

An Innovative Approach to Enhance Collaboration in the Biomedical Field

Georgia Tsiliki, Biomedical Research Foundation, Academy of Athens, Athens, Greece

Manolis Tzagarakis, Research Academic Computer Technology Institute, University of Patras, Rio Patras, Greece

Spyros Christodoulou, Research Academic Computer Technology Institute, University of Patras, Rio Patras, Greece

Sophia Kossida, Biomedical Research Foundation, Academy of Athens, Athens, Greece

Nikos Karacapilidis, Research Academic Computer Technology Institute, University of Patras, Rio Patras, Greece

ABSTRACT

Web 2.0 technologies applications have been suggested as potential enablers for the accumulation of multidisciplinary knowledge, for instance in the biomedical field. Such applications offer new ways of creating, collaborating and sharing user-generated content online. Under this context, The authors' present an innovative Web 2.0 approach that exploits prominent high-performance computing paradigms and large data processing technologies to meaningfully search, analyze and aggregate data existing in diverse, extremely large and rapidly evolving sources. The underlying tool is designed to support the entire life cycle of a biomedical collaboration, with specifically implemented services. Preliminary evaluation results are also presented and discussed.

Keywords: Biomedical Data, Collaboration, Decision-Making, Multidisciplinary Knowledge, Workbench

1. INTRODUCTION

The field of biomedical research has recently seen a vast growth in publicly available biomedical resources, including multiple types of datasets and databases, and thus statistical methodologies and analyses tools. A major advance is that now researchers have access

to complementary views of a single organism by analyzing multiple types of data, including whole genome sequencing, expression profiling and other high-throughput experiments. Those data, which are often called '-omics' data, include the genome sequencing data (genomics), the complete set of RNA transcripts produced by the genome and analysed via microarray, Real-time PCR or Next-Generation Sequencing platforms (transcriptomics), protein structures

DOI: 10.4018/ijssbt.2013010105

and function (proteomics), or any other data available for the organism under study, and provide novel views of cellular components in the biological systems (Tsiliki & Kossida, 2011). As a consequence, an enormous amount of digital content is produced everyday (i.e. information that is created, captured, or replicated in digital form as well as hundreds of analysis systems), resulting in high rates of new information being distributed and demanding attention (Karacapilidis, Tzagarakis, Christodoulou, & Tsiliki, 2012).

Most of those data sets are well organised in publicly available databases, although there are existing limitations in accessing, storing, mapping and managing the increasing amount of data available (Sullivan, Gabbard, Shukla, & Sobral, 2010), which could be overridden when supported by appropriate algorithmic analysis and software tools (Koschmieder, Zimmermann, Tribl, Stoltmann, & Leser, 2011). For instance, cloud and distributed computing, schema-free solutions, domain-specific and process-oriented programming languages or special statistical algorithmic solutions can be applied (Huttenhower, Schroeder, Chikina, & Troyanskaya, 2008; Pennisi, 2011; Baker, 2012). Apart from meaningfully mapping, emphasis has been given to algorithmically unify the data above (Lukk et al., 2010) and their supplementary views (Joyce & Palsson, 2006). However, given the biological or statistical question of interest, choosing the right datasets, databases and tools for a given project is difficult even for an expert (Pennisi, 2011).

Within this environment, research in biostatistics and biomedical fields has become increasingly multidisciplinary and collaborative in nature (Lee, 2007; Baker, 2012). The progressively specialized resources show that the way forward is to form collaboration teams in order to address complex research questions. Such multidisciplinary teams would better meet challenges relative to various problems such as how to store, access, analyze and integrate multiple types of data (Pennisi, 2011); or, how to work with multiple databases simultaneously (Finholt, 2003); or even, how to make

data accessible and usable to life sciences researchers (Sullivan, Gabbard, Shukla, & Sobral, 2010). In addition, tools facilitating sense- and decision-making by appropriately capturing the collective intelligence that emerges during such collaboration are lacking. Biomedical researchers need such tools to efficiently and effectively collaborate and make decisions by appropriately assembling and analyzing enormous volumes of complex multi-faceted data residing in different sources. Supporting team collaboration under such circumstances is still considered as a challenging task (Spencer, Zimmerman, & Abramson, 2011). Towards this goal, we present a Web-based approach to support communities of bio-scientists during their scientific collaboration, which is being developed in the context of the Dicode FP7 EU ongoing project (<http://dicode-project.eu/>).

2. SUPPORTING COLLABORATION IN THE BIOMEDICAL FIELD

The number of publicly available bioinformatics tools and databases on the Web is rapidly increasing over the last few years (Brazas, Yamada, & Ouellette, 2009), together with the increasing demand for programmatic interfaces that would allow bioinformatics tools and databases to be linked in an automated analysis pipeline. Remote access to such resources via web pages has been provided by major bioinformatics organizations, such as KEGG (<http://www.genome.jp/kegg/>), GO (<http://www.geneontology.org/GO.cite.shtml>), NCBI (<http://www.ncbi.nlm.nih.gov/>), which have utilized the standardized set of web services protocols to provide programmatic access to their computational resources (Goble et al., 2010).

Overall, there is a growing demand for programmatic interfaces to link bioinformatics tools and databases and allow computing resources to be readily reused, repurposed and extended by other scientists (Baker, 2012). There are examples of repositories such as BioGRID (Stark et al., 2006), the Biochemical

12 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/innovative-approach-enhance-collaboration-biomedical/78392

Related Content

Implementation of n-gram Methodology for Rotten Tomatoes Review Dataset Sentiment Analysis

Prayag Tiwari, Brojo Kishore Mishra, Sachin Kumar and Vivek Kumar (2017). *International Journal of Knowledge Discovery in Bioinformatics* (pp. 30-41). www.irma-international.org/article/implementation-gram-methodology-rotten-tomatoes/178605

Bifurcation Analysis of a Model Accounting for the 14-3-3s Signalling Compartmentalisation

S. Nikolov, J. Vera and O. Wolkenhauer (2013). *Bioinformatics: Concepts, Methodologies, Tools, and Applications* (pp. 851-859). www.irma-international.org/chapter/bifurcation-analysis-model-accounting-signalling/76099

An Update on the H7N9 Strain of the Influenza A Virus

Dimitrios Vlachakis, Argiro Karozou, Spyridon Champeris Tsaniras and Sophia Kossida (2013). *International Journal of Systems Biology and Biomedical Technologies* (pp. 59-66). www.irma-international.org/article/an-update-on-the-h7n9-strain-of-the-influenza-a-virus/97742

Discovering Lethal Proteins in Protein Interaction Networks

Kar Leong Tew and Xiao-Li Li (2009). *Biological Data Mining in Protein Interaction Networks* (pp. 183-202). www.irma-international.org/chapter/discovering-lethal-proteins-protein-interaction/5565

Similarity Searching of Medical Image Data in Distributed Systems: Facilitating Telemedicine Applications

Amalia Charisi, Panagiotis Korveis and Vasileios Megalooikonomou (2013). *Methods, Models, and Computation for Medical Informatics* (pp. 58-77). www.irma-international.org/chapter/similarity-searching-medical-image-data/73071