Learning Reference-enriched Approach towards Large scale Active Ontology Alignment and Integration

Qiong Cheng
School of Engineering and Technology
Miami Dade College
Miami, USA
chengqiong7@gmail.com

Oleg Ursu, Tudor I. Oprea
Translational Informatics Division,
Department of Internal Medicine,
University of New Mexico
Albuquerque, NM 87131,
oursu, toprea@salud.unm.edu

Stephan Schurer
Department of Molecular and Cellular Pharmacology
University of Miami
Miami, USA
stephan.schurer@gmail.com

Abstract—With the increasing number of ontologies being designed to represent and manage knowledge in all sorts of sectors, ontology alignment and integration become more and more important in aggregating intelligent efforts on homogenous and heterogeneous data. From the computational perspective, it is challenging due to the ubiquitous existence of diverse classifications of same data. In this paper, we propose an active ontology integration and alignment system, which plugs in expandable learning reference context pool. In the reference context pool, we have integrated WordNet, MeSH, and external curated mapping sources (ICD9 to SNOMED-CT) with an extension to injecting UMLS. The active ontology integration and alignment system takes account of not only subsumption tree but also directed acyclic graph underlying ontologies. It allows 1) finding exact one-to-one matching terms of pairwise ontologies, 2) finding inexact one-to-one term mappings, where two terms have at least a concept in common on basis of the lexical context, and 3) finding one-to-many concept mappings, where one concept can be lexically mapped to the combination of multiple exclusive concepts.

Keywords—ontology, ontology alignment, ontology integration, bioassay data

I. INTRODUCTION

With the emerging growth of biomedical big data, intelligent efforts have been exerted to classify data and represent the annotated data in ontologies. However due to limited resources in the ontology development time or the possible delay in sharing reliable data and performing further exploration, ontologies are developed without being able to coordinate with all necessary and related ontologies. The integration activities are sometimes in need of spanning data in different dimensions and at different levels or even in different language systems, which has become more impractical. The problem of ontologies representing diverse classifications of same data is also significant, and prone to confusion. While the National Library of Medicine's Unified Medical Language System (UMLS) [14] has offered the UMLS Metathesaurus which is a distribution of clinical terminologies and classifications and provides a few of pairwise mappings of same type of ontologies, the above-mentioned problem still exists and do not be fundamentally solved.

Especially, disease or disease-related ontology alignment is critical in guiding and improving the prevention and management of diseases, and facilitating disease studies and drug discovery. The alignment algorithms and tools are meaningful for the cross-species gene-disease association studies or other fundamental bioinformatics research, for example, identifying significant biomarkers associated with specific phenotype profiles of rare diseases in the systematic way.

In this paper, we propose an active ontology integration and alignment system, which plugs in expandable learning reference context pool. The proposed method aims at helping with the set of issues.

Our early study of aligning the Observational Medical Outcomes Partnership (OMOP) FDA indication with SNOMED-CT is encouraging. Among the 1428 * 3925 term pairs, we found 47% exact or partial matches. In the study, we expanded our reference pool from the WordNet to MeSH and also to external mappings of ICD9 to SNOMED-CT since the coverage of medical terms in WordNet is limited. The alignment on the basis of WordNet as the only reference performed unpleasantly but combining WordNet with MeSH brought forth a reasonable number of mappings.

The Systematized Nomenclature of Medicine--Clinical Terms (SNOMED-CT) consists of 230,768 active terms and 1,215,361 active concepts. In order to speed up the computation, we limited our search into the "disorder"-related terms in SNOMED-CT, which reduce the SNOMED-CT to 68,653 unique terms and 141,514 concepts and thus, the search space is reduced to ~1,320 * 141,514 pairs. Assuming that computing a pair needs half a second, the serial algorithm may take at least a year to complete a run.

978-1-5090-3050-7/17/$31.00 ©2017 IEEE

We then upgraded our software into the parallel version OAS v1, which simply partitioned the subsumption trees into diverse subtrees and run our algorithm in parallel on pairwise subtrees coming from different ontologies. We launched and run our programs in 500 computing nodes of the IBM IDataPlex system (cluster computing system) and took ~22 hours to run, including the pre- and post-processing. In the study, we found 5,288 unique term-to-term mappings which have 10,411 SNOMED-CT mapping concepts involved. We used different approaches to evaluate our preliminary results.

We further upgraded our software in order to handle not only subsumption tree but also DAG (directed acyclic graph), where the concepts sharing same destination term would be treated in the unique context. The DAG-based version requires more delicate algorithms to partition the DAG in a balanced way since an ideal data partitioning for the parallel computing would prefer to grouping data into independent subsets of approximately same scale of the size.

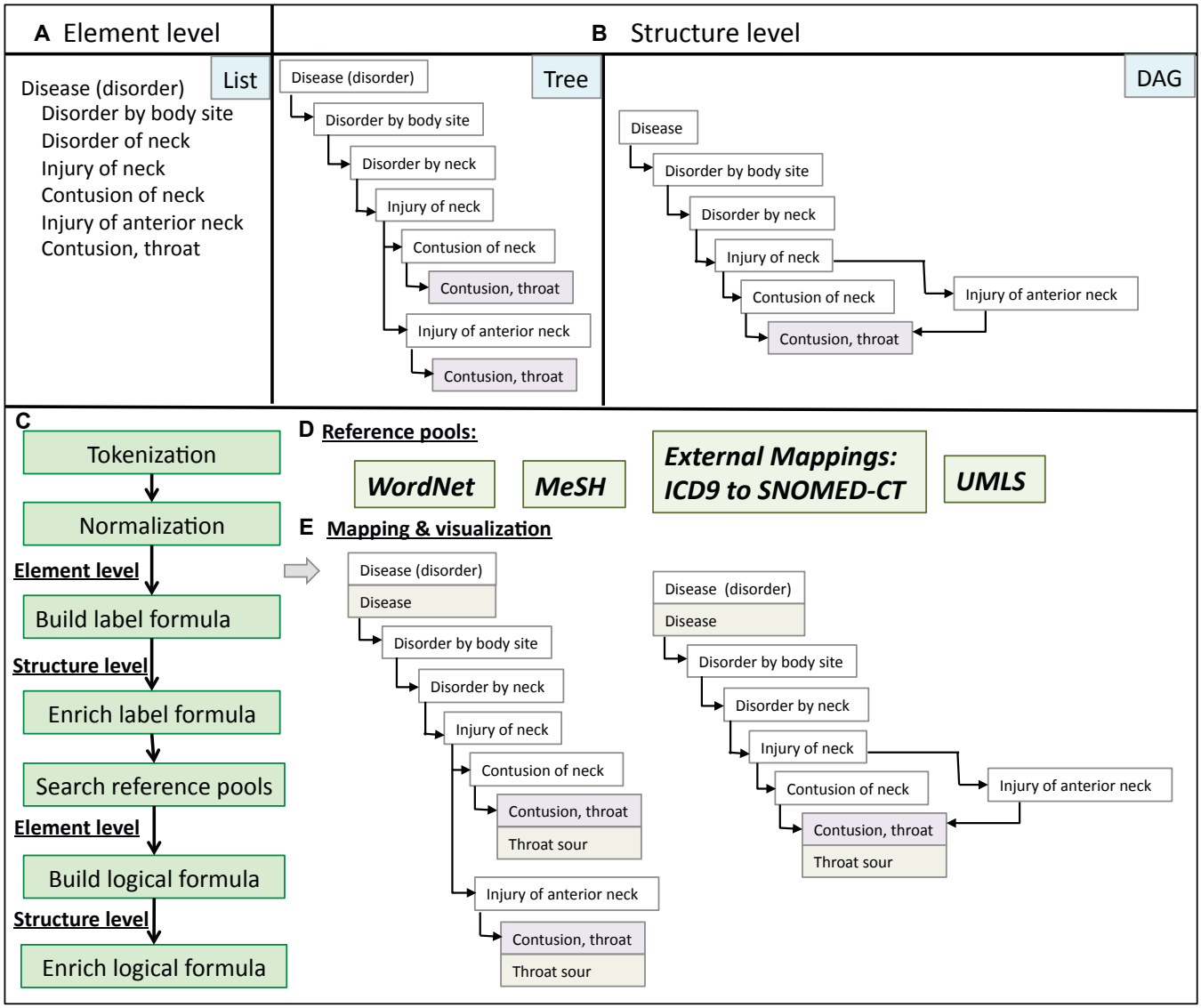

Figure 1. Ontology alignment enriched by the expandable reference pool. A. Element level view of subsumptions underlying the ontology. B. Structure level views of subsumptions underlying the ontology (a tree structure view and a DAG structure view). C. Process flow of building logical formulae of ontology terms. D. Reference pools under different combinations of lexical databases, such as WordNet, MeSH, external mapping from ICD9 to SNOMED-CT, and UMLS. E. Mapping examples on the basis of tree-to-tree alignment and DAG-to-DAG alignment

II. LEARNING REFERENCES-ENRICHED APPROACH

We propose an active ontology alignment and integration framework which allows 1) finding exact one-to-one matching terms of pairwise ontologies, 2) finding inexact one-to-one term mappings, where two terms have at least a concept in common on basis of the lexical context, and 3) finding one-to-many concept mappings, where one concept can be mapped to

the combination of multiple exclusive concepts. The alignment results can be enriched and polished by not only an assembly factory of alignment algorithms but also continually learning reference context pool, both of which are main components of the proposed system. In the algorithm factory, multiple alignment algorithms will be included, for example, the lexical term-to-term mapping, hierarchy-based structure mapping, and the concept-to-concept edit distance [40]. A specified algorithm or the optimal algorithm in the factory will be iteratively run and the process will finally converge when no more references can be used. The advantage of our system arises from the integration of WordNet [19], MeSH[24], and external curated mapping sources (ICD9 to SNOMED-CT), and UMLS[14]. It would also be benefited by an interactive learning framework which is designed for verifying and reusing those reasonable mappings generated from previous runs.

In our current version, we assume that the classification represented by an ontology is organized as a hierarchy, which is a directed acyclic graph, including term nodes and subsumption associated edges linking term nodes. In the structure, each term is labeled by a natural language sentence, which helps compute the lexical meaning of that node as a Description Logic (DL) formula. The DL formula associated to each node is subsumed by the formula of the node above, which is attributed by the subsumption relationship. For example, the term path associated with "Contusion, throat", Disease (disorder) -> Disorder by body site -> Disorder of neck -> Injury of neck -> Injury of anterior neck -> Contusion, throat, demonstrates the underlying subsumption relationship.

The subsumption relationship in the ontology can be viewed in the element level or structure level (Refer to Figure 1A and 1B). The element level takes all terms independently, disregarding the subsumption relationship. Thus, the ontology represents a list of isolated terms.

In the tree-based structure level, a path from the root term to the leaf term represents a concept. For a term, there may be multiple paths associating with it, where the leaf term and possibly some other terms are shared among those paths. In the tree-based structure, all concepts are assumed to be independent and complete. Thus, the shared nodes are cloned and their multiple copies appear in the tree-based structure. For example, the disease term "Contusion, throat" has been cloned and shown in two term paths (see Figure 1B.Tree).

In the DAG-based structure level (DAG, directed acyclic graph), the complete concept of a term is built from the induced subgraph, which includes all concept paths through the root term to the destination term (see Figure 1B.DAG). The DAG-based structure breaks the above-mentioned assumption that each path is independent and complete in the tree-based structure. It views in a background context all concepts that are associated with a term. For example, the disease term "Contusion, throat" is shown as either an "injury of anterior neck" or a "Contusion of neck" (see see Figure 1B). The DAG-based structure has more expressive power in specifying a term than the tree-based structure.

In our software, we first preprocessed each ontology in order to build the logic formulae of ontology terms, which takes account of the context that terms stay. Then we searched for synonyms in the reference pool(s) for each lemma, which is the smallest unit in a logical formula (also called token). By examining the generalization-specification relationship between any two concepts, we determined whether their formulae are logically equal or a "crosstalk" happens in the pairwise alignment. The "crosstalk" represents the case that a supper class of term i in the ontology A is more specific than a subclass of term j in ontology B and at the same time a subclass of the term i is more specific than a supper class of the term j.

In the above-mentioned preprocess (see Figure 1C), we tokenized all terms by removing unmeaningful words or symbols, for example, "of", "and", "on", "to", "a", "an", '(', ')', punctuations, and white spaces. Then we normalized all tokens into the single canonical form and built the label formula by the token conjunction (the "and" logic, also annotated by the ∧ sign). In the structure level, each term is enriched by the formula conjunction. Namely the label formula being identified in the element level is replaced with the formula conjunction, combining all label formulae in the path through the root term to the destination term by the "and" logic. For example, in the structure level, the label formula of the disease term "Contusion, throat" is disease ∧ disorder ∧ body ∧ site ∧ injury ∧ contusion ∧ anterior ∧ neck ∧ throat.

By searching in the reference pool, we found the synonyms for each lemma in the formula. We refined the label formulae by creating the logic formulae on the basis of the reasoning of semantic network underlying the reference pool(s). The theory of description logic allows us to effectively reason the term-to-term relationship and reduce the consistency checking. In our current software, we have imported WordNet, MeSH, and external mappings from ICD9 to SNOMED-CT. We have used our software in three case studies, where we aligned different types of ontologies. A mapping example is shown in the bottom-right panel, where "Throat sour" is lexically mapped to "Contusion, throat".

III. THREE CASE STUDIES

A. Finding exact term-to-term mapping of DTO (Drug target ontology) to ChEMBL

With the detailed sequence of the Human Genome available, the druggable genome has attracted more and more interest of scholars. The druggable genome is defined as the subsets of genes in the human genome that express proteins capable to bind drug-like small molecules [19]. The Illuminating the Druggable Genome (IDG) project (http://targetcentral.ws/) has

developed an integrated knowledgebase for the human druggable genome with a goal of improving our understanding of the properties and functions of proteins that are currently unannotated within the four most commonly drug-targeted protein families: the G-protein coupled receptors (GPCRs), nuclear receptors, ion channels, and protein kinases.

In the context of the IDG program, the Drug Target Ontology (DTO) was developed in [38], which provides a structured knowledge resource of drug target classifications and relevant annotations for these four protein families in the IDG knowledgebase. We extracted the druggable proteins and its classification tree from the DTO, which is an integrative ontology of coordinating, mapping and integrating protein target concepts across a multitude of relevant data resources and external ontologies with individual curation.

Another widely used database in public is ChEMBL [39]. This database includes a large number of well-annotated literature-curated bioactivity data, including small molecule protein binding data. We extracted subsumption trees and terms from ChEMBL [39]. By mapping 1,724 DTO protein Uniprot ID entries to 6,962 ChEMBL entries, we identify 957 common protein entries between them.

B. Coordinating BioAssay Ontology (BAO) with BARD

BARD has offered a curated hierarchical dictionary of terms which act as an ontology to facilitate managing and capturing annotated preclinical data from multiple scientific disciplines, with a special emphasis on bioassay data [1]. The top-to-bottom hierarchy in BARD allows that a class has diverse subclasses and lower level classes have been subsumed under their higher-level classes. Namely, under the subsumption, a class could have multiple direct superclasses subsuming it, and no superclass has been subsumed by a subclass.

BAO 2.0 is an advanced version of BioAssay Ontology (BAO) [2], where a formal Description Logic (DL) ontology has been developed. The DL-based BAO aims at building a comprehensive and systematic set of constructs for logic reasoning and effective prediction. It is organized into several component classes and object properties [11]. Each of classes includes a subsumption tree.

The common properties allow us to model them as directed acyclic graphs, where vertices (leaf nodes and non-leaf nodes) and edges (lines connecting the vertices) act as the terms and the rule(s) of the ontologies [3]. Due to the general subsumption rule, we further built text-based hierarchy tree for both ontologies, where all existing subpaths from the root to any other nodes have been listed.

Our analysis started with the lexical matching of all nodes between BARD and BAO on basis of the *s-match* tool [5,6], where exact match, synonym match, and pattern match of node's label concepts have been measured as the relationship of either "more generality" or "more specificity". We further casted the node-node matching into two groups, exact match and inexact match. In the version, a term-term lexical match or mapping is identified if and only if a pair of leaf nodes is an exact match or a pair of their superclasses can be lexically matched. We evaluated the inexact match by the following procedure.

For an inexact match (u, v), where u is a BAO term and v a BARD term, we walked through hierarchy trees individually from both BAO 2.0 and BARD and searched for lexical matching node pairs through all paths passing from the root to the corresponding nodes u and v through intermediate nodes.

Through the study, we found 856 unique mapping pairs, among which 622 are from exact mapping and 234 from inexact mappings. We also realized that lexicon mapping may perform more pleasantly if biological context information and ontology hierarchy description is introduced reasonably since existing general WordNet-oriented NLP-based ontology matching tools do not work well for the specific bioassay knowledge domain, especially handling numbers and abbreviation in the specific domain. Considering that it is an interesting computational problem with wide applications, we have been proposing new algorithms to improve the precision of the domain-specific lightweight ontology matching [4,7-11].

C. Coordinating OMOP with SNOMED-CT

The SNOMED CT is a core clinical healthcare terminology maintained by International Health Terminology Standards Development Organization (IHTSDO®) [16]. 521,845 SNOMED CT terms are organized into hierarchies consisting of 4,658,378 relationships, out of which 1,159,402 are is-a subtype relationships. The subtype relationship is same as subsumption relationship, representing that upper-level concepts in the hierarchy are more general than lower-level concept. Each term is the destination of a sequence of one or more is-a relationships from the root term. The generalization-specification relationship commits each concept to a unique meaning with specific a formal logic associated. We have downloaded the SNOMED Snapshot core release (20150131) from [16].

The OMOP [18] includes 877,012 concepts, among which 774,824 are unique concepts, which means that multiple concepts get involved in a term and which motivated us to consider not only tree-to-tree mapping but also DAG-to-DAG mapping. FDA Indication is kind of a disease ontology embedded in OMOP as one of 55 types of vocabularies.

We carried out our current alignment routine of mapping OMOP FDA indications concepts to SNOMED-CT, after extracting and building the ontology from SNOMED-CT database flat text file as well as retrieving a subsumption tree from OMOP FDA indications database. The experiment

started with a serial version and then evolved into a parallel version. In the serial version, the scalability was an issue even if we limited our search into the "disorder"-related terms in SNOMED-CT, which reduced the search space to ~1,320 * 141,514 pairs. The serial version would take at least a year to complete a run if computing a pair needs half a second. In the parallel version, we simply partitioned the subsumption trees into independent subtrees and run our algorithm in parallel on pairwise subtrees coming from different ontologies. We launched and run our programs in 500 computing nodes of the IBM IDataPlex system (cluster computing system) and took ~22 hours to run, including the pre- and postprocessing. .

For validating our approach, we selected a subset of concepts respectively from OMOP FDA indication and SNOMED-CT and performed the same routine of our ontology alignment. In the experiment, concepts with key words of "mental", "Immune", and "nervous" have been extracted. The subset-to-subset alignment identified 43 unique term-to-term pairs (out of 1428 * 3,925 pairs), which cover 89 concept-to-concept mappings.

In the experiment, we have explored different learning references, for example, WordNet, Medical Subject Headings (MeSH), and external mappings. Our observation is that the coverage of medical terms in WordNet is quite limited. The reference pool performed better with the MeSH terminology being injected into WordNet electronical dictionary. It led to 1,487 unique term mappings being found, which covered 4,767 mapping concept pairs. One good mapping example is that the SNOMED-CT term 'Disorder of speech and language development (disorder)' is mapped to the OMOP term 'Communication Disorders'.

CONCLUSION

Ontology alignment and integration plays a critical role in knowledge engineering, especially in the context of big data. Disease ontology alignment can be facilitative in systematic analysis of cross-species phenotype profiles and it also has wide usage in healthcare applications. In the paper, we propose an active ontology integration and alignment system, which plugs in an expandable learning reference context pool. We practiced ontology alignment and integration in three case studies, where WordNet, Mesh, and external mapping of ICD9 to *SNOMED-CT* have been injected in the pool. Further experiments are in need of making full use of the Unified Medical Language System (UMLS) Metathesaurus and its semantic networks in the pool to promote ontology alignment and integration.

ACKNOWLEDGMENT

This work was supported by grants U54CA189205 (Illuminating the Druggable Genome Knowledge Management Center, IDG-KMC) and U54HL127624 (BD2K LINCS Data Coordination and Integration Center, DCIC). The IDG-KMC (http://targetcentral.ws/) is a component of the Illuminating the Druggable Genome (IDG) project (https://commonfund.nih.gov/idg) awarded by the NCI. The BD2K LINC DCIC is awarded by the National Heart, Lung, and Blood Institute through funds provided by the trans-NIH Library of Integrated Network-based Cellular Signatures (LINCS) Program (http://www.lincsproject.org/) and the trans-NIH Big Da-ta to Knowledge (BD2K) initiative (https://datascience.nih.gov/bd2k). Both IDG and LINCS are NIH Common Fund projects. Q. C. has also been supported by the grant P031S150046 funded by the U.S. Department of Education.

REFERENCES

[1] Kroeze, W. K., Sassano, M. F., Huang, X. P., Lansu, K., McCorvy, J. D., Giguère, P. M., ... & Roth, B. L. (2015). PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nature structural & molecular biology*, *22*(5), 362-369.

[2] Besnard J, Ruda GF, Setola V, Abecassis K, Rodriguiz RM, Huang XP, Norval S, Sassano MF, Shin AI, Webster LA, Simeons FR, Stojanovski L, Prat A, Seidah NG, Constam DB, Bickerton GR, Read KD, Wetsel WC, Gilbert IH, Roth BL, Hopkins AL. (2012) Automated design of ligands to polypharmacological profiles. *Nature. 2012 Dec 13;492(7428):215-20. doi: 10.1038/nature11691.*

[3] Guilio Rastelli and Luca Pinzi. (2015) Computational polypharmacology comes of age. Frontiers in Pharmacology. Volume 6 | article 157. doi: 10.3389/fphar.2015.00157.

[4] Uhlén M et al, 2015. Tissue-based map of the human proteome. *Science* PubMed: 25613900 DOI: 10.1126/science.1260419

[5] Mestres, Jordi, et al. (2008) Data completeness—the Achilles heel of drug-target networks. *Nature biotechnology* 26.9: 983-984.

[6] A.P. Bento, A. Gaulton, A. Hersey, L.J. Bellis, J. Chambers, M. Davies, F.A. Kruger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos and J.P. Overington (2014) The ChEMBL bioactivity database: an update. Nucleic Acids Res., 42 1083-1090.

[7] Stephan C. Schürer and Steven M. Muskal (2012) Kinome-wide Activity Modeling from Diverse Public High-Quality Data Sets. J Chem Inf Model. 2013; 53(1): 27–38 DOI: 10.1021/ci300403k

[8] LINCS. http://lincs.hms.harvard.edu/db/

[9] Roth, B.L., Sheffler, D.J. & Kroeze, W.K. (2004) Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. Nat. Rev. Drug Discov. 3, 353–359.

[10] Wesley K. Kroeze and Bryan L. Roth. (2012) Polypharmacological Drugs: " Magic Shotguns " for Psychiatric Diseases. Peters, Jens-Uwe, ed. Polypharmacology in Drug Discovery. Somerset, NJ, USA: John Wiley & Sons.

[11] Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., … Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research, 13(11), 2498–2504. http://doi.org/10.1101/gr.1239303

[12] Santos, R., Ursu, O., Gaulton, A., Bento, A.P., Donadi, R.S., Bologa, C., Karlsson, A., Al-Lazikani, B., Hersey, A., Oprea, T.I., & Overington, J.P. A Comprehensive Map of Molecular Drug Targets. Nature Rev. Drug Discov. doi:10.1038/nrd.2016.230 (2016)

[13] Przydzial, M. J., Bhhatarai, B., Koleti, A., Vempati, U., & Schürer, S. C. (2013). GPCR ontology: development and application of a G protein-coupled receptor pharmacology knowledge framework. Bioinformatics. 2013 Dec 15;29(24):3211-9. doi: 10.1093/bioinformatics/btt565. Epub 2013 Sep 29.

[14] Lin, Y., Mehta, S., McGinty, H. K., Turner, J. P., Vidovic, D., Forlin, M., ... & Mathias, S. L. (2017). Drug Target Ontology to Classify and Integrate Drug Discovery Data. bioRxiv, 117564.

[15] Cannon, D.C., Yang, J.J., Mathias, S.L., Ursu, O., Mani, S., Waller, A., Schürer, S.C., Jensen, L.J., Sklar, L.A., Bologa, C.G. and Oprea, T.I., 2017. TIN-X: target importance and novelty explorer. Bioinformatics, p.btx200.

[16] Ursu, O., Holmes, J., Knockel, J., Bologa, C.G., Yang, J.J., Mathias, S.L., Nelson, S.J. and Oprea, T.I., 2017. DrugCentral: online drug compendium. Nucleic acids research, 45(D1), pp.D932-D939.

[17] Loging, W.T. ed., 2016. Bioinformatics and Computational Biology in Drug Discovery and Development. Cambridge University Press.

[18] Acharya, A. and Garg, L.C., 2016. Drug Target Identification and Prioritization for Treatment of Ovine Foot Rot: An In Silico Approach. International journal of genomics, 2016

[19] Russ, A. P., & Lampel, S. (2005). The druggable genome: an update. Drug discovery today, 10(23), 1607-1610.

[20] Kolaja, K. (2012). Drug discovery: computer model predicts side effects. Nature, 486(7403), 326-328.

[21] Lounkine, E., Keiser, M.J., Whitebread, S., Mikhailov, D., Hamon, J., Jenkins, J., Lavan, P., Weber, E., Doak, A.K., Côté, S. and Shoichet, B.K., 2012. Large scale prediction and testing of drug activity on side-effect targets. Nature, 486(7403), p.361.

[22] Hughes, J. P., Rees, S., Kalindjian, S. B., & Philpott, K. L. (2011). Principles of early drug discovery. British journal of pharmacology, 162(6), 1239-1249.

[23] Kobourov, S. G. (2004). Force-directed drawing algorithms.

[24] M. Horridge. OWLViz. http://protegewiki. stanford.edu/wiki/OWLViz, 2010.

[25] Liebig, T., & Noppens, O. (2005). OntoTrack: A semantic approach for ontology authoring. Web Semantics: Science, Services and Agents on the World Wide Web, 3(2), 116-131.

[26] Hop, W., de Ridder, S., Frasincar, F., & Hogenboom, F. (2012, March). Using Hierarchical Edge Bundles to visualize complex ontologies in GLOW. In Proceedings of the 27th Annual ACM Symposium on Applied Computing (pp. 304-311). ACM.

[27] Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. Scientific american, 284(5), 28-37.

[28] García-Peñalvo, F. J., Colomo-Palacios, R., García, J., & Therón, R. (2012). Towards an ontology modeling tool. A validation in software engineering scenarios. Expert Systems with Applications, 39(13), 11468-11478.

[29] Storey, M. A., Noy, N. F., Musen, M., Best, C., Fergerson, R., & Ernst, N. (2002, January). Jambalaya: an interactive environment for exploring ontologies. In Proceedings of the 7th international conference on Intelligent user interfaces (pp. 239-239). ACM.

[30] S. Falconer. OntoGraf. http://protegewiki. stanford.edu/wiki/OntoGraf, 2010.

[31] Falconer, S. M., Callendar, C., & Storey, M. A. (2010, October). A visualization service for the semantic web. In International Conference on Knowledge Engineering and Knowledge Management (pp. 554-564). Springer, Berlin, Heidelberg.

[32] Vercruysse, S., Venkatesan, A., & Kuiper, M. (2012). OLSVis: an animated, interactive visual browser for bio-ontologies. BMC bioinformatics, 13(1), 116.

[33] L. Wachsmann. OWLPropViz. http://protegewiki. stanford.edu/wiki/OWLPropViz, 2008.

[34] Bosca, A., Bonino, D., & Pellegrino, P. (2005, December). OntoSphere: more than a 3D ontology visualization tool. In Swap.

[35] Guo, S. S., & Chan, C. W. (2010, May). A tool for ontology visualizaiton in 3D graphics: Onto3DViz. In Electrical and Computer Engineering (CCECE), 2010 23rd Canadian Conference on (pp. 1-4).

[36] M. Sintek. OntoViz. http://protegewiki. stanford.edu/wiki/OntoViz, 2007.

[37] Eklund, P., Roberts, N., & Green, S. (2002). Ontorama: Browsing rdf ontologies using a hyperbolic-style browser. In Cyber Worlds, 2002. Proceedings. First International Symposium on (pp. 405-411). IEEE

[38] Lin, Y., Mehta, S., McGinty, H. K., Turner, J. P., Vidovic, D., Forlin, M., ... & Mathias, S. L. (2017). Drug Target Ontology to Classify and Integrate Drug Discovery Data. bioRxiv, 117564.

[39] A.P. Bento, A. Gaulton, A. Hersey, L.J. Bellis, J. Chambers, M. Davies, F.A. Kruger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos and J.P. Overington (2014) The ChEMBL bioactivity database: an update. Nucleic Acids Res., 42 1083-1090.

[40] Euzenat, J., & Shvaiko, P. (2007). Ontology matching (Vol. 18). Heidelberg: Springer.