Automated extraction and semantic analysis of mutation impacts from the biomedical literature

Submitted by nona [1] on Mon, 2012-06-18 06:40

- Open Mutation Miner [2]
- Corpora [3]
- Bioinformatics [4]
- Ontology [5]
- Text Mining [6]

Title	Automated extraction and semantic analysis of mutation impacts from the biomedical literature
Publication Type	Journal Article
Year of Publication	2012
Authors	Naderi, N. [7], and R. Witte [8]
Refereed Designation	Refereed
Journal	BMC Genomics
Volume	13
Issue	Suppl 4
Pagination	S10
Date Published	06/2012
ISSN	1471-2164

Abstract

BACKGROUND: Mutations as sources of evolution have long been the focus of attention in the biomedical literature. Accessing the mutational information and their impacts on protein properties facilitates research in various domains, such as enzymology and pharmacology. However, manually curating the rich and fast growing repository of biomedical literature is expensive and time-consuming. As a solution, text mining approaches have increasingly been deployed in the biomedical domain. While the detection of single-point mutations is well covered by existing systems, challenges still exist in grounding impacts to their respective mutations and recognizing the affected protein properties, in particular kinetic and stability properties together with physical quantities.

RESULTS: We present an ontology model for mutation impacts, together with a comprehensive text mining system for extracting and analysing mutation impact information from full-text articles. Organisms, as sources of proteins, are extracted to help disambiguation of genes and proteins. Our system then detects mutation series to correctly ground detected impacts using novel heuristics. It also extracts the affected protein properties, in particular kinetic and stability properties, as well as the magnitude of the effects and validates these relations against the domain ontology. The output of our system can be provided in various formats, in particular by populating an OWL-DL ontology, which can then be queried to provide structured information. The performance of the system is evaluated on our manually annotated corpora. In the impact detection task, our system achieves a precision of 70.4%-71.1%, a recall of

71.3%-71.5%, and grounds the detected impacts with an accuracy of 76.5%-77%. The developed system, including resources, evaluation data and end-user and developer documentation is freely available under an open source license at http://www.semanticsoftware.info/open-mutation-miner [9].

CONCLUSION: We present Open Mutation Miner (OMM), the first comprehensive, fully open-source approach to automatically extract impacts and related relevant information from the biomedical literature. We assessed the performance of our work on manually annotated corpora and the results show the reliability of our approach. The representation of the extracted information into a structured format facilitates knowledge management and aids in database curation and correction. Furthermore, access to the analysis results is provided through multiple interfaces, including web services for automated data integration and desktop-based solutions for end user interactions.

URL	http://www.biomedcentral.com/1471-2164/13/S4/S10 [10]
DOI	10.1186/1471-2164-13-S4-S10 [11]
Copyright	Copyright © 2012 Naderi and Witte; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0 [12]), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Impact Factor	4.21
History	Published: 18 June 2012

Attachment	Size
1471-2164-13-S4-S10.pdf [13]	1.96 MB

Source URL (retrieved on *2025-12-22 07:51*):
https://www.semanticsoftware.info/biblio/automated-extraction-and-semantic-analysis-mutation-impacts-biomedical-literature

Links:
[1] https://www.semanticsoftware.info/users/nona
[2] https://www.semanticsoftware.info/category/project/mutation-miner
[3] https://www.semanticsoftware.info/category/topic/nlp/corpora
[4] https://www.semanticsoftware.info/category/topic/bioinformatics
[5] https://www.semanticsoftware.info/category/topic/ontology
[6] https://www.semanticsoftware.info/category/topic/text-mining
[7] https://www.semanticsoftware.info/biblio/author/22
[8] https://www.semanticsoftware.info/biblio/author/1
[9] http://www.semanticsoftware.info/open-mutation-miner
[10] http://www.biomedcentral.com/1471-2164/13/S4/S10

[11] http://dx.doi.org/10.1186/1471-2164-13-S4-S10

[12] http://creativecommons.org/licenses/by/2.0

[13] https://www.semanticsoftware.info/system/files/1471-2164-13-S4-S10.pdf

Semantics for the Masses