

ELIXIR Text-mining Workshop / Biocreative & BeCalm Workshop

Co-located text-mining workshops

Advances in high-throughput technology have resulted in tremendous growth of biological data, and the number of research papers published continues to increase. This proves to be a great challenge for curators and researchers in assimilating scientific facts described in articles. This workshop aims to bring together experts in Text mining (TM), Data curation and Knowledge Management working in the Life Sciences domain, to formulate ideas on how to leverage text-mining infrastructures.

On April 25th, 26th and 27th, 2017, two text-mining workshops will be held in Barcelona: ELIXIR WP3 workshop (April, 25th & 26th), and BioCreative V.5 - BeCalm workshop (April, 26th & 27th).

ELIXIR-EXCELERATE workshop

This workshop, in the context of [ELIXIR-EXCELERATE](#) WP3, will specifically focus on:

- Identifying current practices in the TM community.
- How do we measure what is 'good' text mining.
- How do we share TM outputs (annotations) widely.
- Addressing gaps that prevent the uptake of TM solutions
- Identifying synergies with the larger scientific community.

The workshop will feature introductory talks and structured discussions; the output of the meeting will be a report that will form the basis of future developments for ELIXIR.

Biocreative V.5 & BeCalm workshop

BioCreative: Critical Assessment of Information Extraction in Biology is a community-wide effort for evaluating text mining and information extraction systems applied to the biological domain. Built on the success of the previous BioCreative Challenge Evaluations and Workshops (BioCreative I, II, II.5, III, 2012 workshop, IV and V).

The BioCreative V.5 BeCalm Challenge evaluation workshop will take place in Barcelona, Spain on April 26th-27th, 2017 in a collocated meeting with the ELIXIR Text-mining Workshop (25th-26th).

[For BioCreative V.5 – BeCalm](#), the selection of the tracks was driven in part by demands from both the text mining and biocuration communities, with the goal of addressing some of the major barriers to the adoption and use of text mining tools: evaluation, accessibility, interoperability, robustness and integration.

Two traditional BioCreative tracks focused on monitoring progress on the recognition of relevant bio-entities (chemicals and gene/proteins) and their evaluation against manually annotated.

A novel track called TIPS (Technical interoperability and performance of annotation servers) focuses on the technical aspects of the evaluation of continuous text Annotation Servers (ASs) for named entity recognition. Aspects that should be addressed in the context of this track during the workshop include:

- Continuous evaluation of text mining systems to promote more stable annotation tools
- Facilitate the interoperability of multiple text annotation systems at the technical level and through the design of compatible annotation schemas
- Extraction of textual content from heterogeneous sources, such as scientific literature (PubMed abstracts) or patent abstracts
- Visualization and detailed comparative evaluation assessment of automatic and manual annotations
- Harmonization of multiple biomedical text annotations
- Usage of standard evaluation metrics for the performance evaluation of component-level tasks
- Management of users corresponding to developers of text annotation systems

With the contribution of invited speakers and panel sessions this workshop will also discuss aspects related to formats, annotations and technical integration of text mining components, the experience of text mining techniques for the DARPA Big Mechanism program and the use of web services and workflows of text processing. Moreover, during this event the [OpenMinted project](#) and open call for tenders will be presented.

Programmes

[BioCreative V.5 Workshop](#) [Text-mining infrastructure requirements Workshop](#)

Organising committees

ELIXIR-EXCELERATE

- Salvador Capella, Spanish National Bioinformatics Institute (INB), Spain
- Aravind Venkatesan, EMBL-EBI, UK
- Laura Furlong, Research Programme on Biomedical Informatics, Spain
- Julien Gobeill, Swiss Institute of Bioinformatics (SIB), Switzerland
- Senay Kafkas, EMBL-EBI, UK
- Jee-Hyub Kim, EMBL-EBI, UK
- Johanna McEntyre, EMBL-EBI, UK

BIOCREATIVE & BECALM

- Cecilia Arighi, University of Delaware, USA
- Donald Comeau, National Center for Biotechnology Information (NCBI), NIH, USA
- Juliane Fluck, Fraunhofer Institute for Algorithms and Scientific Computing SCAI, Germany
- Rezarta Islamaj Dogan, National Center for Biotechnology Information (NCBI), NIH, USA
- Lynette Hirschman, MITRE Corporation, USA
- Sun Kim, National Center for Biotechnology Information (NCBI), NIH, USA
- Martin Krallinger, Spanish National Cancer Centre, CNIO, Spain
- Zhiyong Lu, National Center for Biotechnology Information (NCBI), NIH, USA
- Fabio Rinaldi, Institute of Computational Linguistics, University of Zurich, Switzerland
- Alfonso Valencia, Spanish National Cancer Centre, CNIO, Spain
- Thomas Wiegers, North Carolina State University, USA
- Cathy Wu, University of Delaware and Georgetown University, USA

- Kevin Cohen, University of Colorado, USA



INB

Spanish National
Bioinformatics Institute

cnio Centro Nacional
de Investigaciones
Oncológicas

Image not found or type unknown



Organized by:

PRB 2



is
Instituto
de Sa
Carlo



inB

Spanish National
Bioinformatics Institute



FEDER

Fondo Europeo de
Desarrollo Regional

UNIÓN EUROPEA
“Una manera de hacer Europa”



Funded by:

Image not found or type unknown



This workshop is part of a project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654021

Image not found or type unknown



Dissemination:

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 29 Mar 2024 - 15:11): <https://www.bsc.es/es/news/events/elixir-text-mining-workshop-biocreative-becalm-workshop>