Accelerating Biomedical Research through Semantic Web Services

Artemis Chaleplioglou, Biomedical Research Foundation, Academy of Athens, Athens, Greece

ABSTRACT

In view of the fact that there is a lot of biomedical research data, rapidly accumulating to relevant repositories, there is an urgent need these data to be delivered, as soon as possible, to the specific scientific audience interesting in it. Unfortunately, the current database technologies often isolate data rather than making it easily and freely accessible. A considerable effort by the information scientists is needed to process the resources that meet the scientific query criteria as well as to index and present them as useful metadata. Taking into account that biomedical data are mostly hidden from the public eye, often stored in not indexed databases or libraries and inaccessible by standard search engines, the retrieval, storing, annotating, and qualification of health information remain major challenges. The evolution of the World Wide Web from a collection of unstructured and predominantly human readable data into the Semantic Web of knowledge with meaningful relationships between resources and machine readable data will significantly improve our ability to conduct bioinformatics analyses and to make better clinical decisions that positively affect healthcare outcomes. To this end novel semantic web services arise, which depend on markup ontologies in order to make biological and clinical data logical analysis computational and reasonable processed through the utilization of appropriate algorithms. Herein, we discuss the use of these technologies for the efficient and reliable retrieval of meaningful biomedical data from the relevant resources and repositories.

Keywords: Algorithms, Biomedical Libraries, Data Mining, Library Services, Semantic Web

INTRODUCTION

Information in the World Wide Web accumulates rapidly with increasing rate. The size of the internet was increased by 37% within 2013, with more than $8 \times 10^8$ websites recorded (http://news.netcraft.com/) and nearly $1.7 \times 10^9$ web pages were indexed collectively by January 2014 (http://www.worldwidewebsize.com/). Importantly, most web pages almost doubled the size of their content in the same year (http://httparchive.org/). At the same time nearly 900,000 new records were added in 2013, twice as much compared to 2003, a decade ago, to the registry of 22,000,000 publication records listed in MEDLINE/PubMed database (http://www.nlm.nih.gov). Information overload, stove-
pipe systemic organization of information, poor content aggregation due to insufficient maintenance, and the Deep Web issue, with trillions of not indexed web pages located to institutional databases, libraries and repositories, inaccessible to search engines, are responsible for the complexity and heterogeneity in biomedical data. The pessimistic perspective suggests that inevitably a crisis will arise in biomedical progress because of the accumulation of unevaluated data, which in turn will significantly delay both basic scientific and applied clinical research.

Today’s global health security issues are attributed to the emergence of global diseases threats of both communicable and non-communicable nature (See Figure 1). In specific, HIV remains the world’s leading infectious killer, accompanied by malaria and other foodborne germs. The globalization of travel and food supply facilitates the spread of new microbes, like the novel Middle East Respiratory Syndrome Coronavirus (MERS-CoV), the norovirus in US and poxvirus in Georgia (http://www.cdc.gov). Furthermore, because of the widespread antibiotic resistance to infections, the prevention of prescription drug abuse and overdose is significant. The acceleration of biological science capabilities raises bioethical questions, and the risk of inadvertent or intentional release of pathogens is high.

Mortality of women and children in low and middle income countries account for over 95% of all maternal and child deaths because of the lack of sanitary conditions. The high occurrence of non-transmittable diseases like cardiovascular disease, cancer, diabetes, chronic respiratory disease are the predominant etiology of human mortality worldwide because of life-style, toxic substance contamination and environmental

Figure 1. Semantic web is connecting global biomedical challenges to current biomedical approaches
Intermediation Structures in Electronic Healthcare Portals
www.igi-global.com/chapter/intermediation-structures-electronic-healthcare-portals/9234?camid=4v1a