



## Ontologies in Biological Data Visualization

**Sheelagh Cpendale**  
*University of Calgary*

**Min Chen**  
*University of Oxford*

**Daniel Evanko**  
*Nature Publishing Group*

**Nils Gehlenborg**  
*Harvard Medical School*

**Carsten Görg and Larry Hunter**  
*University of Colorado*

**Francis Rowland**  
*European Bioinformatics Institute*

**Margaret-Anne Storey**  
*University of Victoria*

**Hendrik Strobel**  
*Polytechnic Institute of New York University*

In computer science, an ontology is essentially a graph-based knowledge representation in which each node corresponds to a concept and each edge specifies a relation between two concepts. Concepts can encompass a variety of real-world entities and abstractions such as names, objects, events, ideas, and types. Relations can encompass different unidirectional or bidirectional associations, including connectivity, hierarchy, membership, functional mapping, and causation. Ontologies often feature additional information such as node and edge attributes; formal-logic operators; structural groupings for defining sets, rules, and systems; and temporal changes of attributes, topologies, and structures.

As computational models, ontologies have been used extensively in artificial intelligence, natural-language processing, and Web sciences. A number of specification languages (for example, OWL, the Web Ontology Language) have been developed for encoding ontologies. (For the URL for OWL and other resources mentioned in the article, see the “Related Resources” sidebar.) Among the many disciplines in which visualization plays a significant role, biology has embraced ontologies purposefully and energetically.

Here, we examine the potential impact of ontologies’ increasing presence in research on biological data visualization, considering both technical challenges and opportunities for new research directions. We have two ultimate goals:

- Support ontology developers, data curators, and data analysts with advanced visualization techniques.
- Stimulate new research programs for developing knowledge-assisted visualization<sup>1</sup> by using ontologies as knowledge representations.

Because ontological development has progressed much further in biology than in many other disciplines, biological data visualization offers an exceptional application field for ontology-based visualization research.

### Ontologies in Biology

Biology is notoriously nonquantitative. In contrast to the law-like equations of physics, scientific claims in biology (especially molecular biology) often hypothesize the existence of specific entities, functions, and relationships among them. Furthermore, the deep evolutionary interconnectedness of all living things means that hypotheses and evidence regarding them in one area of biology are often relevant to hypotheses about apparently quite distant aspects of biology. This problem is particularly acute now that we have the technology to easily and cheaply determine the chemical sequence of a genome (which contains or encodes most of the central biological entities), without the complementary ability to help people understand these complex, multifaceted functions and relationships.

## Related Resources

Here are descriptions and URLs for some of the resources mentioned in the main article, along with a couple of other useful resources.

### Ontologies in Biology

- *The Gene Ontology* is a collection of ontological data and tools for accessing and processing the data. <http://geneontology.org>
- *The National Center for Biomedical Ontology* is a community hub that gathers a collection of open source software for biomedical ontologies. <http://bioontology.org>
- *The Neuroscience Lexicon* is a wiki for community-based curation of neuroscience terms, with collected information about gene ontology tools. [http://neurolex.org/wiki/Category:Resource:Gene\\_Ontology\\_Tools](http://neurolex.org/wiki/Category:Resource:Gene_Ontology_Tools)
- *The OBO Foundry* is a collaborative experiment for developers of open biological and biomedical ontologies to facilitate interoperability between diverse ontologies. [www.obofoundry.org](http://www.obofoundry.org)

### Ontology Editing and Visualization

- *BioMixer* is a software tool for visualizing multiple biological ontologies and the mappings across them. <http://thechiselgroup.org/2012/01/01/biomixer-visualizing-mappings-of-biomedical-ontologies-2>
- *Protégé* is an ontology editor. <http://protege.stanford.edu>
- *Jambalaya* is a Protégé plug-in for visualizing ontologies. <http://thechiselgroup.org/2004/07/06/jambalaya>

- *OntoViz* is a Protégé plug-in for visualizing ontologies. <http://protegewiki.stanford.edu/wiki/OntoViz>
- *TGVizTab* is a Protégé plug-in for visualizing ontologies. <http://users.ecs.soton.ac.uk/ha/TGVizTab>

### Ontology-Supported Exploration

- *BioJigSaw* is a visual index that supports the investigation of connections between biological entities. [www.cc.gatech.edu/~john.stasko/papers/kes10-biojigsaw.pdf](http://www.cc.gatech.edu/~john.stasko/papers/kes10-biojigsaw.pdf)
- *The Bohemian Bookshelf* is an interactive tool for exploring digital book collections. <http://innovis.cpsc.ucalgary.ca/Research/BohemianBookshelf>
- *Docuburst* is an ontology-based document visualization tool. <http://vialab.science.uoit.ca/portfolio/docuburst-visualizing-document-content-using-language-structure>
- *GoPubMed* is an ontology-based search engine for biomedical-research publications. [www.gopubmed.org/web/gopubmed](http://www.gopubmed.org/web/gopubmed)

### Other Resources

- *OWL* is the Web Ontology Language. [www.w3.org/TR/owl2-overview](http://www.w3.org/TR/owl2-overview)
- *WordNet* is a large lexical database of English. <http://wordnet.princeton.edu>
- *TimeViz* is a survey of time-oriented visualization techniques. <http://survey.timeviz.net>
- *MatrixBrowser* is network visualization software. <http://dl.acm.org/citation.cfm?id=506504>

To make this problem more concrete, consider a long-standing, fundamental hypothesis: homologous genes in different organisms have conserved (that is, identical or very similar) functions. Homologous genes arose from a common ancestor; genes with very similar sequences are likely homologous. Genomic sequencing makes determining molecular homology straightforward. There's also an enormous amount of scientific literature that (among other things) makes claims about the functions of millions of different genes in thousands of organisms. The combination of these factors should have enabled the testing of the conserved-function hypothesis.

Unfortunately for biologists, statements about gene functions are in hundreds of thousands of long, complex journal articles written in human language. Although the data needed to test this hypothesis is clearly in the literature, the difficulty of extracting all that information and putting it in a form that facilitates testing makes the task utterly impractical.

In the late 1990's, Michael Asburner, a fruit fly geneticist at Cambridge and the first director of the

European Bioinformatics Institute, began advocating for a standardized language to describe critical biological concepts—for example, gene functions. Adopting both the philosophical principles of ontology and the computational approach that Thomas Gruber proposed,<sup>2</sup> biologists working on different model organisms (fruit flies, mice, and yeast) began coordinating their functional vocabulary to create an ontology of biological processes, molecular functions, and subcellular locations. In 2000, they published the first version of the Gene Ontology.<sup>3</sup> This resource contained not only the ontology but also mappings of genes from different organisms to the newly standardized way of describing qualitative aspects of biology. The computational approach to ontologies can exploit this standardization, additionally including subsumption hierarchies, mereologies, and other types of relationships among terms. This will provide value beyond just a standardized vocabulary.

Since then, ontologies have covered an increasing number of biological areas (for example, chemical entities, sequence phenomena, and anatomy), and the number of genes annotated with

Several lines of evidence, based on the results of overexpression studies, indicate that PGC-1 $\alpha$  is sufficient to promote mitochondrial biogenesis and regulate mitochondrial respiratory capacity. First, PGC-1 $\alpha$  activates the transcription of mitochondrial uncoupling protein-1 (UCP-1) in BAT through interactions with the nuclear hormone receptors PPAR $\gamma$  and thyroid receptor [2]. Second, forced expression studies in adipogenic and myogenic mammalian cell lines demonstrated that PGC-1 $\alpha$  activates mitochondrial biogenesis through a group of transcription factor targets including nuclear respiratory factors 1 and 2 (NRF-1 and -2) and mitochondrial transcription factor A (Tfam), key transcriptional regulators of mitochondrial DNA transcription and replication [8]. Third, studies in primary cardiac myocytes in culture and in the hearts of transgenic mice have demonstrated that overexpression of PGC-1 $\alpha$  promotes mitochondrial biogenesis [10,16]. Lastly, forced expression of PGC-1 $\alpha$  in skeletal muscle of transgenic mice triggers mitochondrial proliferation and the formation of mitochondrial-rich type I, oxidative ("slow-twitch") muscle fibers [17]. Collectively, these results indicate that PGC-1 $\alpha$  is sufficient to drive mitochondrial biogenesis.

Figure 1. A piece of biology text annotated with multiple ontologies. Different-color highlights represent different classes as defined by several ontologies. (Source: Colorado Richly Annotated Full Text Corpus; <http://bionlp-corpora.sourceforge.net/CRAFT>; used with permission.)

ontological terms has grown dramatically. The US National Institutes of Health, recognizing ontological research's importance to all biomedical research, created the National Center for Biomedical Ontology and funds the Gene Ontology Consortium, in addition to providing other wide-ranging ontology-related research support. The Gene Ontology's structure and the annotations of genes to terms in the ontology have unleashed a vast, tremendously useful torrent of valuable applications, many of which are far from Ashburner and his colleagues' original intent. For example, biomedical ontologies turn out to be an important target for biomedical text mining (see Figure 1).

The creation and dissemination of ontologies for entities in the published biomedical literature is only a first necessary step. An arguably larger challenge is accurately annotating scientific texts with these ontological terms. The same sequence homology that allows assignment and coordination of terms across organisms can contribute to authors' laxity in fully defining the source organism of named entities in manuscripts. And despite ontology developers' best efforts, researchers often continue to use deprecated or slang entity names in their writing. This situation complicates text mining of the published literature. Some publishers, such as Nature Publishing Group, try to correct these deficiencies during manuscript copyediting to ensure that all entity names are the current official ontological terms.

Even when a manuscript uses the correct terms for entities, each term still must be linked to the correct entity. Such linking can be done automati-

cally through text mining, using a tool such as Reflect, or the journals themselves can create these links. At select *Nature* research journals, technical editors perform such linking of manuscript text to ontologies of chemical compounds, genes, and proteins. By manually assigning terms and links during production, editors can query authors for accuracy and correct mistakes.

Instead of tagging individual entities, many publishers assign ontological terms to entire manuscripts. This ignores the fine-grained information in the manuscript but allows improved searching, browsing, and linking of manuscripts for strategic reading. Assigning the most relevant terms for any article is much more subjective than annotating individual entities. This can add considerable editorial value by

- filtering out articles that are only superficially related to a term in them or
- grouping articles on a subject regardless of whether the manuscript used that subject name.

On the other hand, this subjectivity can lead to undesirable inconsistencies in the ontologies themselves and in their use.

## Challenges

Many groups of people are working with ontologies on different kinds of tasks. For example, the OBO (open biological and biomedical ontologies) Foundry, a collaborative experiment for ontologists, lists over 120 proposed ontologies in different biological and biomedical areas.

Three main types of scientists work with ontologies. Ontologists focus on creating, maintaining, and extending ontologies. Data curators use ontologies to annotate and classify results from experiments, datasets, and the literature. Data analysts use ontologies indirectly (through computer applications) to analyze experimental data, search databases, and navigate and read the scientific literature. These multifaceted tasks create the following distinct challenges.

### Ontology Evolution

Ontologies are dynamic. Many ontologies constantly change and evolve owing to discoveries and new knowledge in their domain. The creation of multiple versions of ontologies is prone to inconsistencies and adds downstream complexity, posing problems for humans and computer programs. Keeping track of evolution becomes even more daunting for ontologies that integrate multiple data sources.

Ontology evolution affects all three types of scientists. Ontologists must keep track of ontologies' structure, study their properties (which parts are worked out in detail and which parts are preliminary or deprecated), and identify terms that aren't yet part of a given ontology. Merging, mapping, and comparing ontologies also become increasingly difficult when ontologies change over time.

Data curators must examine which parts of an ontology have changed and then re-annotate their data.

Data analysts must keep track of the version they used for all their analyses and potentially rerun their analyses after one or more ontologies have changed. In this context, a big problem they face is the automatic integration of data from multiple sources based on ontological annotations.

### **Scale**

Many ontologies represent an overwhelming amount of data. For example, the US National Center for Biomedical Ontology's BioPortal has collected over 300 ontologies, featuring more than 6 million terms and more than 26 billion annotations. The 2013-07-04 version of the popular Gene Ontology contains nearly 40,000 classes and some 645,000 gene products, with an 18-level hierarchy.

These large ontologies are usually developed and maintained by a team of ontologists; this requires a framework supporting collaborative work. A first step in this direction was the development of the Changes and Annotation Ontology.<sup>4</sup>

Data curators often have to deal with finding the most appropriate concepts in these large ontologies when they annotate terms in documents or samples in experimental data.

For data analysts, the number of ontology annotations can easily be overwhelming, especially if documents or data are annotated with multiple ontologies (as in Figure 1). Developers of programs that use ontologies must find a balance between hiding as much complexity as possible and effectively using the complex representations in an ontology. One potentially efficient way to reduce the complexity that must be shown explicitly is to apply task- and context-specific degree-of-interest functions.

### **Relationships and Types**

So far, most applications don't exploit the complex set of relationships in ontologies but rather reduce them to a simple hierarchy. In the context of ontologies, gene expression data is often visualized as a treemap.<sup>5</sup> Although treemaps are certainly useful for data analysts, they have difficulties depict-

ing rich semantics and complex relationships in ontologies. The lack of an effective visualization technique for observing and exploring ontologies, combined with their large sizes, makes the manual maintenance of ontologies a considerable effort for ontologists.

To help meet this challenge, researchers have developed cross-product extensions of the Gene Ontology that computationally define relationships in the ontology.<sup>6</sup> Reasoners can use these extensions to assist ontologists with these daunting tasks and thus minimize errors and inconsistencies.

### **Analysts' Requirements**

Finally, to make ontologies more useful for data analysis, ontologists must understand what analysts want to investigate and how they can use ontologies for their specific tasks. For many applications and datasets, we don't yet know how ontologies could support investigations and where the ontologies' real value lies. To this end, we need more empirical research to improve the understanding of analysts' workflows, tasks, and aims.

### **Research Opportunities**

The previous challenges are also opportunities. Winston Churchill said, "a pessimist sees the difficulty in every opportunity; an optimist sees the opportunity in every difficulty." Biological ontologies' availability, coupled with the abundance of biological data, presents exciting possibilities for widening the current research scope or establishing new frontiers.

### **Visualizing Ontologies**

Ontologies provide powerful, concise conceptualizations of domain knowledge that must be machine-processable but also human-readable to improve ease of use and to ease communication across stakeholders. Improving cognitive support for ontology understanding is particularly important in the biomedical and biological domains. Biomedical ontologies are typically large and complex and usually have been developed collaboratively by a community of stakeholders over a long time.<sup>7</sup>

Ontology authors need cognitive support as they try to understand what's already in an ontology that they are developing further. Ontology users need cognitive support to help them find relevant concepts for a data annotation task. To address the human-readable requirements for ontology use, researchers have designed many visualization tools for specific tasks such as finding pertinent terms for data annotation or to support ontology authoring or data alignment.<sup>8</sup>

Some visualization tools support such key tasks as providing ontology overviews, visualizing taxonomic relationships, browsing ontology classes and instances, and displaying numerous role relationships and properties of ontological concepts. Other tools present existing or proposed mappings between ontologies or support axiom validity and error checking in more sophisticated ontologies. In addition, some tools support searching and advanced querying of ontologies that can be helpful during data curation.

Ontology visualization tools have applied such methods as

- indented lists (for example, Protégé),
- tree layouts (for example, Jambalaya),
- network views (for example, OntoViz),
- matrix views (for example, MatrixBrowser), and
- treemaps (for example, Jambalaya).

---

## ***We need powerful, easy-to-use tools for both data curators and users browsing annotated content (such as journal articles).***

---

Most tools provide various interaction techniques such as zoomable views (Jambalaya), hyperbolic views, concept-anchored exploration (TouchGraph in TGViz), and time-based views (TimeViz). Collaborative ontology visualization is also increasing, as ontology-authoring tools themselves become more collaborative (for example, BioMixer).

### ***Visualizing Annotated Content***

Despite much research on ontology visualization, most tools focus on visualizing or navigating the ontologies themselves rather than visualizing ontologically annotated content. We need powerful, easy-to-use tools for both data curators and users browsing annotated content (such as journal articles). Such tools can also enable the users to do richer analyses of the content.

Text is probably the most important medium for knowledge exchange in most sciences. Although text has the advantage of general expressiveness, we perceive it sequentially, which slows down its consumption. Text-mining and text visualization approaches aim to accelerate specific tasks, such as searching, browsing, or recognition, at different aggregation levels (for example, paragraph, document, or collection). But these approaches

require automatic text-processing methods that try to retrieve information from unstructured data sources that allow considerable freedom in expressing the information. The freedom involves not only which facts to represent but also which terminology to use, which rhetorical form to use, and so on.

Using ontologies as structured background knowledge to allow consistent information extraction from or annotation of unstructured text seems promising. Text-mining methods employ ontological knowledge<sup>9</sup> and allow semiautomatic extension of ontologies by text analysis.<sup>10</sup> The opportunity here is to focus on how visualizations of text and documents can beneficially include ontological background knowledge.

First, you can enrich plain text with information retrieved from concept mapping. For example, highlighted text snippets can represent ontological concepts. When the cursor hovers over these highlights, additional concept information appears. Advanced highlighting methods such as variable text scaling<sup>11</sup> allow better discrimination between highlights and regular text.

Second, you can use metadata and background knowledge to navigate documents and document collections. For example, GoPubMed sorts PubMed search results according to the Gene Ontology hierarchy by investigating document abstracts for Gene Ontology concepts. Users can select sub-hierarchies to further filter the search results. The Bohemian Bookshelf interlinks visualizations of a digital-library collection. It provides access to books, using different views showing meta-information such as the historical period of the content and the cover color. This approach not only supports navigation but also enables serendipitous discoveries. Docuburst represents, for a given document, the subtree of a background structure based on WordNet. It allows background-structure-guided analysis of documents.

### ***Automated Visualization Generation***

The availability of domain-specific ontologies provides the exciting opportunity to develop automated visualization methods and services. Although interaction remains important for facilitating data exploration, learning to use a visualization can be costly, especially for occasional users, and interacting with the system can be time-consuming. In many application scenarios, automatically generated visualizations might serve us more efficiently and effectively and might facilitate knowledge sharing. For instance, an organization might have datasets that arrive regularly or have many users,

including new ones, who need to handle a variety of data. In such scenarios, overemphasizing the need for interaction would be a great waste of human resources.

To automate visualization, Owen Gilson and his colleagues proposed using three ontologies.<sup>12</sup> A Domain Ontology (DO) captures a collection of domain concepts and their relationships. A Visual Representation Ontology (VRO) stores the knowledge about visualization tools, styles, and the parameter space typically applicable to the domain. A Semantic Bridging Ontology (SBO) brings together the DO and VRO by gathering common mappings from the DO to the VRO.

For example, a VRO might store a set of commonly used visual representations for bio-information, such as pixel-based representations, network representations, time series plots, and molecular representations. Each representation is connected to a set of encoding parameters, such as axes or a color map. Given a biological dataset, the visualization system can obtain its main facets (for example, categories, components, and attributes), typically by reading the metadata. It can also determine the corresponding domain concepts (nodes) in the DO. For each DO concept, the SBO provides one or a few connections to visualization concepts in the VRO. The domain concepts associated with the dataset trigger a collection of visualization concepts with different weights stored in the SBO. This collection likely might suggest more than one visual representation, with the weights ranking each representation. The representations composed by the VRO can easily be translated to instructions for the visualization system (for example, scripting-language constructs or API functions and parameters).

With the availability of digital knowledge in the form of ontologies, future visualization tools might harvest such knowledge to make connections between biological data to be visualized and the semantic information in relevant ontologies. Such tools might let users explore visualization images' semantic context and might generate illustrative visualizations with semantic annotations based on relevant ontologies.

Automation doesn't necessarily prohibit choices. On the contrary, it often might facilitate more efficient and effective user interaction in creating appropriate visualizations. For example, a visualization system could automatically generate a "design gallery" offering users multiple optional visual representations.<sup>13</sup> Simply by selecting preferred designs, users could explore the design space and

refine a design iteratively. In addition, they could provide the SBO with new knowledge about relationships between domain concepts in the DO and visualization concepts in the VRO.

Min Chen and his colleagues outlined a vision for knowledge-assisted visualization.<sup>1</sup> In a visualization infrastructure supporting many users, users can acquire and share domain-specific knowledge about various visualization aspects (for example, commonly used visual representations, the performance of different tools and APIs, and the most effective parameter set). Such knowledge can help reduce the burden on users who are

---

## ***The application of ontologies and even multiple ontologies to annotate text corpora offers new potential and new challenges.***

---

learning to use complex visualization techniques. It also enables the visualization community to learn and model best practices, so that powerful visualization infrastructures can develop and evolve. Ontologies will doubtlessly be indispensable in developing infrastructures for knowledge-assisted visualization.

### ***Visualizing Ontological Context to Support Search***

Ontologies intrigue us because they offer a conceptualization of a domain. This conceptualization contains considerable structure combined with juxtaposed semantics; however, it also contains considerable variation. So, challenges abound, alongside tempting potential that invites consideration of ontologies as a key, or at least a major player, in unraveling our information maze.

Many problems remain unsolved regarding the use of visualization and interaction to explore ontologies. However, the application of ontologies and even multiple ontologies to annotate text corpora offers new potential and new challenges.

For example, even though ontologies have powerful hierarchical features, they're not just hierarchies. So, although you can use tree visualizations to create an ontology visualization, you must address the additional structure. In addition, ontologies usually have more structure than general graphs, but exactly how to exploit this additional structure in visualizations isn't clear.

Also, ontologically annotated text provides semantically rich metadata. Because visualizing

## Exploring Ontologies in Biological Visualization

During the 2012 Dagstuhl Seminar on Biological Data Visualization, a discussion group (comprising the main article's authors) explored the topic of ontologies in biological visualization. The group followed a classic design process.<sup>1-3</sup> The discussion was broadly divided into a problem phase and a solution phase. Each phase featured divergent and convergent stages: discover and define for the problem phase and develop and deliver for the solution phase.

### Discover and Define

The group first broke down the topic into the "Four Cs":

- *components* (parts) of the topic,
- *characters* (people) who are involved in some way,
- *challenges* to be aware of, and
- *characteristics* (features or behaviors) you might encounter.

This approach helped the group better define the topic while gaining a holistic view of the design problem.

### Develop and Deliver

To establish a shared vision of the outcome, the group

engaged in the Draw the Box activity. The members worked together as a team, imagining a software system as an end product that would be provided in a box on a shelf, just like any product in a shop. By focusing on different users (some of the characters explored earlier), the group designed the product's packaging.

This activity was not only great fun but also a good tool for pooling ideas and visualizing an outcome. The tangible artifact transformed the two-day meeting to a memorable vision shared by the group members. It also sparked further research activities, including the preparation of this article.

### References

1. "The Design Process," Design Council; [www.designcouncil.org.uk/designprocess](http://www.designcouncil.org.uk/designprocess).
2. D. Gray, *Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers*, O'Reilly, 2010.
3. L. Hohmann, *Innovation Games: Creating Breakthrough Products through Collaborative Play*, Addison-Wesley, 2006.

even simple metadata (for example, the Bohemian Bookshelf) can enhance serendipity in information exploration, visualizing annotated text is very promising. However, the complexity of text visualization coupled with the complexity of ontology visualization makes this a big challenge.

Finally, researchers are exploring search within ontologies. This can expand to consider search across multiple related ontologies or be inverted to include search via multiple ontologies.

Serendipity in day-to-day life can be enhanced through the organizational efforts of our fellow humans. For example, browsing for books on traditional library bookshelves frequently leads to unexpected, useful findings. These findings are at least somewhat due to librarians organizing and cataloging the books. Working with ontologies that are both machine-processable and human-readable could bring such richness to interactive computer search.

Although this article has focused on biological data visualization, the viewpoints we've presented are relevant to visualization and visual analytics in general. For example, visualization techniques and editing tools developed for creating, maintaining, and extending biological ontologies will most likely be applicable to other disciplines. This could facilitate the development

and use of domain-specific ontologies in a wider spectrum of science, engineering, medicine, the social sciences, and the humanities. It could also enable visualization to play a more significant role in core computer science subjects such as artificial intelligence, Web sciences, and software engineering.

By capitalizing on ontologies as knowledge representations, we'll be able to take a significant step toward realizing knowledge-assisted visualization. Such visualization might take the form of automated visual annotation of texts, documents, and corpora; automated construction of visualizations for novice users; automated recommendation of best visualization practices; or automated visualization of ontological context in information retrieval. Because ontologies can be dynamically updated, especially through their deployment in visualization processes, the realization of these research directions will let us systematically capture, record, and reuse visualization knowledge. This will greatly improve our capacity to deal with the ever-increasing data deluge. 

### Acknowledgments

This article resulted from a discussion group at the 2012 Dagstuhl Seminar on Biological Data Visualization. For more information, see the "Exploring Ontologies in Biological Visualization" sidebar.

## References

1. M. Chen et al., "Data, Information, and Knowledge in Visualization," *IEEE Computer Graphics & Applications*, vol. 29, no. 1, 2009, pp. 12–19.
2. T.R. Gruber, "A Translation Approach to Portable Ontology Specifications," *Knowledge Acquisition*, vol. 5, no. 2, 1993, pp. 199–220.
3. M. Ashburner et al., "Gene Ontology: Tool for the Unification of Biology," *Nature Genetics*, vol. 25, no. 1, 2000, pp. 25–29.
4. N.F. Noy, "A Framework for Ontology Evolution in Collaborative Environments," *The Semantic Web—ISWC 2006*, LNCS 4273, Springer, 2006, pp. 544–558.
5. J. Bernhardt et al., "Visualizing Gene Expression Data via Voronoi Treemaps," *Proc. 6th Int'l Symp. Voronoi Diagrams*, 2009, pp. 233–241.
6. C.J. Mungall et al., "Cross-Product Extensions of the Gene Ontology," *J. Biomedical Informatics*, vol. 44, no. 1, 2011, pp. 80–86.
7. N.F. Noy et al., "Developing Biomedical Ontologies Collaboratively," *Proc. Am. Medical Informatics Assoc. Ann. Symp.*, 2008, pp. 520–524.
8. A. Katifori and C. Halatsis, "Ontology Visualization Methods—a Survey," *ACM Computing Surveys*, vol. 39, no. 4, 2007.
9. A. Hotho et al., "Ontologies Improve Text Document Clustering," *Proc. 3rd IEEE Int'l Conf. Data Mining (ICDM 03)*, 2003, pp. 541–544.
10. T. Wächter and M. Schroeder, "Semi-automated Ontology Generation within OBO-Edit," *Bioinformatics*, vol. 26, no. 12, 2010, pp. i88–i96.
11. A. Stoffel et al., "Document Thumbnails with Variable Text Scaling," *Computer Graphics Forum*, vol. 31, no. 3, 2012, pp. 1165–1173.
12. O. Gilson et al., "From Web Data to Visualization via Ontology Mapping," *Computer Graphics Forum*, vol. 27, no. 3, 2008, pp. 959–966.
13. J. Marks et al., "A General Approach to Setting Parameters for Computer Graphics and Animation," *Proc. Siggraph*, 1997, pp. 389–400.

**Sheelagh Carpendale** is a professor of information visualization in the University of Calgary's Department of Computer Science. Contact her at [sheelagh@ucalgary.ca](mailto:sheelagh@ucalgary.ca).

**Min Chen** is a professor of scientific visualization at the University of Oxford. Contact him at [min.chen@oerc.ox.ac.uk](mailto:min.chen@oerc.ox.ac.uk).

**Daniel Evanko** is the chief editor of *Nature Methods* at Nature Publishing Group. Contact him at [d.evanko@us.nature.com](mailto:d.evanko@us.nature.com).

**Nils Gehlenborg** is a research associate at Harvard Medical School. Contact him at [nils@hms.harvard.edu](mailto:nils@hms.harvard.edu).

**Carsten Görg** is an instructor in the University of Colorado School of Medicine's Computational Bioscience Program. Contact him at [carsten.goerg@ucdenver.edu](mailto:carsten.goerg@ucdenver.edu).

**Larry Hunter** is a professor of pharmacology, preventive medicine and biometrics, computer science, and biology at the University of Colorado. Contact him at [larry.hunter@ucdenver.edu](mailto:larry.hunter@ucdenver.edu).

**Francis Rowland** is a user experience designer and Web developer at the European Bioinformatics Institute. Contact him at [frowland@ebi.ac.uk](mailto:frowland@ebi.ac.uk).

**Margaret-Anne Storey** is a professor of computer science at the University of Victoria. Contact her at [mstorey@uvic.ca](mailto:mstorey@uvic.ca).

**Hendrik Strobelt** is a postdoctoral fellow in information visualization and visual analytics at the Polytechnic Institute of New York University. Contact him at [hendrik@strobelt.com](mailto:hendrik@strobelt.com).

Contact department editor Theresa-Marie Rhyne at [theresamarierhyne@gmail.com](mailto:theresamarierhyne@gmail.com).

 Selected CS articles and columns are also available for free at <http://ComputingNow.computer.org>.



**NEW**  
IEEE  computer society  
**STORE**

Find the latest trends and insights for your

- presentations
- research
- events

[webstore.computer.org](http://webstore.computer.org)

**Save up to 40%**  
on selected articles, books, and webinars.