# Preface

## NETTAB 2012
## Workshop on "Integrated Bio-Search"



**Marco Masseroli**[1]✉, **Paolo Romano**[2], **Frédérique Lisacek**[3]

[1]Politecnico di Milano, Milan, Italy
[2]IRCCS San Martino IST, Genova, Italy
[3]SIB Swiss Institute of Bioinformatics, Geneva, Switzerland

NETTAB Workshops are a series of International meetings on "Network Tools and Applications in Biology" held annually in Italy. They are aimed at introducing participants to the most promising among those innovative Information and Communication Technologies (ICTs) that are being applied to the biomedical application domain. Workshops include many focused sessions which are devoted to tools, systems, applications, and perspectives. Keynote lectures introduce the sessions' topics, and are followed by presentations selected from among the submitted contributions after peer review by members of the Scientific Committee. Discussion is a key factor, both within sessions and in a special Panel Discussion. Tutorials and poster sessions complete the agenda of the NETTAB workshops. Each year, the workshop is focused on a different technology or domain. Since 2001, many different topics, often related to data integration issues, were discussed. These included, e.g., Standardization for data integration (Genoa, 2001), Multi agent systems (Bologna, 2002), Scientific workflows (Naples, 2005), Grid and Web Services (Santa Margherita di Pula, 2006), The Semantic Web (Pisa, 2007), Collaborative research and development (Catania, 2009), Biological wikis (Naples, 2010), and Clinical Bioinformatics (Pavia, 2011).

The NETTAB 2012 workshop, the twelfth in the series, was held in Como, Italy, on November 14-16 2012.

Its rationale is based on the consideration that the data deluge of the current post-genomic era is providing scientists with potentially very valuable information, but it makes difficult to find and extract from the increasing available high-throughput omics data those information that are most reliable, specific and most related to the biomedical questions to be answered. Such questions are increasingly complex and they often simultaneously regard many heterogeneous aspects of an organism, tissue, cell and the role of all biomolecular entities. Several of these questions can be addressed only by comprehensively searching different types of data, which generally are distributed in many heterogeneous sources. Usually, scientists explore these data by using the individual search services and tools available in Internet and they then struggle in combining the essential information in order to answers their global questions. In this context, moreover, quality and consistency checking is a central issue that should be brought up

Searching and combining all open and linked data and algorithmic sources has the potential of reshaping the scenario of current bioinformatics applications, going beyond the capabilities of conventional tools, Web services and existing search engines. Yet, it also presents new technological challenges.

Solving data integration and automatic extraction problems requires new solutions, including the use of universal URIs, efficient indexing, partial or approximate value matching, rank aggregation, continuous or push-based search, exploratory methods and context-aware paradigms, collaborative and social search, and building new efficient information retrieval approaches, based on automation of workflows too that may contribute to new "good practices" in data searching, retrieval, and integration, with the specific goal of ensuring quality of procedures, as well as their reproducibility coupled with efficiency and efficacy.

On these premises, then, the NETTAB 2012 workshop has been focused on "Integrated Bio-Search", which includes all aspects that relate to technologies, methods, architectures, systems and applications for searching, retrieving, integrating and analyzing data, information, knowledge, infrastructures, services and tools that are required to answer complex bio-medical-molecular questions.

Workshop topics included four main areas. The first area relates to data integration. It includes syntactic and semantic methods and algorithms for biological and clinical data and knowledge integration, information and knowledge retrieval, data and knowledge query, data, information and knowledge extraction, and data and knowledge mining.

The second area refers to new and optimized technologies for data management. It includes federated databases, data warehouses, and triple stores. It also includes topics as biomedical terminologies and ontologies, systems' interoperability, natural language processing, and scientific workflow processing.

Tools and platforms for molecular data management and storage, deep sequencing analysis, omics data computing, search computing, decision support, and clinical bioinformatics are the third topic area, while the fourth area includes examples of applications of these methods, technologies and tools in different biomedical domains, such as knowledge assessment, integration, discovery, and validation, drug design, diagnosis and prognosis support, and personalized medicine.

The Call for abstracts was able to attract 34 submissions for oral communications. From these submissions, the Scientific Committee of the workshop was able to select 12 oral communications, seven short oral communications, and three technological communications from industry. All submissions underwent peer review by at least two members of the Scientific Committee. At the workshop, 29 posters were also presented. Submissions for posters were also peer reviewed by one or two members of the Scientific Committee. This Supplement therefore includes about 50 abstracts, all revised according to reviews, which are grouped by submission type and ordered by first author name.

The NETTAB 2012 workshop has been a great meeting for all researchers involved in data search and integration in biology and medicine. It was possible to discuss ideas, and doubts, with such scientists as Erik Bongcam-Rudloff, Barend Mons, Eric Neumann, Alexander Kel, Katy Wolstencroft, who accepted to give invited lectures and tutorials, and many others who enthusiastically joined the workshop.

And, of course, the workshop has been a great occasion to enjoy Italian lifestyle....

# Speakers

## Keynote Speakers

**Erik Bongcam-Rudloff**
Swedish University of Agricultural Sciences, and Uppsala University, Sweden

**Barend Mons**
Leiden University Medical Center, and Netherlands Bioinformatics Center, The Netherlands

**Eric Neumann**
PanGenX, and Clinical Semantics Technologies, USA

## Tutorials

**Alexander Kel**
GeneXplain GmbH, Wolfenbüttel, Germany, and Biosoft.ru, Skolkovo Center of Bioinformatics, Novosibirsk, Russian Federation

**Katy Wolstencroft**
School of Computer Science, University of Manchester, United Kingdom

## Oral presentations

Claudio Angione
Bachir Balech
Esra Erdem
Giovanni Felici
Francesca Finotello
Matteo Gabetta
Francisco Gómez-Vela
Alejandra Gonzalez-Beltran
Claudia Gugenmus
Víctor Martínez
Marco Masseroli
Marco Muselli
Carmen Navarro
Piergiorgio Palla
Uberto Pozzoli
François Rechenmann
Fabio Rinaldi
Patrick Ruch
Bahar Sateli
Saif Ur-Rehman

# Editors

| | |
|---|---|
| **Marco Masseroli** | Politecnico di Milano, Milan, Italy |
| **Paolo Romano** | IRCCS San Martino IST, Genoa, Italy |
| **Frédérique Lisacek** | SIB Swiss Institute of Bioinformatics, Geneva, Switzerland |

## Editorial Staff

| | |
|---|---|
| **Davide Chicco** | Politecnico di Milano, Milan, Italy |
| **Arif Canakoglu** | Politecnico di Milano, Milan, Italy |

# Chairs and Conference Committees

## Chairs

| | |
|---|---|
| **Marco Masseroli** | Politecnico di Milano, Milan, Italy |
| **Paolo Romano** | IRCCS San Martino IST, Genoa, Italy |
| **Frédérique Lisacek** | SIB Swiss Institute of Bioinformatics, Geneva, Switzerland |

## Scientific committee

| | |
|---|---|
| **Francisco Azuaje** | Centre de Recherche Public de la Santé, Luxemburg |
| **Riccardo Bellazzi** | University of Pavia, Italy |
| **Olivier Bodenreider** | National Institutes of Health, USA |
| **Mario Cannataro** | University "Magna Græcia" of Catanzaro, Italy |
| **Bastien Chopard** | Swiss Institute of Bioinformatics, University of Geneva, Switzerland |
| **Marie-Dominique Devignes** | CNRS LORIA, France |
| **Christine Froidevaux** | University Paris-Sud, France |
| **Carole Goble** | The University of Manchester, United Kingdom |
| **Nicolas le Novère** | European Bioinformatics Institute, United Kingdom |
| **Ulf Leser** | Humboldt University, Germany |
| **Frédérique Lisacek** | SIB Swiss Institute of Bioinformatics, Geneva, Switzerland |
| **Paolo Magni** | University of Pavia, Italy |
| **Roberto Marangoni** | University of Pisa, Italy |
| **Marco Masseroli** | Politecnico di Milano, Milan, Italy |
| **Luciano Milanesi** | Biomedical Technologies Institute, CNR, Italy |
| **Paolo Missier** | Newcaste University, United Kingdom |
| **Heiko Muller** | Istituto Italiano di Tecnologia, Italy |
| **Norman Paton** | University of Manchester, United Kingdom |
| **Horacio Pérez-Sánchez** | University of Murcia, Spain |
| **Paolo Romano** | IRCCS San Martino IST, Genoa, Italy |
| **Patrick Ruch** | University of Applied Sciences, Geneva, Switzerland |
| **Indra Neil Sarkar** | University of Vermont, USA |

## Organising Committee

| | |
|---|---|
| **Marco Masseroli** | Politecnico di Milano, Milan, Italy |
| **Arif Canakoglu** | Politecnico di Milano, Milan, Italy |
| **Davide Chicco** | Politecnico di Milano, Milan, Italy |

## Organization Staff

Centro Studi Scientifici "Alessandro Volta", Como, Italy

Centro di Cultura Scientifica
"Alessandro Volta"

# Supporting Institutes, Scientific Societies and Projects

The workshop is held under the patronage of

**Bioinformatics Italian Society**
http://www.bioinformatics.it/

**EMBnet: the Global Bioinformatics Network**
http://www.embnet.org/

and with support from

**Politecnico di Milano, Italy**
http://www.polimi.it/

**San Martino IST, Genoa, Italy**
http://www.hsanmartino.it/

**SIB Swiss Institute of Bioinformatics**
http://www.isb-sib.ch/

**Flagship INTEROMICS Project**

**Search Computing (SeCo) Project**
http://www.search-computing.it/

**CNR BIOINFORMATCS Project**
http://www.cnr.it/sitocnr/

# Sponsors

**CRC Press**
http://www.crcpress.com/

**Camera di Commercio di Como**
http://www.co.camcom.gov.it/