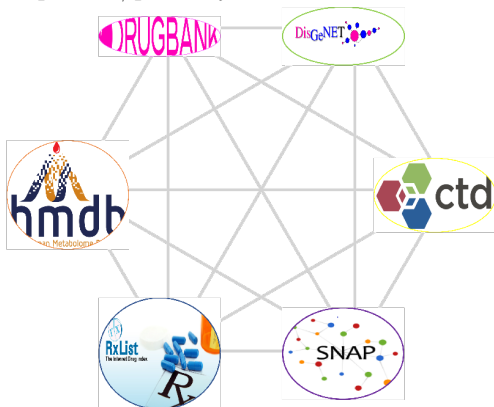


Web Scraping
Service



Structured data

Nodes: genes, SNPs, diseases, phenotypes, drugs, metabolites/compounds/pathways



Data Source Information

HMDB

Metabolites, Metabolic and Disease Pathways

- 114,003 Metabolites entries with chemical taxonomy
- ~25,000 human metabolic and disease pathways



CTD

Diseases, Synonyms and Phenotypes

- 7,212 Diseases with mapped synonyms
- 4,340 Disease Related Phenotypes



SNAP

Disease Ontology and Synonyms

- 8,803 Disease terms with related synonyms





Mining the Web

Database contents

Data Source Information

DB

Drugs, Drugs Interactions, Drug-Target Association

- 11,926 Drugs
- 18,969 Drug-Targets Associations



DGNET

Gene-Disease, Disease-Variant Associations

- 628,685 associations, between 17,549 genes and 24,166 diseases
- 210,498 associations, between 117,337 variants and 10,358 diseases



RXLIST

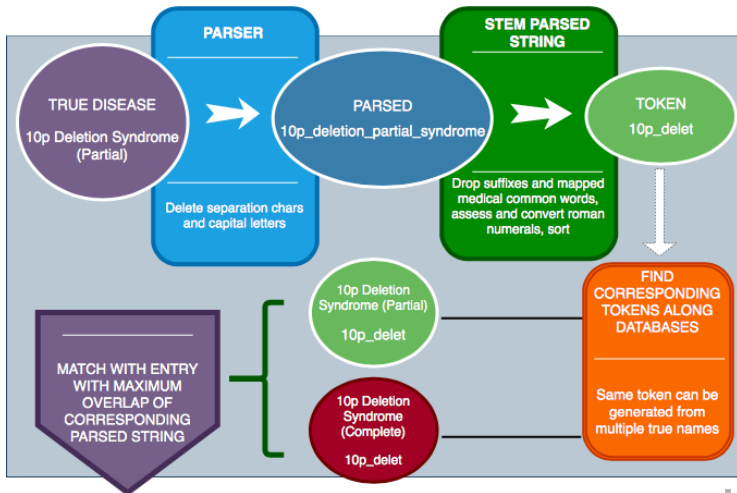
Diseases, Related, Causes and Drugs

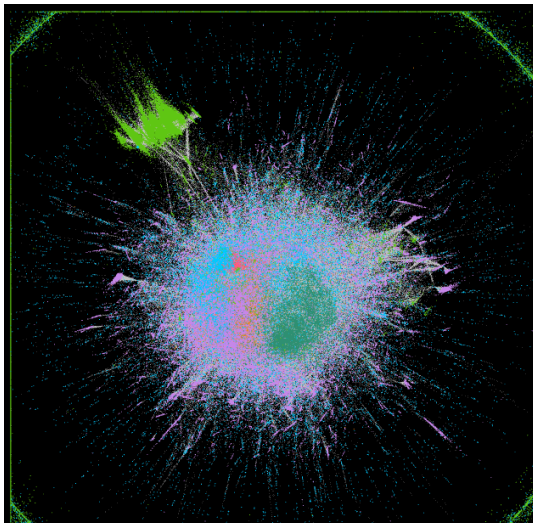
- Associations between related disease and causes
- Drug-Disease associations



String Processing

Pre-processing Pipeline





Sample Raw Info Overlap

CTD-SNAP = 24.17%

RXLIST-HMDB = 8.03%

SNAP-HMDB = 0.39%

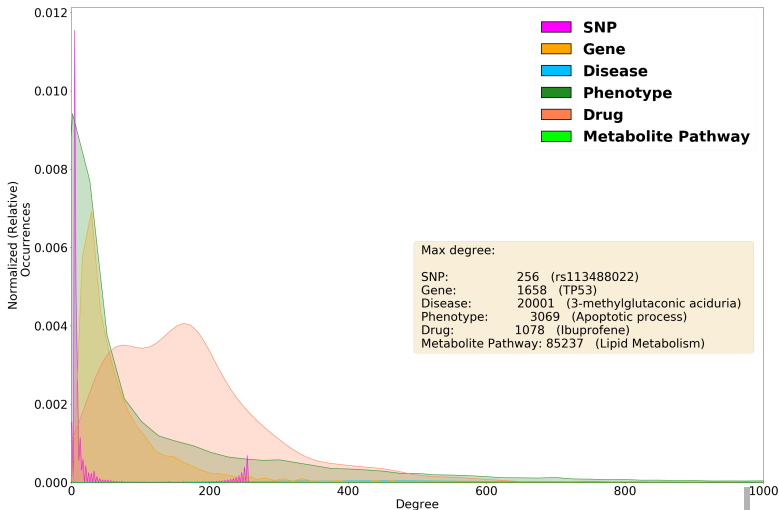
DGNET-RXLIST = 19.78%

variant	(36,01%)
Metabolite_Pathway	(35,21%)
disease	(19,1%)
gene	(5,38%)
phenotype	(4,05%)
drug	(0,25%)



Preliminary Analysis

Degree distributions





CHIMeRA as Service

- Convert Network structure to queryable DataBase (**ArangoDB**);
- Query node according (disease, gene, SNP, drug, metabolite, phenotype):
 - Query single node: neighborhood, equivalent to single database query;
 - Breadth first search: percolation from a given root node (n-th neighborhood);
 - Query multiple root nodes (**in progress**);
 - Query shortest paths between two entries (**in progress**).
- Graphical visualization of resulting networks;

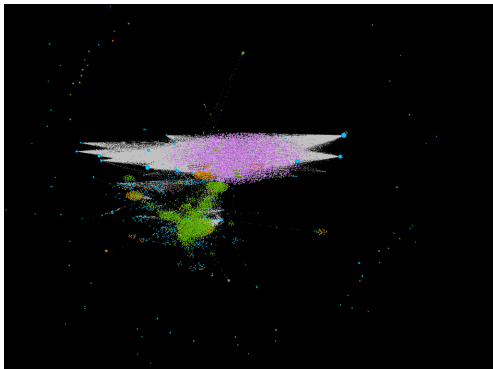
Example: Leukemia disease

Looking for “Leukemia”
disease into CHIMeRA db:

- 291 types of Leukemia
- 82 connected components

Node types:

- 838 diseases (subtypes, related);
- 2463 genes;
- 5195 phenotypes;
- 765 SNPs;
- 154 metabolite pathways;
- 40 metabolites;
- 5 drugs;





Next

- New build based on enhanced processing
- Network-of-Networks analysis
- Improve query efficiency and user interface.
- Implement as a public querable web service.

Thanks to the research group

