Xournals

Academic Journal of Bioinformatics

ISSN UA | Volume 01 | Issue 01 | June-2018

A Study of Bio-ontologies

Vipul Joshi[1]

Available online at: www.xournals.com

Received 22th February 2017 | Revised 27th February 2018 | Accepted 19th April 2018

Abstract:

Ontology has been taken up by biologists that is widely used in biological and biomedical research which is themselves as a means to consistently annotate features from genotype to phenotype. Artifacts called ontologies in medical informatics that have been used for a longer period of time to produce controlled vocabularies for coding schemes. It is combination of four main features that is present in almost all ontologies in which success is lies: provision of standard identifiers for classes and relations that represent the phenomena within a domain; provision of a vocabulary for a domain; provision of metadata that describes the intended meaning of the classes and relations in ontologies; and the provision of machine-readable axioms and definitions that enable computational access to some aspects of the meaning of classes and relations. In this paper, discuss about the main features of ontologies, categories of gene ontologies and future trend of ontologies.

Keywords: *Ontology, Phenotype, Genotype, Vocabulary*

Authors:

1. *Maulana Azad College, Aurangabad, Maharashtra, INDIA*

Introduction

Bioinformatics have relatively recent achievement which is Bio-ontologies that effort toward an efficient organization and distribution of biological data. Bio-ontologies provide a structure and controlled vocabulary through which data especially those gathered through sequencing and genomics but results are increasing from other types of research. The use of ontologies began in biological sciences around 1998 with the development of gene ontology. Their interest and activity are increase in the area to merit national and international coordination efforts such as Open Biomedical Ontologies Foundry or National Center for Biomedical Ontologies in 2007. Ontology is referred to as mainly ‘ontologies’ as well as ‘vocabularies’ and ‘thesauri’ that provide several main features and these features are used in all their applications, which is as follows:

- Classes and relations are referred to as identifier such as an Internationalized Resource Identifier (IRI), a Uniform Resource Identifier (URI), or a database identifier string.
- A domain vocabulary may be a list of terms associated with the ontology’s classes and relations.
- Textual definitions and descriptions that provide additional information about what kind of things a class or relation.
- Formal definitions and axioms that provide a computational counterpart to textual definitions which is accessed and exploited automatically using specialized software and axioms about domain.

Here, discuss ontologies containing features and these features provide a ‘functional’ perspective on ontologies. These features are improve data analysis in biology and biomedicine, which is discuss in below table.

Table: Features provided by ontologies in Biological and Biomedical Research

Ontology feature	**Utility in research**
Classes and relations	In ontologies, use of standard identifiers for classes and relations, is what enables data integration across multiple databases because the same identifiers can be used across multiple, disconnected databases, files or web sites.
Domain vocabulary	Through labels associated with classes and relations, ontologies provide a domain vocabulary that can be exploited for applications ranging from natural language processing, creation of user interfaces, etc.
Metadata and descriptions	In ontologies, classes are linked with Textual definitions, descriptions, examples and further metadata that are what enable domain experts to understand the precise meaning of class in the ontology. The meaning of classes in ontologies are consistent understood by definitions and related metadata.
Axioms and formal definitions	Formal definitions and axioms enable automated and computational access to (some parts of) the meaning of a class or relation.

Here, we also discuss the categories of gene ontologies which is divided into three parts that is discuss into below:

- Molecular function
- Biological process
- Cellular component

Molecular function: It is defined as biological activity that include specific binding to ligands or structures which is a gene product of activity. It is applies to the capability that a gene product carries as a potential. It is describes only that what is done without specifying the where or when event actually occurs. Examples of narrower functional terms are ‘adenylate cyclase’ or ‘Toll receptor ligand’ and example of broad functional terms are ‘enzyme’, ‘transporter’ or ‘ligand’.

Biological process: In which biological process refers to a biological objective to which the gene or gene product contributes. A process is accomplished through one or more ordered assemblies of molecular functions. This process involve a chemical or physical transformation, in the sense that something goes into process and something is different that come out of it. Example of broad (high level) biological process terms are "cell growth and maintenance" or "signal transduction" and more specific (lower level) process terms are 'translation', 'pyrimidine metabolism' or 'Cyclic adenosine monophosphate (cAMP) biosynthesis'.

Cellular component: It is refers to the place in the cell where a gene product is active. These terms reflect the understanding of eukaryotic cell structure. It is true for other ontologies but not for all terms which are applicable to all organisms; the set of terms is meant to be inclusive. It includes terms such as 'ribosome' or 'proteasome', 'nuclear membrane' and 'Golgi apparatus', specifying where multiple gene products would be found.

Molecular function, biological process and cellular component are all attributes of genes, gene products or gene-product groups. All of these are assigned as an independently and simply recognizing that biological process, molecular function and cellular location represent independent attributes is by itself clarifying in many situations. The relationships between a gene product (or gene-product group) to biological process, molecular function and cellular component are one-to-many, reflecting the biological reality that a particular protein may function in several processes, contain domains that carry out diverse molecular functions, and participate in multiple alternative interactions with other proteins, organelles or locations in the cell.

Future and role of biomedical ontologies

Biomedical ontologies provide the semantic structure that supports integration and comparison of large, complex data sets in biology. To practical knowledge of ontology development and application has been on capture of domain knowledge such as function and location. Ontologies support uniform data encoding thus providing semantic integration of information derived from multiple resources such as free-text publications or functional annotations of sequences from multiple providers. Exchange and integration of data is the facilities of semantic integration that is between the resources and is essential for future development of semantic web services. It is useful as terminologies for annotation systems and that they can be used in formal ontological representations in biomedicine as these resource develop. Bio-ontologies have significant impact on data integration, access and analysis through their use in capturing and structuring biological data.

Review of Literature

Jupp et al., concluded that in life science ontologies, Simple Knowledge Organization System (SKOS) vocabulary has some adoption-in particularly the labeling and mapping properties. It allow existing tools that consume SKOS access to terminological information of bio-ontologies by taking standard approaches. It is need for a better tool support to enable life scientists to work with SKOS more easily. This paper demonstrates that how to separate the protein with concern knowledge and use presentation into layers and adopting standards such as SKOS offer new possibilities for new data.

Blake, Botstein and Butler 2000, by all eukaryotes, large fraction of genes are made by genomic sequencing that is specifying the core biological functions. Biological role knowledge of such shared proteins in one organism can often be transferred to other organism. The aim of gene ontology consortium is to produce a dynamic, controlled vocabulary that can be applied to all eukaryotes which have a knowledge of gene and protein in cells is accumulating and changing.

Blake and Bult 2006, stated that generation of genome-scale data sets are supported by new technologies and sets of sequences, sequence variants, transcripts and proteins; genetic elements underpinning understanding of biomedicine and disease. These data is manage by information systems and by analysis of these data, biological knowledge have to come, heterogeneous data sets for new biologically relevant patterns, to generating hypotheses for experimental validation and ultimately, to building models of how biological system work. Data integration and comparative genomics are two key approaches in role of bio-ontologies that is effective interpretation of genome-scale data sets.

Huntley et.al 2014, in this paper, describing the some examples of gene ontology that is linked with annotation can change over time, using example how to manage the changes of UniPort. For analysis to understand that who make use of gene ontology and why these changes occur in order to make the appropriate conclusions for their interpretations.

Malone et.al 2016, dictated that metadata is describing by bio-ontologies that is important tool which is an increasingly important consideration as scientific community aims for open, reusable data. Ontology is choices by pick and not use as straightforward that is demonstrated by number of times and authors are asked to recommend a particular ontology for a given problem. First to select the bio-ontology is to understand requirements before deciding to engage with a particular ontology or indeed before minting one's own ontology.

Conclusion

Bio-ontologies play an important role is able to recognize the vital biology and medical research that is use within biomedicine is now a mainstream activity. It is widely used to deliver vocabularies for describing data and future aspect will see greater analysis of data due to increasing formality of these ontologies. By using this formality, will see the growth of reference ontologies in biology and biomedicine. Gene ontologies found accurately resembles modules that drives gene and protein interaction networks, leading to a data-driven way for validating ontology. By doing the scientific research, increasing the methods for improving the use of ontologies for evaluating ontology quality.

References:

"Taking a view on bio-ontologies". Available at: https://pdfs.semanticscholar.org/694b/5db6306f7a983f64df849ccc7637214bd8d5.pdf

Blake, Judith A., and Carol J. Bult. "Beyond the Data Deluge: Data Integration and Bio-Ontologies." *Journal of Biomedical Informatics*, vol. 39, no. 3, 2006, pp. 314–320.

Blake, Judith A., et al. "Gene Ontology: Tool for the Unification of Biology." *2000 Nature America Inc*., vol. 25, May 2000, pp. 25–29.

Bodenreider, Olivier, and Robert Stevens. "Bio-Ontologies: Current Trends and Future Directions." *Briefings in Bioinformatics*, vol. 7, no. 3, 2006, pp. 256–274.

Hoehndorf, R., et al. "The Role of Ontologies in Biological and Biomedical Research: a Functional Perspective." *Briefings in Bioinformatics*, vol. 16, no. 6, Oct. 2015, pp. 1069–1080.

Huntley, Rachael P, et al. "Understanding How and Why the Gene Ontology and Its Annotations Evolve: the GO within UniProt." GigaScience, vol. 3, no. 1, 2014.

Leonelli, Sabina. "Bio-Ontologies as Tools for Integration in Biology." *Biological Theory*, vol. 3, no. 1, 2008, pp. 7–11.

Malone, James, et al. "Ten Simple Rules for Selecting a Bio-Ontology." *PLOS Computational Biology*, vol. 12, no. 2, Nov. 2016.