

REVIEW

Open Access



The eXtensible ontology development (XOD) principles and tool implementation to support ontology interoperability

Yongqun He^{1*} , Zuoshuang Xiang¹, Jie Zheng², Yu Lin³, James A. Overton⁴ and Edison Ong⁵

Abstract

Ontologies are critical to data/metadata and knowledge standardization, sharing, and analysis. With hundreds of biological and biomedical ontologies developed, it has become critical to ensure ontology interoperability and the usage of interoperable ontologies for standardized data representation and integration. The suite of web-based Ontoanimal tools (e.g., Ontofox, Ontorat, and Ontobee) support different aspects of extensible ontology development. By summarizing the common features of Ontoanimal and other similar tools, we identified and proposed an “eXtensible Ontology Development” (XOD) strategy and its associated four principles. These XOD principles reuse existing terms and semantic relations from reliable ontologies, develop and apply well-established ontology design patterns (ODPs), and involve community efforts to support new ontology development, promoting standardized and interoperable data and knowledge representation and integration. The adoption of the XOD strategy, together with robust XOD tool development, will greatly support ontology interoperability and robust ontology applications to support data to be Findable, Accessible, Interoperable and Reusable (i.e., FAIR).

Keywords: Ontology, Interoperability, eXtensible ontology development, Software, Ontoanimal tools, Ontofox, Ontobee, Ontorat, Semantic alignment, And ontology design pattern

Background

In informatics, an ontology is a set of computer- and human-interpretable terms and relations that represent entities and their relations in a specific domain of the world. Hundreds of biological/ biomedical ontologies have been developed in the last two decades. The Open Biological and Biomedical Ontologies (OBO) Foundry is a collaborative initiative aimed at establishing a set of ontology development principles and incorporating ontologies following these principles in an evolving non-redundant and interoperable suite [1]. The OBO library currently includes over 160 ontologies covering >3 million terms in biological and clinical domains. NCBO BioPortal [2] has over 400 ontologies including both

OBO and non-OBO ontologies. Given hundreds of ontologies developed, a critical issue is the lack of ontology interoperability, preventing the seamless understanding and exchange of semantic information between different resources.

Ontologies are widely used in different areas [3, 4], including: (1) Naming “things”; (2) Knowledge base construction, e.g., the Ontology of Vaccine Adverse Events (OVAE) representing the knowledge of adverse events induced by FDA-licensed vaccines [5]; (3) Data exchange, e.g., BioPAX for representing molecular and cellular pathways and facilitating the exchange of biological pathway data [6]; (4) Data integration, e.g., the Ontology for Biomedical Investigations (OBI) [7] for integrative representations of data in various areas of life-science and clinical investigations; (5) Data analysis, as exemplified by the wide usage of the Gene Ontology (GO) [8] to support high-throughput gene expression data analyses; (6) Natural language processing [9, 10]; (7) Metadata standard generation [11–13]. (8) Information retrieval and new knowledge discovery [14–16].

* Correspondence: yongqunh@med.umich.edu

¹Unit for Laboratory Animal Medicine, Department of Microbiology and Immunology, Center for Computational Medicine and Bioinformatics, University of Michigan Medical School, Ann Arbor, MI, USA
Full list of author information is available at the end of the article

To support various needs in ontology development and applications, different software programs have been developed. The Protégé OWL editor [17] is likely the most popular tool for manual processing and editing of ontology OWL documents. However, manual ontology development is typically tedious and inefficient, especially when the structure of ontology is enormous. Over the years, we have developed a collection of web-based “Ontoanimal” tools including Ontofox [18], Ontodog [19], Ontorat [20], Ontobee [21], Ontobeeep [22], Ontobull [23], Ontokiwi [24], and Ontobat [20]. Each Ontoanimal tool has its specific functions, and the collective use of these tools enables users to cover the full development of ontology and linked data (i.e., data published on the Web that is explicitly defined, machine-readable, and interlinked with external data sets [25]), including: extracting ontology subsets for term reuse and semantic alignment, providing ontology community views, adding and editing multiple ontology terms, visualizing and comparing ontology terms, supporting community editing and discussion, and creating ontology-based linked data. The back-end He group RDF triple store serves as the default ontology RDF triple store for the OBO Foundry ontologies [21]. Complementary to Protégé, Ontoanimal tools are widely used for efficient and flexible ontology development without requiring programming skills. For example, according to Google Analytics and Google Scholar, Ontobee has been used by over 77,000 users from 181 countries, Ontofox has been used by over 17,000 users from 147 countries, and Ontoanimal tools have been cited by >400 publications in the last 5 years.

The Ontoanimal tools and other similar tools have significantly enhanced the speed and quality of ontology development and improved ontology interoperability. Given an increasing number of these tools, it would be important to identify the common features of these tools. After retrospective examination and careful summary of these tools, we realized the most common feature of these tools being their support for “extensible” ontology development. Such extensibility is crucial to increase the interoperability among the ever increasing number of ontologies. However, a systematic view of “extensible” ontology development is not available. Thus, we propose the “eXtensible Ontology Development” (XOD) strategy and four XOD principles in this paper. Such an XOD strategy is complementary to the OBO principles and the OBO goal of achieving interoperable ontology suite [1], and it is also complementary to the ten simple rules proposed for biomedical ontology development [26]. We also believe that the adoption of the XOD strategy and principles support the FAIR Guiding Principles proposal that all research data should be Findable, Accessible, Interoperable and Reusable (FAIR) for both machine and human users [27].

XOD: eXtensible ontology development

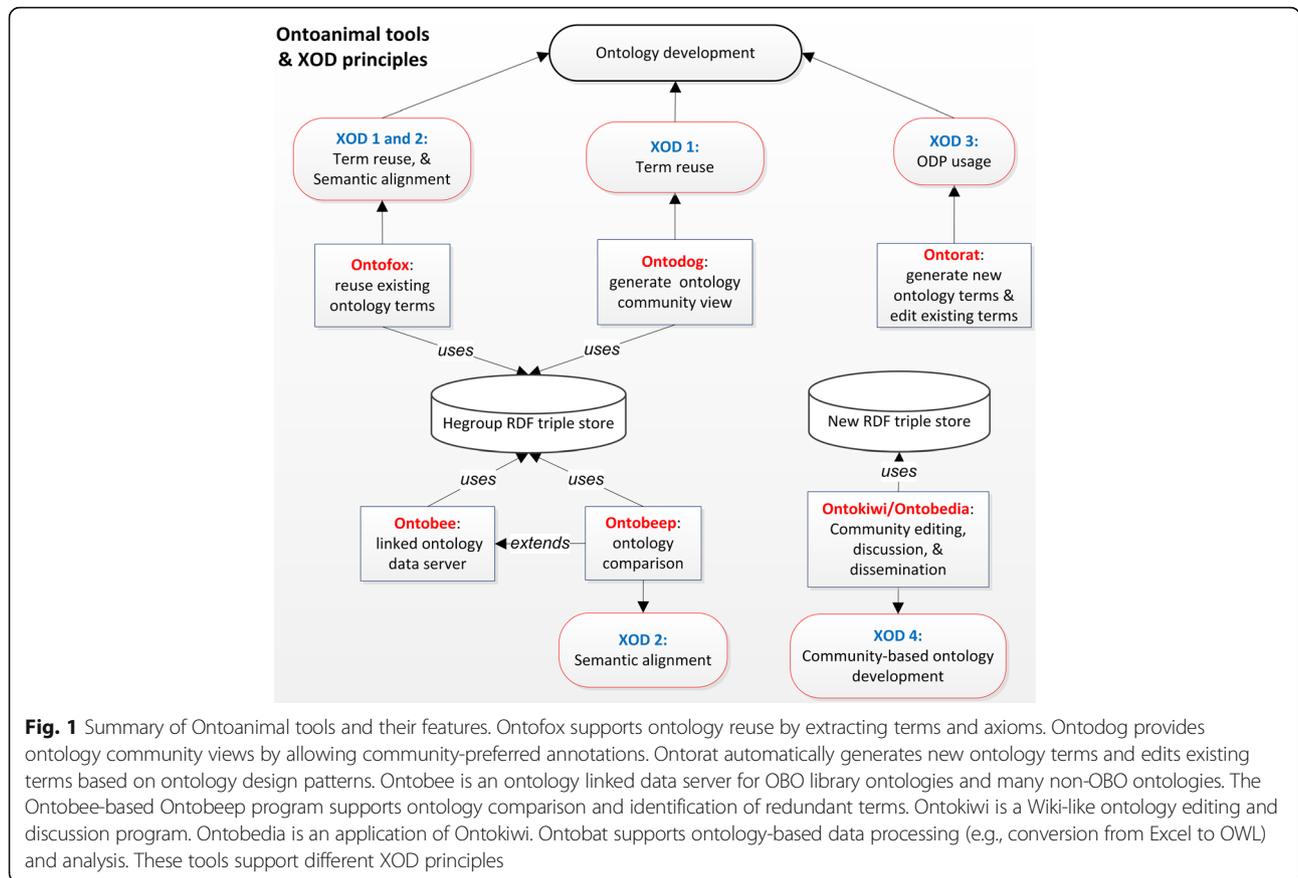
In information technology, extensible describes something (e.g., a program or protocol) that is designed so that users/developers can expand or add to its capabilities with no or minimal change in the system’s internal structure and data flow. For example, extensibility is a primary feature of the eXtensible Markup Language (XML) system. Being “eXtensible”, XOD contains four key principles that are extensible at different levels of ontology development (Fig. 1):

- (i) Ontology term reuse. Instead of reinventing the wheel when generating new ontologies, XOD emphasizes the reuse of terms from existing reliable ontologies that are well constructed and commonly used by the ontology community [1, 28, 29].
- (ii) Ontology semantic alignment. For ontology interoperability, it is important to align imported terms from existing ontologies and newly added terms with the same semantics.
- (iii) ODP usage for new term generation and existing term editing. Instead of adding one term at a time, XOD emphasizes the addition or editing of a group of terms based on ontology design patterns (ODPs).
- (iv) Community extensibility. While the development of an ontology might be initiated by a small group with one or a few use cases, the ontology should be co-developed and applied to more use cases by more people in a broader community.

Ontoanimal tools (Fig. 1) and many other programs support XOD principles (Table 1). In the following sections, different principles and associated tools are described with details.

XOD 1: Ontology term reuse

Reusing terms from reliable reference ontologies is better than reinventing the wheel to generate new terms in an ontology [18, 30]. The reference ontology should be registered in an ontology library (e.g., OBO Foundry) to make these terms more findable, accessible, and reusable. To improve reusability and extensibility, ontology terms in the reference ontologies should be expressive and generalizable and endure consistency checking and evaluation. To maintain ontology interoperability, ontology term mapping is often used to map terms from different ontologies with the same meaning [31]. Compared to ontology term reuse, term mapping is less ideal since it is time-consuming, often inaccurate, redundant, and increases maintenance cost and confusion. Given multiple ontologies without using the ontology term reuse strategy, ontology mapping becomes a core task for ontology interoperability [32]. The wide usage of the



term reuse principle would make the mapping among different ontologies unneeded.

An initial method of ontology term reuse was to import a full ontology, which was not ideal since it might import too many unrelated terms. Instead of importing external ontologies as a whole, the Minimum Information to Reference an External Ontology Term (MIREOT) strategy, introduced by OBI developers [30], proposes the usage of the minimal information of an external ontology term that is of direct interest to a target ontology [30]. Specifically, MIREOT suggests the following minimal set: (1) source ontology URI; (2) source term URI; and (3) target direct superclass URI. With the set of information, the source ontology term can be extracted to under the target direct superclass. Since it is often hard to maintain semantic consistency among ontologies, the popular MIREOT strategy provides a simple solution with possible semantics loss.

Ontofox, Ontodog, and Ontobull support term reuse. Originally named OntoFox, Ontofox was the first web tool to support the MIREOT strategy (Fig. 2) [18]. Ontofox is able to quickly and easily fetch user-specified terms and their annotations from source ontologies and assign them under defined superclass(es) in target ontologies (Fig. 2a and b). Ontofox also extends MIEROT by retrieving semantical axioms with different options (see next section). Ontodog is also able to extract a subset of ontology terms and axioms [19]. Unlike plain text definition in Ontofox, Ontodog uses Excel input files to identify terms to retrieve. To match possible updates of source ontologies, e.g., the upper-level Basic Formal Ontology (BFO) [33], Ontobull is developed for automatic conversion and updating [23].

Several other tools also support ontology term reuse (Table 1). The Protégé MIREOT plugin [34] and

Table 1 XOD principles and supporting software programs

XOD principle #	XOD principle names	Tool name
XOD 1	Ontology term reuse	Ontodog, Ontofox, OntoMATON, Protégé MIREOT plugin, ROBOT
XOD 2	Semantic alignment	Ontobeeep, Ontofox, ROBOT
XOD 3	ODP usage	MappingMaster, Ontorat, Populous, ROBOT, TermGenie, Webulous
XOD 4	Community extensibility	Ontodog, Ontokiwi/Ontobedia, WebProtege

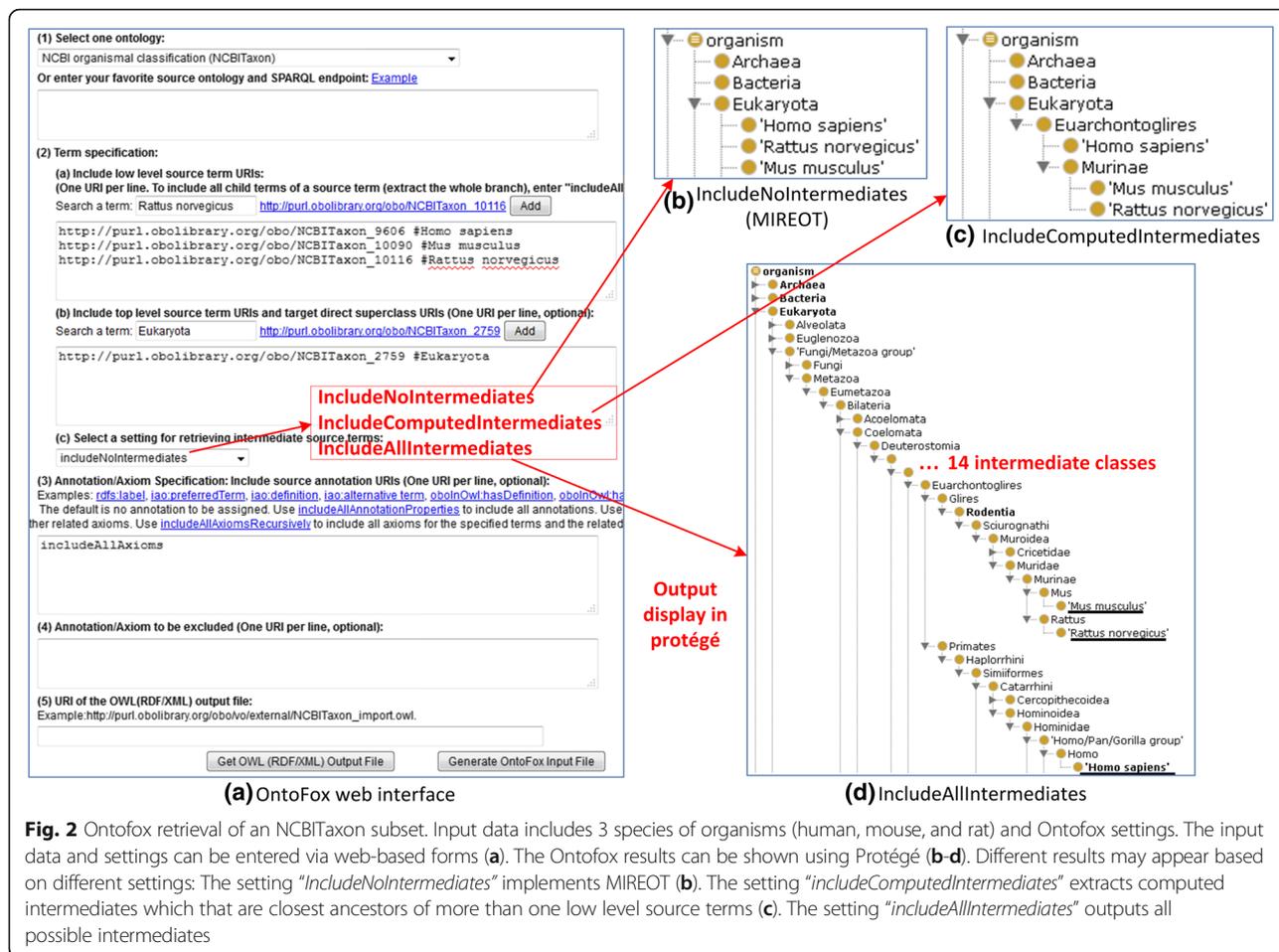


Fig. 2 Ontofox retrieval of an NCBITaxon subset. Input data includes 3 species of organisms (human, mouse, and rat) and Ontofox settings. The input data and settings can be entered via web-based forms (a). The Ontofox results can be shown using Protégé (b-d). Different results may appear based on different settings: The setting "IncludeNoIntermediates" implements MIREOT (b). The setting "includeComputedIntermediates" extracts computed intermediates which that are closest ancestors of more than one low level source terms (c). The setting "includeAllIntermediates" outputs all possible intermediates

OntoMaton [35] support term reuse as a plugin of the Protégé OWL editor or Google Spreadsheets, respectively. ROBOT is a command-line Java tool supporting the extraction of ontology terms and subsets [36]. ROBOT also has many other features and supports multiple XOD principles (Table 1) as described below [36].

To better support ontology reuse and community-based ontology development, it would also be valuable to have the authors of the source ontology know of the reuse of a term in their ontology. The Ontobee program [21] includes a feature in the web page of an ontology term that shows all the other ontologies reusing the term, which supports ontology interoperability.

XOD 2: Ontology semantic alignment

The XOD 2 principle proposes to align imported ontology terms and newly added terms with the same or compatible semantics. Such semantic alignment has two specific meanings. First, in addition to the term reuse in XOD 1, the semantic relations among reused terms should also be reused and aligned. If different relation types (i.e., object properties) mean the same thing, they should be merged. Correspondingly, the axioms of

reused terms and any additional terms specified in the axioms should also be retrieved and imported. Second, the semantics related to newly developed terms should be aligned and compatible with imported semantics, and the same or compatible relations be used in the new ontology. If a well-defined relation already exists, we should reuse the relation instead of defining another relation with the same meaning. Such semantic alignments support ontology semantic interoperability.

Ontofox and Ontodog support ontology semantic alignment. Ontofox and Ontodog extract semantic axioms and terms related to user-specified terms from the source ontologies. Given different options, Ontofox allows the computation and extraction of (i) intermediate terms that are the shared parent terms of multiple low level terms (Fig. 2c), or (ii) all intermediate terms between the required terms and a top level term (Fig. 2d). These subset semantic axioms and terms can then be retrieved and become a part of the new ontology. Note that manual intervention and judgment may still be needed now to ensure the semantic alignment between retrieved subset and target ontology semantics [37]. It will also be important to have computer-supported

semantic capture and synchronization of ontology evolution and updates. To foster reliability, an overall formal evaluation and consistency checking would be needed.

XOD 3: ODP-based ontology development

An Ontology Design Pattern (ODP) represents a reusable solution to solve a recurrent modeling problem in the context of ontology engineering [20, 38, 39]. ODPs provide extensible representations of entities and relations, make ontologies more maintainable, and improve ontology quality. This XOD 3 principle requires an ODP-based strategy to develop and edit new terms, annotations, and relations. This principle extends XOD 1 and XOD 2 and provides a specific, feasible, and robust mechanism to achieve interoperable ontology term generation/annotation and semantic consistency.

The Ontorat program (<http://ontorat.hegroup.org>) supports ODP-based creation of new ontology terms, annotations and logical axioms [20]. Fig. 3 illustrates

an example of using Ontorat to add new terms, annotations, and axioms to the Ontology of Adverse Events (OAE) [40]. Ontorat uses reusable ODPs (Fig. 4a) to automatically generate and edit ontology terms and axioms and provides term annotations. A specific ODP can be used to derive an Excel template of different terms/annotations and a set of rules that define the relations among those terms/annotations (Fig. 3b). The Ontorat template, similar to a QTT (Quick Term Template) originated by OBI developers [41], can be populated with specific terms or annotations to define or annotate specific ontology terms, or generate axioms illustrating logic relations between ontology terms. With the support of the Ontorat settings (Fig. 3c), the populated template spreadsheet can then be converted into an OWL file with newly generated ontology terms and axioms (Fig. 3d and e). The setting and template files can also be saved and reused.

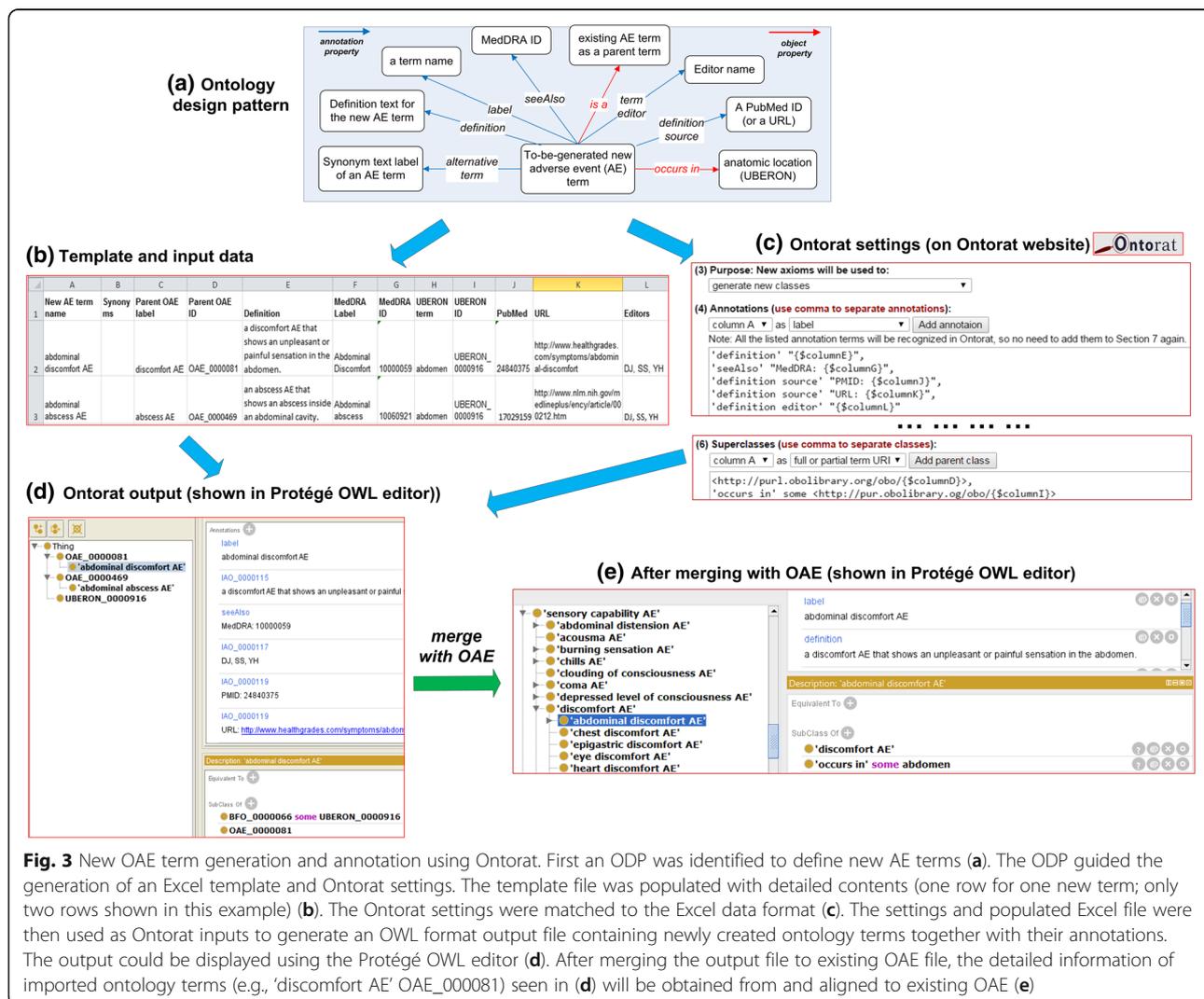
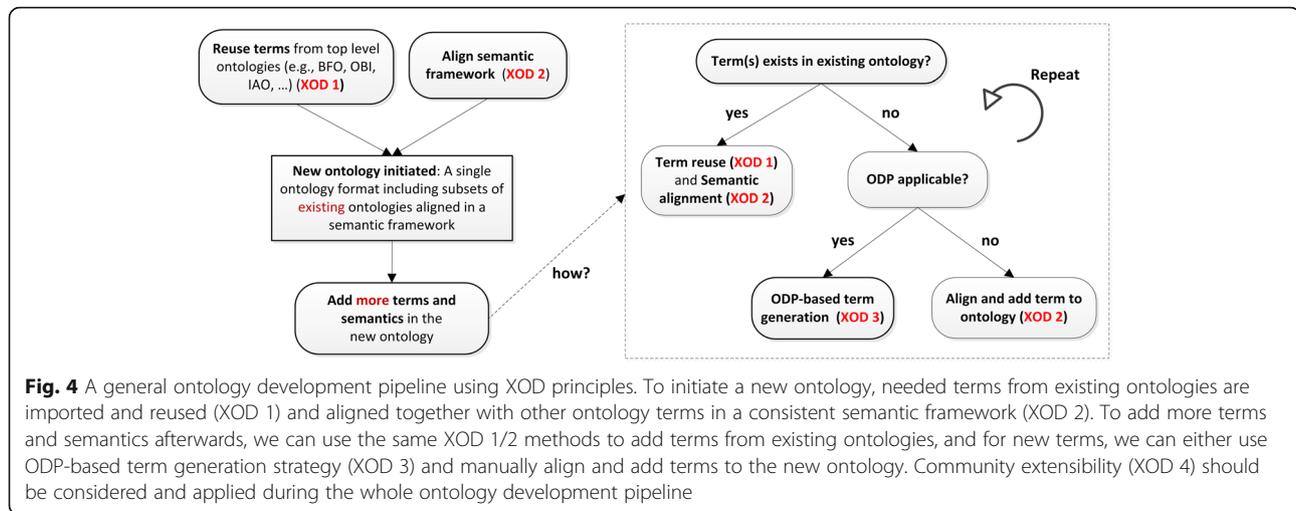


Fig. 3 New OAE term generation and annotation using Ontorat. First an ODP was identified to define new AE terms (a). The ODP guided the generation of an Excel template and Ontorat settings. The template file was populated with detailed contents (one row for one new term; only two rows shown in this example) (b). The Ontorat settings were matched to the Excel data format (c). The settings and populated Excel file were then used as Ontorat inputs to generate an OWL format output file containing newly created ontology terms together with their annotations. The output could be displayed using the Protégé OWL editor (d). After merging the output file to existing OAE file, the detailed information of imported ontology terms (e.g., ‘discomfort AE’ OAE_000081) seen in (d) will be obtained from and aligned to existing OAE (e)



Other ODP-based XOD tools include MappingMaster [42], TermGenie [43], Populous [44], Webulous [45], and ROBOT [36] (Table 1). Developed as a Protégé plugin, MappingMaster can only be used with old version Protégé 3.4 and is not available for newer Protégé 4 and 5 [42]. Targeting domain experts, TermGenie provides a web application that supports new GO term generation based on predefined patterns [43]. Populous requires software installation but provides a user-friendly interface [44]. Webulous is Google Add-On application usable with Google Spreadsheets [45]. ROBOT also has a template system for converting spreadsheets of terms to OWL files.

XOD 4: Community extensibility

The community's involvement during the developing phase of an ontology is the key for wide adoption of the ontology in the future. However, this step is often a bottleneck for ontology development, since the wider the community is, the more difficult it is to reach agreements on term definitions and classifications. In the reality, an ontology is often initiated by a small group and often driven by one or more use cases. To enhance its quality and broad recognition, XOD 4 recommends that a broader community with more developers and users participate in the ontology development and applications. This XOD community extensibility principle emphasizes the community participation to further extend, develop, and apply an ontology. With the principle of community extensibility, one ontology can be extended to cover different use cases in the same project and different projects from a wide range of research communities. The nature of such a practice will require more people to participate, make the ontology community bigger, and achieve better data interoperability.

WebProtege [46] and Ontokiwi [24] support community-based ontology development. WebProtege is a web-based ontology editor that supports collaborative OWL ontology

development [46]. WebProtege has been used by many groups. It includes full change tracking and revision history, and many community collaboration features such as sharing and permissions, threaded notes and discussions, watches and email notifications. Ontokiwi is the user-friendly Wiki-like web program that supports community-wide ontology editing, annotation, discussion, and distribution [24]. Ontobedia is an Ontokiwi application preloaded with existing biomedical ontologies [24]. The Wiki-like addition and editing of text that is not part of the ontology makes Ontokiwi/Ontobedia a unique platform for community-wide ontology discussion and distribution.

For community-wide ontology development and applications, tools to support ontology query, comparison, and evaluations are also needed. NCBO BioPortal [2], OLS [47], Ontobee [21], and AberOWL [48] are commonly used ontology registry and repositories that also provide ontology visualization, queries and analysis features, which facilitates the community involvement principle. The SPOT ontology toolkit (<http://www.ebi.ac.uk/spot/ontology/>) also provides a list of community-driven open source ontology tools. For example, Ontobee (<http://www.ontobee.org>) is an ontology browser and a linked ontology data server for dereferencing ontology terms [21]. Ontobeeep is an ontology comparison program that compares ontologies and identifies common terms existing in two or three ontologies by aligning 2–3 ontologies from the roots of these ontologies [22]. Ontobeeep also detects inconsistency and term duplication in one or more ontologies.

Demonstrations of XOD implementation for interoperable ontology development

Figure 4 outlines a simple pipeline of how the XOD principles can be used together for productive ontology development. Basically, a new ontology can be initiated by reusing existing terms from different ontologies (XOD 1) and aligning these terms in a semantic

framework (XOD 2), and new terms can be added by extending the semantic framework (XOD 2) and if ODPs identifiable, applying ODP-based approach (XOD 3). Ontology development often uses top-down and bottom-up approaches simultaneously [49]. The ontology initiation step is usually achieved by the top-down approach, i.e., developing the top level semantic framework by reusing and aligning upper level terms and semantics from existing ontologies (XOD 1/2). The same top-down approach can also be used to generate the upper level new terms commonly identified in the new ontology. Meanwhile, the bottom-up approach is use case driven and focuses on adding new terms to address specific use cases. For the bottom-up approach, XOD 1–3 principles are all important, and if possible, ODP-based design and term generation (XOD 3) is often critical to ensure development efficiency and consistency.

Here we will demonstrate our pipeline by using the complete use case of developing the community-based Vaccine Ontology (VO) [10, 50, 51]. As outlined in the extensive ontology development pipeline (Fig. 4), at the early stage of the VO development, we performed ontology survey and reused terms from several existing ontologies including BFO [33], OBI [7], GO [8], and the Information Artifact Ontology (IAO) [52]. The original VO version reported in 2009 included ~1000 imported terms from 10 existing ontologies and ~1000 VO-specific terms [50]. Since then more terms have been added to VO. As of November 20, 2017, out of 6541 terms in VO, approximately 1600 terms were imported and reused from approximately 30 ontologies (<http://www.ontobee.org/ontostat/VO>).

Many VO-specific terms were added to VO by semantic alignment with the upper BFO ontology or middle level ontologies (e.g., OBI) (XOD 2). For example, VO term ‘vaccine’ (VO_0000001) is asserted as a subclass of OBI term ‘processed material’ (OBI_0000047). This assertion means that any non-processed material (e.g., an infectious bacterium that exists in the air) that causes an infection in human and eventual immune responses and protection in the human is not counted as a vaccine. Similarly, the VO term ‘vaccination’ (VO_0000002) is asserted as a subclass of OBI term ‘administering substance in vivo’ (OBI_0600007). The alignment with ‘administering substance in vivo’ differentiates VO vaccination (i.e., administering a vaccine to in vivo) from immunization (i.e., to make one immune to something). In comparison, vaccination is considered as the synonym of immunization in MedDRA, a controlled terminology system commonly used for representation of regulatory activities [53].

In many cases, we can generate a number of new terms simultaneously by developing and following specific ODPs (XOD 3). For example, the VO developers

retrieved from the US Department of Agriculture (USDA) and other public databases the information of approximately 800 licensed animal vaccines. Manually adding these animal vaccines to VO would be time consuming. To speed up the inclusion of the large number of licensed animal vaccines to VO, an ODP was developed to include different entities (e.g., vaccine name, manufacturer, animal species, animal pathogen, and disease), annotations, and the semantic relations among these entities. Such an ODP was further used to design an Excel template which was then applied to include the categorized information of these vaccines. Ontorat was finally used to automatically transfer the ODP and the information recorded in the Excel file to an OWL file and then imported to VO [20]. Furthermore, the same ODP could be used later to add new animal vaccines to VO. The Ontorat use case of ODP-based VO addition of veterinary vaccines was first presented in the 2012 International Conference for Biomedical Ontology (ICBO) [54]. Since then other ODPs were also developed for further VO development [51]. Meanwhile, it is noted that not all new terms can be fit under identifiable common design patterns. In this case, we can generate the term by aligning it with existing framework (XOD 2) (Fig. 4).

As a community-based open source ontology, the VO development has involved the broader community in its continuous development (XOD 4). The community participation helps further extend the VO and its interoperability with other biomedical ontologies. For example, according to BioPortal and Ontobee, the VO term vaccine (VO_0000001) has been reused by more than ten other ontologies such as OBI and Apollo Structured Vocabulary (<https://github.com/apollodev/>), and the VO term vaccination (VO_0000002) has been reused by ten other ontologies such as the Prescription of Drugs Ontology (<https://github.com/OpenLHS/PDRO>). In addition, the VO community involvement makes it achieve better data interoperability with other ontologies. Meanwhile, the community involvement extends the applications of VO, such as vaccine-related T cell and B cell response analysis and queries [55], epitope data management [56], vaccine-related literature mining [10], and vaccine-related network analysis [57, 58].

In addition to VO, many other ontologies, e.g., Beta Cell Genomics Ontology (BCGO) [37], Micro ontology for representing microorganism phenotypic and metabolic characters [59], and BioAssay Ontology (BAO) [60], have been developed using the same or similar strategies.

Discussion and perspectives

The XOD strategy and principles reflect the growing maturity of biological and biomedical ontology development. When only a small number of ontologies were developed, such XOD strategy was not needed. However, with

hundreds of ontologies developed now, it is critical to ensure ontology interoperability, and the XOD principles provide a practical solution. Given the importance of ontologies in the integration, sharing, and analysis of the increasing large and heterogeneous data/metadata and knowledge, the XOD strategy is very significant and critical to meet the challenges in the current big data era.

Among the four XOD principles, the first three principles emphasize the requirements to reuse ontology terms, extend and align semantic structures, and build new terms and semantics among terms using design patterns. XOD 2 is a more general principle which covers the semantic interoperability among terms including terms from the target ontology and terms newly generated or imported from source ontology. Extending XOD 1 and 2, XOD 3 provides a more specific mechanism (i.e., ODP-based term generation and editing) to achieve consistent ontology term generation and annotation. While the first three principles provide more technical guidance, XOD 4 emphasizes the community collaboration and involvement in new ontology development.

XOD is complementary to the OBO principles [1] and the ten simple rules proposed for ontology development [26]. The OBO principles (e.g., open, common format, versioning, scope, relations, users, collaboration, and locus of authority) provide general principles for the development of an ontology (<http://obofoundry.org/principles/fp-000-summary.html>). The ten simple rules proposed by Malone et al. include ontology term reuse, design patterns, and community engagement [26], which are directly associated with XOD principles. The other 7 rules (e.g., scope, license, versioning) are not directly related. In comparison to the OBO principles and the ten simple rules, the XOD principles address the single important point of ontology extensibility and emphasize different scales of extensible relations among ontologies, with the aim to achieve ontology interoperability. Since different ontologies extend and are aligned with existing reliable ontologies, applying the XOD principles will support the OBO aim of establishing non-redundant and interoperable suite of ontologies.

The XOD strategy supports the FAIR Guiding Principles, which propose that various data be Findable, Accessible, Interoperable and Reusable [27]. Ontologies lay out the basic foundation for the data FAIRness. Adopting the XOD strategy will lead to the development of extensible ontologies and the generation of ontology-extended data and metadata representations. Such ontology-supported data sharing and integration will result in natural data access, interoperability, and usability, query, and advanced analysis. For example, the KaBOB knowledge base uses the OBO ontologies to semantically integrate data from 18 prominent biomedical databases [61]. Millions of RDF triples were also generated in KaBOB, enabling findable, accessible,

interoperable, and reusable queries of the underlying data from these databases. Therefore, the XOD strategy supports the eventual achievement of the FAIRness of data.

Many challenges exist in adopting and achieving the goals defined in the XOD strategy and principles. First, the interoperability among current hundreds of ontologies is still limited and challenging [32, 62, 63]. Term redundancy among ontologies cannot be solved easily, leading to issues of achieving data FAIRness. Second, only a small amount of data resources (including a large number of databases) adopt ontology-guided strategy, which restricts data interoperability and analysis. Third, while many linked data systems [25] standardize data using ontologies, the ontologies underlying linked data are often non-interoperable, making linked data systems become individual silos and difficult to integrate [64, 65]. To address these challenges, it is important to adopt the XOD strategy and XOD principles. Active ontology training and outreach will be beneficial.

The suite of Ontoanimal tools has provided different features to address several real issues in ontology development. Each of these tools focuses on one or more primary tasks, and all together they are combined to strongly support XOD principles. Given the complexity of these tools, there are concerns about their usability and sustainability. Since these tools are more about research in ontology development, it is important to have a strong evaluation system to be used to better understand the strengths and limitation of each tool.

While Ontoanimal tools and other similar XOD tools have already supported the XOD strategy, existing tools require further improvements, and more user-friendly integrative tools are needed. Tools are critical to make more efficient extensible ontology development. For example, although it was recognized that term reuse was a better strategy, the term reuse principle was not widely implemented until Ontofox and other tools were developed. Currently XOD tool usage often requires extensive training. More easy-to-use and integrative XOD tools are desired for ontology developers and users with no or limited programming background. We believe that the adoption of the XOD principles together with robust XOD tools would greatly support interoperable ontology development and data FAIRness.

Conclusion

Our examination of Ontoanimal tools and similar programs discovered their shared features of extensibility. We proposed the “eXtensible Ontology Development” (XOD) strategy and four XOD principles to support extensible ontology development and usage. We propose to adopt these XOD principles for active development and usage of extensible ontologies and tools, leading to better data FAIRness.

Abbreviations

BFO: Basic formal ontology; GO: Gene ontology; MIREOT: Minimum Information to reference an external ontology term; NCBO: National Center for Biomedical Ontology; OAE: Ontology of adverse events; OBO: Open biological and biomedical ontologies; ODP: Ontology design pattern; OWL: Web ontology language; RDF: Resource description framework; XML: eXtensible markup language; XOD: eXtensible ontology development

Acknowledgements

We acknowledge the OBO Foundry consortium and OBO ontology developers for their support and collaborations.

Funding

The Ontoanimal tool research was partly supported by the NIH National Institute of Allergy and Infectious Diseases grant 1R01AI081062.

Availability of data and materials

Not applicable.

Authors' contributions

YH Ontoanimal tool project design and management, initial XOD concept proposer, preparation of the first manuscript draft; ZX Primary developers of Ontoanimal tools Ontofox, early version of Ontobee, Ontobeeep, Ontodog, Ontorat, and Ontobat; JZ Co-developer of Ontodog, Ontorat, and Ontobee; YL Ontoanimal tool testing, discussion, and evaluation; JO ROBOT tool developer, Ontoanimal tool testing and evaluation; EQ Primary developers of Ontokiwii/Ontobedia, later version of Ontobee, Ontobull, and Ontoanimal tool maintenance. All authors participated in discussion, manuscript preparation and editing, and approved the manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Unit for Laboratory Animal Medicine, Department of Microbiology and Immunology, Center for Computational Medicine and Bioinformatics, University of Michigan Medical School, Ann Arbor, MI, USA. ²Department of Genetics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA. ³Center for Computational Science, University of Miami, Coral Gables, FL, USA. ⁴Nocean Inc., Toronto, ON, Canada. ⁵Department of Computational Medicine and Bioinformatics, University of Michigan Medical School, Ann Arbor, MI, USA.

Received: 1 May 2017 Accepted: 7 December 2017

Published online: 12 January 2018

References

- Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, Goldberg LJ, Eilbeck K, Ireland A, Mungall CJ, et al. The OBO foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat Biotechnol.* 2007;25(11):1251–5.
- Whetzel PL, Noy NF, Shah NH, Alexander PR, Nyulas C, Tudorache T, Musen MA. BioPortal: enhanced functionality via new web services from the National Center for biomedical ontology to access and use ontologies in software applications. *Nucleic Acids Res.* 2011;39(Web Server issue):W541–5.
- Schulz S, Balkanyi L, Cornet R, Bodenreider O. From concept representations to Ontologies: a paradigm shift in health informatics? *Healthc Inform Res.* 2013;19(4):235–42.
- Hoehndorf R, Schofield PN, Koutos GV. The role of ontologies in biological and biomedical research: a functional perspective. *Brief Bioinform.* 2015; 16(6):1069–80.
- Marcos E, Zhao B, He Y. The ontology of vaccine adverse events (OVAE) and its usage in representing and analyzing adverse events associated with US-licensed human vaccines. *J Biomed Semantics.* 2013;4(1):40.
- Demir E, Cary MP, Paley S, Fukuda K, Lemer C, Vastrik I, Wu G, D'Eustachio P, Schaefer C, Luciano J, et al. The BioPAX community standard for pathway data sharing. *Nat Biotechnol.* 2010;28(9):935–42.
- Bandrowski A, Brinkman R, Brochhausen M, Brush MH, Bug B, Chibucos MC, Clancy K, Courtot M, Derom D, Dumontier M, et al. The ontology for biomedical investigations. *PLoS One.* 2016;11(4):e0154556.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat Genet.* 2000;25(1):25–9.
- Doms A, Schroeder M. GoPubMed: exploring PubMed with the gene ontology. *Nucleic Acids Res.* 2005;33(Web Server issue):W783–6.
- Ozgun A, Xiang Z, Radev DR, He Y. Mining of vaccine-associated IFN-gamma gene interaction networks using the vaccine ontology. *J Biomed Semantics.* 2011;2(Suppl 2):S8.
- Gonzalez-Beltran A, Maguire E, Sansone SA, Rocca-Serra P. linkedISA: semantic representation of ISA-tab experimental metadata. *BMC Bioinformatics.* 2014;15(Suppl 14):S4.
- Malladi VS, Erickson DT, Podduturi NR, Rowe LD, Chan ET, Davidson JM, Hitz BC, Ho M, Lee BT, Miyasato S, et al. Ontology application and use at the ENCODE DCC. *Database.* 2015;2015:1–11. <https://www.ncbi.nlm.nih.gov/pubmed/?term=25776021>.
- Dugan VG, Erlich SJ, Giraldo-Calderon GI, Harb OS, Newman RM, Pickett BE, Schriml LM, Stockwell TB, Stoeckert CJ Jr, Sullivan DE, et al. Standardized metadata for human pathogen/vector genomic sequences. *PLoS One.* 2014;9(6):e99979.
- Guo A, Racz R, Hur J, Lin Y, Xiang Z, Zhao L, Rinder J, Jiang G, Zhu Q, He Y. Ontology-based collection, representation and analysis of drug-associated neuropathy adverse events. *J Biomed Semantics.* 2016;7:29.
- He Y. Ontology-based vaccine and drug adverse event representation and theory-guided systematic causal network analysis toward integrative pharmacovigilance research. *Curr Pharmacol Rep.* 2016;2(3): 113–28.
- Zheng J, Stoyanovich J, Manduchi E, Liu J, Stoeckert CJ Jr. AnnotCompute: annotation-based exploration and meta-analysis of genomics experiments. *Database.* 2011;2011:bar045.
- Rubin DL, Noy NF, Musen MA. Protege: a tool for managing and using terminology in radiology applications. *J Digit Imaging.* 2007;20(Suppl 1):34–46.
- Xiang Z, Courtot M, Brinkman RR, Ruttenberg A, He Y. OntoFox: web-based support for ontology reuse. *BMC Res Notes.* 2010;3(175):1–12.
- Zheng J, Xiang Z, Stoeckert CJ Jr, He Y. Ontodog: a web-based ontology community view generation tool. *Bioinformatics.* 2014;30(9):1340–2.
- Xiang Z, Zheng J, Lin Y, He Y. Ontorat: automatic generation of new ontology terms, an-notations, and axioms based on ontology design patterns. *J Biomed Semantics.* 2015;6(1):4. (10 pages)
- Ong E, Xiang Z, Zhao B, Liu Y, Lin Y, Zheng J, Mungall C, Courtot M, Ruttenberg A, He Y. Ontobee: a linked ontology data server to support ontology term dereferencing, linkage, query and integration. *Nucleic Acids Res.* 2017;45(D1):D347–52.
- Xiang Z, He Y. IDO extensions alignment using Ontobee. In: *Infectious Disease Ontology (IDO) Workshop 2010: December 8-9, vol. 2010.* Baltimore, MD, USA.
- Ong E, Xiang Z, Zheng J, Smith B, He Y. Ontobull and BFOConvert: web-based programs to support automatic ontology conversion. In: *Proceedings of the Joint International Conference on Biological Ontology and BioCreative: August 1-2, 2016 2016.* Corvallis, Oregon, USA: CEUR Workshop proceedings. p. 1–2. URL: http://ceur-ws.org/Vol-1747/D201_ICBO2016.pdf.
- Ong E, He Y. Community-based Ontology Development, Annotation and Discussion with MediaWiki extension Ontokiwii and Ontokiwii-based Ontobedia. *AMIA Jt Summits Transl Sci Proc.* 2016;2016:65–74.
- Bizer C, Heath T, Berners-Lee T. Linked data - the story so far. *Int J Semantic Web Inform Syst.* 2009;5(3):1–22.
- Courtot M, Malone J, Mungall C: Ten simple rules for biomedical ontology development. In: *International Conference on Biomedical Ontology and BioCreative (ICBO BioCreative 2016): 2016; Corvallis, Oregon, United States, 2016.* CEUR-ws.org: http://ceur-ws.org/Vol-1747/IT404_ICBO2016.pdf.
- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, et al. The FAIR guiding principles for scientific data management and stewardship. *Sci Data.* 2016;3:160018.

28. Malone J, Stevens R, Jupp S, Hancocks T, Parkinson H, Brooksbank C. Ten simple rules for selecting a bio-ontology. *PLoS Comput Biol*. 2016; 12(2):e1004743.
29. Hoehndorf R, Dumontier M, Gkoutos GV. Evaluation of research in biomedical ontologies. *Brief Bioinform*. 2013;14(6):696–712.
30. Courtot M, Gibson F, Lister A, Malone J, Schober D, Brinkman R, Ruttenberg A. MIREOT: the minimum information to reference an external ontology term. *Appl Ontol*. 2011;6:23–33.
31. Kalfoglou Y, Schorlemmer M. Ontology mapping: the state of the art. *Knowl Eng Rev*. 2003;18(1):1–31.
32. Tang J, Liang B-Y, Li J-Z. Toward detecting mapping strategies for ontology interoperability. In: *The 14th International World Wide Web Conference (WWW 2005)*. Japan; 2005. <https://pdfs.semanticscholar.org/e40f/885e4c9397a538513952d8ab0cabc1e15eba.pdf>.
33. Arp R, Smith B, Spear AD. *Building Ontologies using basic formal ontology*. Cambridge, MA, USA: MIT Press; 2015.
34. Hanna J, Cheng C, Crow A, Hall R, Liu J, Pendurthi T, Schmidt T, Jennings S, Brochhausen M, Hogan W. Simplifying MIREOT; a MIREOT Protege Plugin. In: *International Semantic Web Conference (Posters & Demos): November 11–15, 2012*. Boston, USA; 2012. p. 1–4. URL: http://ceur-ws.org/Vol-914/paper_48.pdf.
35. Maguire E, Gonzalez-Beltran A, Whetzel PL, Sansone SA, Rocca-Serra P. OntoMaton: a biportal powered ontology widget for Google spreadsheets. *Bioinformatics*. 2013;29(4):525–7.
36. Overton JA, Dietze H, Essaid S, Osumi-Sutherland D, Mungall CJ. ROBOT: A command-line tool for ontology development. *Proceedings of the 2015 International Conference on Biomedical Ontologies (ICBO)*, Lisbon, Portugal, July 27–30, 2015. 2015, 1515:1–2. URL: <http://ceur-ws.org/Vol-1515/demo6.pdf>.
37. Zheng J, Manduchi E, Stoeckert CJ Jr. Development of an Application Ontology for Beta Cell Genomics based on the Ontology for Biomedical Investigations. *4th Int Confer Biomed Ontol (ICBO-2013)*. 2013;1060:62–7.
38. Aranguren ME, Antezana E, Kuiper M, Stevens R. Ontology Design Patterns for bio-ontologies: a case study on the Cell Cycle Ontology. *BMC Bioinformatics*. 2008;9(Suppl 5):S1.
39. Hoehndorf R, Ngonga Ngomo AC, Pyysalo S, Ohta T, Oellrich A, Rebholz-Schuhmann D. Ontology design patterns to disambiguate relations between genes and gene products in GENIA. *J Biomed Semantics*. 2011; 2(Suppl 5):S1.
40. He Y, Sarntivijai S, Lin Y, Xiang Z, Guo A, Zhang S, Jagannathan D, Toldo L, Tao C, Smith B. OAE: the ontology of adverse events. *J Biomed Semantics*. 2014;5:29.
41. Rocca-Serra P, Ruttenberg A, O'Connor MJ, Whetzel PL, Schober D, Greenbaum J, Courtot M, Brinkman R, Sansone SA, Scheuerman R, et al. Overcoming the ontology enrichment bottleneck with quick term templates. *Appl Ontol*. 2011;6(1):13–22.
42. O'Connor M, Halaschek-Wiener C, Musen M. M2: a language for mapping spreadsheets to OWL. San Francisco, CA, USA: OWLED 2010; 2010.
43. Dietze H, Berardini TZ, Foulger RE, Hill DP, Lomax J, Osumi-Sutherland D, Roncaglia P, Mungall CJ. TermGenie - a web-application for pattern-based ontology class generation. *J Biomed Semantics*. 2014;5:48.
44. Jupp S, Horridge M, Iannone L, Klein J, Owen S, Schanstra J, Wolstencroft K, Stevens R. Populous: a tool for building OWL ontologies from templates. *BMC Bioinformatics*. 2012;13(Suppl 1):S5.
45. Jupp S, Burdett T, Welter D, Sarntivijai S, Parkinson H, Malone J. Webulous and the Webulous Google add-on—a web service and application for ontology building from templates. *J Biomed Semantics*. 2016;7:17.
46. Horridge M, Tudorache T, Nuyals C, Vendetti J, Noy NF, Musen MA. WebProtege: a collaborative web-based platform for editing biomedical ontologies. *Bioinformatics*. 2014;30(16):2384–5.
47. Cote R, Reisinger F, Martens L, Barsnes H, Vizcaino JA, Hermjakob H. The Ontology Lookup Service: bigger and better. *Nucleic Acids Res*. 2010; 38(Web Server issue):W155–60.
48. Slater L, Gkoutos GV, Schofield PN, Hoehndorf R. Using AberOWL for fast and scalable reasoning over BioPortal ontologies. *J Biomed Semantics*. 2016;7(1):49.
49. Zheng J, Harris MR, Masci AM, Lin Y, Hero A, Smith B, He Y. The ontology of biological and clinical statistics (OBCS) for standardized and reproducible statistical analysis. *J Biomed Semantics*. 2016;7(1):53.
50. He Y, Cowell L, Diehl AD, Mobley HL, Peters B, Ruttenberg A, Scheuermann RH, Brinkman RR, Courtot M, Mungall C, et al. VO: Vaccine Ontology. In: *The 1st International Conference on Biomedical Ontology (ICBO-2009): July 24–26 2009*; Buffalo, NY, USA: Nature Precedings: <http://precedings.nature.com/documents/3552/version/1>; 2009.
51. Lin Y, He Y. Ontology representation and analysis of vaccine formulation and administration and their effects on vaccine immune responses. *J Biomed Semantics*. 2012;3(1):17.
52. The Information Artifact Ontology (IAO) [<https://github.com/information-artifact-ontology/IAO/>].
53. Brown EG, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). *Drug Saf*. 1999;20(2):109–17.
54. Xiang Z, Lin Y, He Y. Ontorat web server for automatic generation and annotations of new ontology terms. In: *The 3rd International Conference on Biomedical Ontologies (ICBO): July 24–27 2012*; Graz, Graz, Austria: CEUR Workshop Proceedings; 2012. URL: http://ceur-ws.org/Vol-897/poster_12.pdf.
55. Vita R, Overton JA, Greenbaum JA, Sette A, consortium OBI, Peters B. Query enhancement through the practical application of ontology: the IEDB and OBI. *J Biomed Semantics*. 2013;4(Suppl 1):S6.
56. Chung JL, Sun J, Sidney J, Sette A, Peters B. IMMUNOCAT—a data management system for epitope mapping studies. *J Biomed Biotechnol*. 2010;2010:856842.
57. Tao C, Wu P, Luo Y, Zhang Y. Linked vaccine adverse event data from VAERS for biomedical data analysis and longitudinal studies. *BioData Mining*. 2014;7:36.
58. Zhang Y, Tao C, He Y, Kanjamala P, Liu H. Network-based analysis of vaccine-related associations reveals consistent knowledge with the vaccine ontology. *J Biomed Semantics*. 2013;4(1):33.
59. Blank CE, Cui H, Moore LR, Walls RL. Micro: an ontology of phenotypic and metabolic characters, assays, and culture media found in prokaryotic taxonomic descriptions. *J Biomed Semantics*. 2016;7:18.
60. Abeyruwan S, Vempati UD, Kucuk-McGinty H, Visser U, Koleti A, Mir A, Sakurai K, Chung C, Bittker JA, Clemons PA, et al. Evolving BioAssay Ontology (BAO): modularization, integration and applications. *J Biomed Semantics*. 2014;5(Suppl 1 Proceedings of the Bio-Ontologies Spec Interest G):S5.
61. Livingston KM, Bada M, Baumgartner WA Jr, Hunter LE. KaBOB: ontology-based semantic integration of biomedical databases. *BMC Bioinformatics*. 2015;16:126.
62. Hyvönen E. Preventing ontology interoperability problems instead of solving them. *Semantic Web*. 2010;1(1, 2):33–7.
63. Euzenat J, Meilicke C, Stuckenschmidt H, Shvaiko P, Trojahn C. Ontology alignment evaluation initiative: six years of experience. In: *Journal on data semantics XV*; Springer; 2011. p. 158–92. https://link.springer.com/chapter/10.1007%2F978-3-642-22630-4_6.
64. Ding L, Michaelis J, McCusker J, McGuinness DL. Linked provenance data: a semantic web-based approach to interoperable workflow traces. *Futur Gener Comput Syst*. 2011;27(6):797–805.
65. Bernstein A, Hendler J, Noy N. A new look at the semantic web. *Commun ACM*. 2016;59(9):35–7.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

