Ontology-Based Interactive Visualization of Patient-Generated Research Questions

David Borland, PhD[1], Laura Christopherson, PhD[1], Charles Schmitt, PhD[2]
[1]RENCI, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA;
[2]National Institute of Environmental Health Sciences, Durham, NC, USA

Abstract

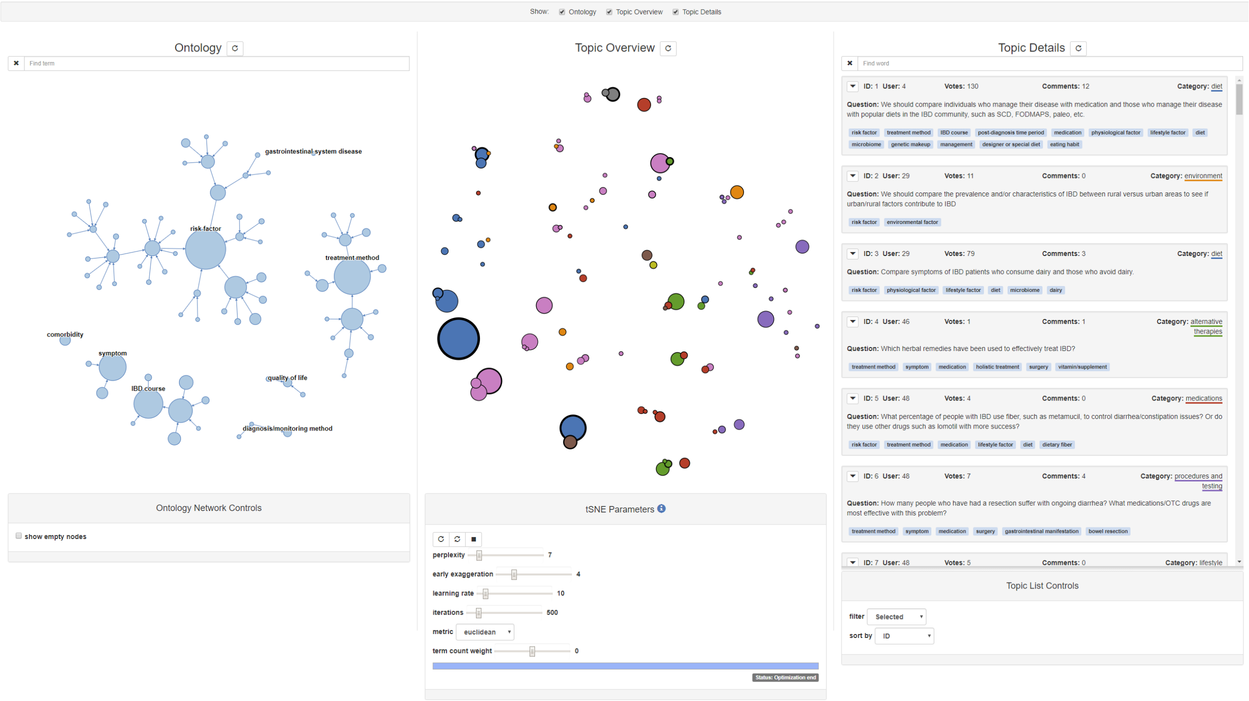
Crohn's disease and colitis are chronic conditions that affect every facet of patients' lives (e.g., social interaction, family, work, diet, and sleep). Thus, treatment consists largely of disease management. The University of North Carolina at Chapel Hill chapter of the Crohn's and Colitis Foundation–IBD Partners–has created an interactive website that, in addition to providing helpful information and disease management tools, provides a discussion forum for patients to talk about their experiences and suggest new lines of research into Crohn's disease and colitis. In order to help researchers and physicians better understand how patients think about their conditions and what research questions these patients would like the researchers to pursue, we have created an interactive visualization tool that incorporates an ontology describing the major themes and topics in the discussion forum. The tool employs linked views of (a) the ontology, (b) a research topic overview clustered by relevant ontology terms, and (c) a detailed view of the discussion forum content. In this paper we describe the creation of the ontology, discuss visualizations and interactions enabled by the visualization tool, provide an example scenario using the tool, and discuss limitations and future work based on feedback from potential users.

1 Introduction

Crohn's disease is an inflammatory bowel disease (IBD) with symptoms that can include diarrhea, inflammation of both the gut and other parts of the body, fatigue, abdominal pain, and weight loss, among others. Colitis refers to inflammation of the inner lining of the colon, and commonly co-occurs with Crohn's disease. There is no known cure for either condition, although certain therapies can help treat their symptoms, sometimes bringing about long-term remission. Thus, treatment largely consists of disease management. Given the varied ways in which these conditions can present themselves in different patients, and their chronic nature that affects every facet of patients' lives (e.g., social interaction, family, work, diet, and sleep), researchers in the University of North Carolina at Chapel Hill chapter of the Crohn's and Colitis Foundation–IBD Partners (formerly CCFA Partners)–are interested in engaging patients to aid them in disease management and to collect information useful for researching potential treatments. To this end, they have created an interactive website that provides a discussion forum for patients to talk about their experiences, suggest and discuss new lines of research into their conditions, and vote on promising research topics[1].

Although such a forum can be invaluable for generating and prioritizing research questions based on patient experiences, it can be time and labor intensive sifting through all of the questions and comments on the discussion forum, trying to effectively interpret such a large volume of text. IBD Partners is interested in developing more efficient approaches for identifying common themes and determining which research questions are most frequently discussed by patients. Interactive visualization offers a potential solution to help clinicians and researchers explore the data and identify the salient questions and needs of the patients.

In this paper we describe an interactive visualization tool developed to help researchers and physicians better understand how patients think about their condition and what research questions these patients would like the researchers to pursue. The CCFA Explorer tool employs linked views of (a) an ontology developed from concepts discussed by patients, (b) a research topic overview clustered by relevant ontology terms, and (c) a detailed view of the discussion forum content (Figure 1). We describe the creation of the ontology, discuss visualizations and interactions enabled by the CCFA Explorer tool, provide an example scenario using the tool, and discuss limitations and future work based on feedback from potential users. We believe this work will be useful more broadly given the growing focus on patient-centered outcomes approaches.



perplexity 7
early exaggeration 4
learning rate 10
iterations 500
metric euclidean
term count weight 0
Status: Optimization end

ID: 1 **User:** 4 **Votes:** 130 **Comments:** 12 **Category:** diet

Question: We should compare individuals who manage their disease with medication and those who manage their disease with popular diets in the IBD community, such as SCD, FODMAPS, paleo, etc.

risk factor · treatment method · IBD course · post-diagnosis time period · medication · physiological factor · lifestyle factor · diet · microbiome · genetic makeup · management · designer or special diet · eating habit

ID: 2 **User:** 29 **Votes:** 11 **Comments:** 0 **Category:** environment

Question: We should compare the prevalence and/or characteristics of IBD between rural versus urban areas to see if urban/rural factors contribute to IBD

risk factor · environmental factor

ID: 3 **User:** 29 **Votes:** 79 **Comments:** 3 **Category:** diet

Question: Compare symptoms of IBD patients who consume dairy and those who avoid dairy.

risk factor · physiological factor · lifestyle factor · diet · microbiome · dairy

ID: 4 **User:** 46 **Votes:** 1 **Comments:** 1 **Category:** alternative therapies

Question: Which herbal remedies have been used to effectively treat IBD?

treatment method · symptom · medication · holistic treatment · surgery · vitamin/supplement

ID: 5 **User:** 48 **Votes:** 4 **Comments:** 0 **Category:** medications

Question: What percentage of people with IBD use fiber, such as metamucil, to control diarrhea/constipation issues? Or do they use other drugs such as lomotil with more success?

risk factor · treatment method · medication · lifestyle factor · diet · dietary fiber

ID: 6 **User:** 48 **Votes:** 7 **Comments:** 4 **Category:** procedures and testing

Question: How many people who have had a resection suffer with ongoing diarrhea? What medications/OTC drugs are most effective with this problem?

treatment method · symptom · medication · surgery · gastrointestinal manifestation · bowel resection

ID: 7 **User:** 48 **Votes:** 5 **Comments:** 0 **Category:** lifestyle

filter Selected
sort by ID

Figure 1: The CCFA Explorer interface: Ontology (left), Topic Overview (middle), and Topic Details (right).

2 Previous Work

We developed an ontology based on the major themes discussed in the forum to enable more effective analysis and visualization. Previous work on biomedical ontologies and interactive visualization is discussed below.

2.1 Ontologies

Ontologies are controlled vocabularies that represent knowledge about a domain of interest[2]. They offer richer representations than other controlled vocabularies (e.g., taxonomies, thesauri) because they enable relationships beyond hierarchical and synonymous. Ontologies have a long history in medicine and biological research[3–5], and are used for a variety of purposes, e.g., classifying literature for information retrieval, mapping and integrating diverse data sources, aggregating/clustering information, and natural language processing applications[6,7].

Biomedical ontologies tend to focus on representing encyclopedic knowledge about a given domain. For example, UBERON[1] contains approximately 20,000 concepts ranging from very granular anatomy (e.g., cell membrane) to larger systems (e.g., digestive system). Other biomedical ontologies cover chemicals, drugs, phenotypes, diseases, adverse events, and more. Each of these ontologies tend to be very large, containing thousands of concepts.

2.2 Interactive Visualization

Interactive visualization has proven to be a useful method for analyzing datasets across a wide range of disciplines, including in the health care domain[8,9]. To develop effective interactive visualizations, Shneiderman's visual information seeking mantra–*overview first, zoom and filter, then details on demand*–has been adopted by a wide range of data visualization tools[10]. We adopt this approach, providing overviews of the CCFA ontology and forum content, along with the ability to filter and obtain detailed forum content based on patterns and relationships discovered from interacting with the overviews. Our approach is similar to that of Jigsaw[11], but also incorporating an ontology to provide structure to the relevant concepts.

[1] http://www.obofoundry.org/ontology/uberon.html

2.3 Visualization of Ontologies

Ontologies are typically organized primarily as hierarchies, and therefore hierarchical visualization techniques, such as tree maps[12], icicle plots[13], and tree diagrams (e.g., tidy trees[14]), can be used to visualize them. Although such visualization techniques are effective for showing hierarchical structure, they are not designed to show other types of ontological relationships. Network diagrams offer the ability to encode different types of relationships via different styles of links in the diagram, and due to this flexibility we adopted this approach, although currently only sowing hierarchical relationships. Kamdar et al. present research analyzing user interactions with biomedical ontologies for different visualization types, including network diagrams, and show that different users interact with ontologies differently[15], suggesting that a suite of approaches may be useful, which will help inform our future work.

3 Forum Data

The data snapshot used when creating the CCFA forum ontology consists of 97 *research topics* (i.e., user posts consisting of a proposed research question and a description of the question), and 121 user comments made by fellow patients on proposed questions, for a total of 17,322 words. An example topic post is the following:

> *Question:*
>
> Nicotine has shown to be effective for UC [ulcerative colitis] in some individuals, both prior- and non-smokers. What is the mechanism? Does nicotine affect the microbiome, the immune system or both?
>
> *Description:*
>
> Big Pharma will not take on the role of studying nicotine as there is no $$$ in it. Few studies with small sample sizes have been done but more research is needed.

Each research topic also has an anonymized user ID (400 unique users), the number of votes for each topic (1246 total votes), and one of 9 predefined categories (*diet, medications, procedures and testing, environment, alternative therapies, lifestyle, genetics, exercise,* and *other*) selected by the topic creator.

4 Ontology Creation

To facilitate visualization and analysis of the forum data, we initially performed some basic linguistic processing on the forum text, such as calculating word and phrase frequencies, but the results did not effectively capture the forum conversation. Terms such as *inflammatory bowel*, *controlled trial*, and *disease activity* appeared frequently, which simply confirmed the obvious: patients were discussing IBD and its research. These frequencies did not capture the nuance of specific lines of research the patients were interested in. We therefore created an ontology of the forum conversation to better capture the depth and breadth of research topics in which the patients were interested.

We first conducted an in-depth, manual exploration of the forum text. Specifically, we applied content analysis to the forum text, sifting through manifest content (i.e., what is seen directly in the text, such as the occurrence of a particular word) to find latent content (i.e., underlying meaning, connotation, nuance). According to Wildemuth, “An example of latent content is the level of research anxiety present in user narratives about their experiences at the library”[16]. In other words, a user may not directly state, “I am so anxious.” Instead, the anxiety may be implied, e.g., “My heart won’t stop beating so fast” or “I wish I could relax.” Wildemuth notes, “Sometimes there is no existing theory or research on your message populations; you may not know what the important variables are. The only way to discover them is to explore the content”[16]. In other words, it may be impossible to identify themes without first immersing one’s self in the text, allowing the themes to be revealed as one becomes more intimate with the conversation. This is reflected in the fact that the most common predefined category assigned by users to their proposed topic was *other* (34 out of 97, over 40%), implying that the categories did not fully capture the breadth of their interests and discussions.

After completing the content analysis, it was clear that no existing ontology would adequately represent the patient conversations. Most biomedical ontologies provide encyclopedic objective knowledge about a particular subject, whereas the CCFA forum text describes personal patient experiences, emotions, and desires. The goal of CCFA physicians and researchers is to understand their patients’ needs and wants, and an effective ontology needs to reflect this goal to help bridge the gap between how clinical practitioners, researchers, and patients view their conditions.

Our ontology structured and classified the raw information in the forum. Concepts (e.g., *medication*, *surgery*, *diet*, *symptom*) discussed in the forum became "classes" in the ontology. Although relationships beyond hierarchical (e.g., medication *treats* symptom) are possible, our ontology does not currently include such relationships, and functions primarily as a concept network. Additional concepts (primarily from the CCFA website to align with their approach to care) were included in anticipation of future forum conversations. Where applicable, classes from pre-existing ontologies (the Ontology for Adverse Events and the Disease Ontology)[2] were used. The resulting ontology describes a hierarchy of 337 total classes, with seven top-level classes: *comorbidity*, *diagnosis/monitoring method*, *IBD course*, *quality of life*, *risk factor*, *symptom*, and *treatment method* (Figure 1, left). The ontology was created using Protégé[17], exported in OWL format, and converted to an OBO Graph using ROBOT for easy ingestion into our visualization tool[3]. Based on the content analysis, each research topic was labeled with one or more terms from the ontology. The ontology structure and linkage to the research topics enable the interactive visualization described in the next section.

5 CCFA Explorer

The CCFA Explorer tool consists of three different interactive visualizations: (a) the CCFA forum ontology, (b) an overview of the patient-generated research topics, and (c) a detailed view of the forum text and other information about each research topic (Figure 1). Users can select visual elements representing different aspects of the data (research topics or ontology terms) in each view, and all three views are linked to automatically highlight relationships from the various visual elements in each view to the selected items.

5.1 Ontology Visualization

We use a force-directed network to show the ontology structure and indicate the most prominent ontology terms (Figure 1, left). Each term is represented by a node in the visualization, and links (i.e., arrows connecting nodes) indicate "is a" hierarchical relationships (e.g., *medicine* is a *treatment method*). Node size is proportional to the number of research topics labeled with that term. For any given term, if a topic has been classified with that term, it is also assumed that all ancestors of that term are also associated with that topic, therefore no child node will ever be larger than its parent. When the visualization initially loads, node labels for top-level terms in the ontology are visible. Labels for other nodes appear when the user hovers over a node, or upon user selection as described in Section 5.4. The user can also show and highlight in red any node label by searching for that term in the term search box.

5.2 Topic Overview

The topic overview uses t-SNE[18] to lay out circular glyphs representing each research topic (Figure 1, middle), placing topics with similar sets of ontology terms closer together, enabling the user to visibly identify clusters of similar topics. The size of each glyph is proportional to the number of forum comments made in response to that topic, and the outline thickness is proportional to the number of user votes for that topic, enabling the user to identify popular topics. The glyph color represents which of the 9 predefined categories was chosen by the topic creator. We introduce three modifications to the standard t-SNE layout to enable more effective visualization of the CCFA forum data. (1) Because two or more topics may be labeled with similar sets of ontology terms, glyphs may overlap and occlude each other, making it difficult to see cluster sizes for very similar topics, and to see and select individual topics. We therefore apply a force-directed layout for overlapping glyphs that separates the centers of each glyph while maintaining some overlap to indicate closely-related clusters (Figure 2-a). (2) Due to the hierarchical nature of the ontology, we enable weighting of higher-level or lower-level ontology terms to determine at which level in the hierarchy topic glyphs are clustered. Weighting higher-level terms results in fewer clusters based on more general terms (Figure 2-b), and weighting lower-level terms results in more clusters based on more specific terms (Figure 2-c). (3) Greater weights can be applied to the currently selected ontology terms, resulting in clusters reflecting combinations of the selected terms. For example, Figure 2-d shows a layout emphasizing two selected ontology terms, with three clusters indicating the presence of only the first term, only the second term, or both terms. This feature enables, for example, easy selection of all topics with a given set of terms.

[2] http://www.oae-ontology.org/ | http://disease-ontology.org/

[3] https://www.w3.org/OWL/ | https://github.com/geneontology/obographs | http://robot.obolibrary.org/

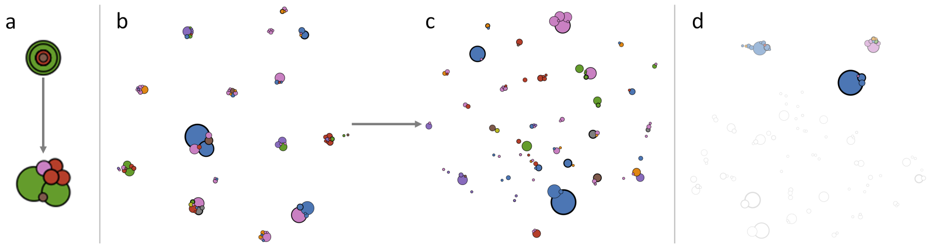
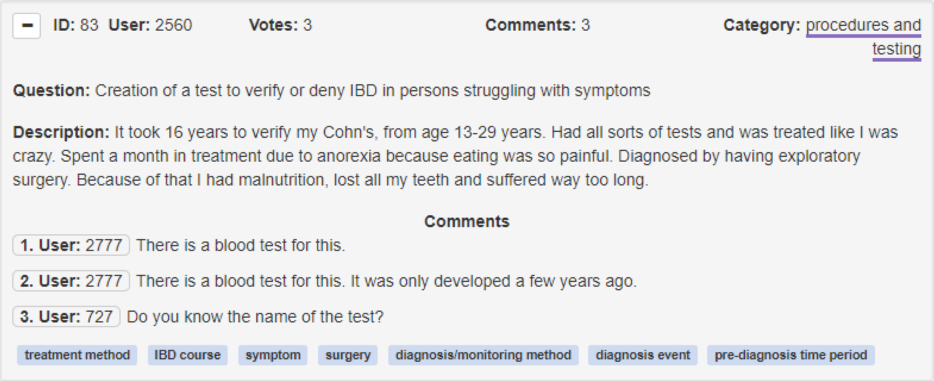


Figure 2: Modifications to the standard t-SNE layout: (a) force-directed layout of overlapping glyphs to increase cluster legibility, (b and c) differential weighting of ontology terms emphasizing (b) higher-level terms resulting in fewer, more general clusters and (c) lower-level terms resulting in a larger number of more specific clusters, and (d) emphasizing the currently selected ontology terms for clustering.



− **ID:** 83 **User:** 2560 **Votes:** 3 **Comments:** 3 **Category:** procedures and testing

Question: Creation of a test to verify or deny IBD in persons struggling with symptoms

Description: It took 16 years to verify my Cohn's, from age 13-29 years. Had all sorts of tests and was treated like I was crazy. Spent a month in treatment due to anorexia because eating was so painful. Diagnosed by having exploratory surgery. Because of that I had malnutrition, lost all my teeth and suffered way too long.

Comments

1. User: 2777 There is a blood test for this.

2. User: 2777 There is a blood test for this. It was only developed a few years ago.

3. User: 727 Do you know the name of the test?

treatment method | IBD course | symptom | surgery | diagnosis/monitoring method | diagnosis event | pre-diagnosis time period

Figure 3: An example topic in the Topic Details view.

5.3 Topic Details

The topic details view is a scrollable list of panels for each topic in the forum. Each topic panel contains the research question, description, and comments for that topic, along with additional information such as the number of user votes, color-coded user-selected category, and tags indicating the ontology terms for that topic (Figure 3). Users may select three different levels of detail to display each topic's text: (1) question only, (2) question and description, and (3) question, description, and comments. The list of topics can be sorted by *topic ID, user ID, number of votes, number of comments,* and *category*. The list can also be filtered based on selected topics or ontology terms, as described in Section 5.4. In addition, the user can highlight in red any text searched for in the search box.

5.4 Interactive Selection and Highlighting

The user can interactively select visual elements representing ontology terms or research topics in any of the three views, and all views will be automatically updated to highlight relationships to the selected items, enabling the user to effectively explore the forum data.

5.4.1 Selection

We define three types of possible relationships between ontology terms and research topics: (1) The *co-occurrence* between two ontology terms is the number of topics that have been labeled with both terms, and therefore is an indication of which ontology terms are discussed together by the forum users. For multiple selected terms, the co-occurrence between a term and the selection is the size of the union of the common topics. (2) The *association* between two topics is the number of ontology terms that the two topics share in common, and is an indication of how closely related the two topics are. For multiple selected topics, the association between a topic and the selection is the size of the union of the common terms. (3) The *connection* between an ontology term and a topic is 1 if the topic is labeled

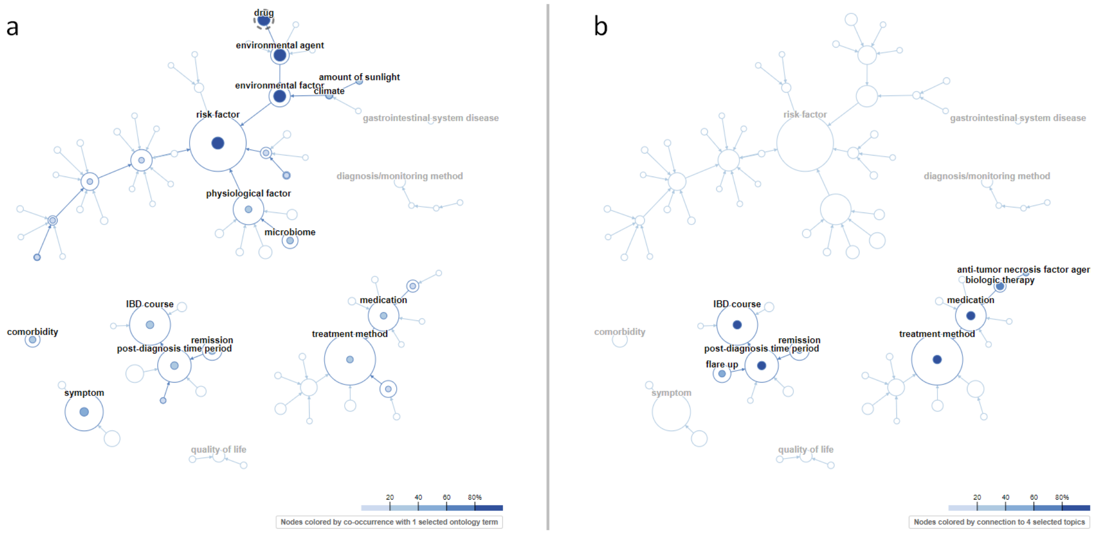


Figure 4: Interactive highlighting of the ontology visualization, enabling (a) highlighting of co-occurrences with a selected ontology term (*drug*), and (b) highlighting of connections to topics selected in one of the other views.

with that term, and 0 otherwise. For multiple selected terms or topics, the connection is the sum of each individual connection.

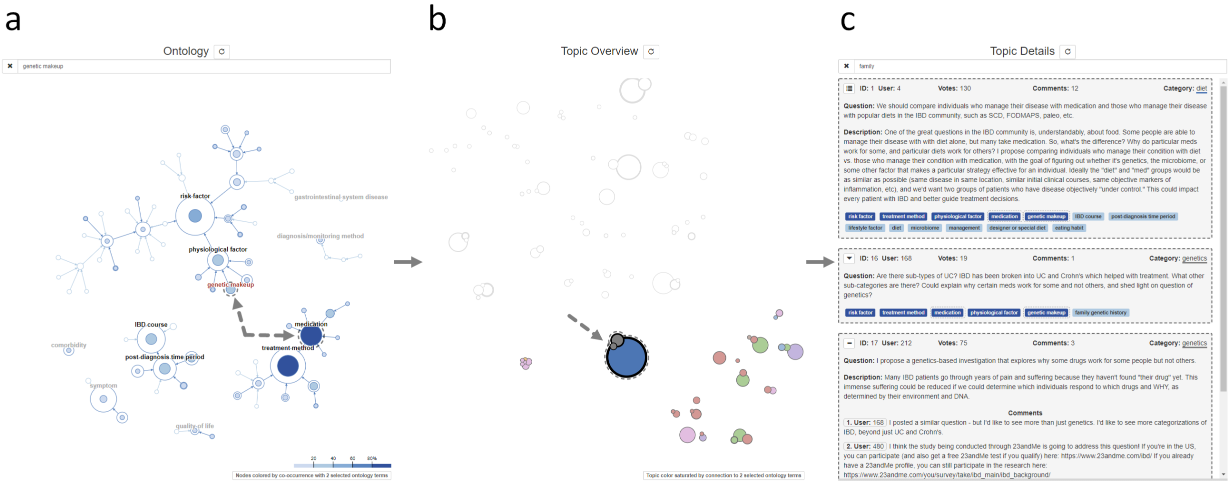
In the ontology view, terms can be selected by clicking on the node for that term. In the topic overview, topics can be selected by clicking on the glyph for that topic. In the topic details view, topics can be selected by clicking on the panel for that topic, and ontology terms can be selected by clicking on the tag for that term in any given topic. In all views, selected visual elements are represented by dashed outlines.

5.4.2 Highlighting

In the ontology visualization, the co-occurrence with any currently selected terms is represented by an inset circle for each node, with size proportional to the co-occurrence and color proportional to the percent co-occurrence (co-occurrence divided by total number of topics connected to the selected terms $\times 100$) with the selected terms (Figure 4-a). Similarly, the association with any currently selected topics is represented by an inset circle with radius proportional to the association, and color proportional to the percent association (association divided by total number of selected topics $\times 100$) with the selected topics (Figure 4-b). In both cases, labels are displayed for any nodes with a percent co-occurrence/association of at least 25%. In the case of selected terms and selected topics, highlighting topic connections takes precedence in the ontology visualization. Automatic highlighting of the ontology visualization enables the user to quickly find ontology terms that are discussed in the same topics, and which ontology terms are related to a given group of topics.

In the topic overview, the connection with any currently selected terms is mapped to glyph color saturation, normalized by the total number of selected terms (Figure 2-d). Similarly, the association with any currently selected topics is also mapped to glyph color saturation, normalized by the by total number of terms for that glyph's topic (such that any selected topic will be fully-saturated). In the case of selected terms and selected topics, highlighting term connections takes precedence in the topic overview. Automatic highlighting of the topic overview enables the user to quickly find topics related to ontology terms of interest, and discover related topics.

In the topic details view, topics can be optionally filtered by *Selected* or *Connected*. For *Selected*, if any topics are selected, only those topics will be shown. For *Connected*, if there are any selected terms or topics, only topics with a non-zero connection or association will be shown. In this manner, the user can quickly drill down to see the forum text related to ontology terms or topics of interest. In addition, the same color map applied to the ontology nodes during highlighting is applied to the ontology term tags for each topic (Figure 5-c).


a
b
c
Ontology
genetic makeup
risk factor
gastrointestinal system disease
diagnosis/monitoring method
physiological factor
genetic makeup
IBD course
comorbidity
post-diagnosis time period
medication
treatment method
symptom
quality of life
20
40
60
80%
Nodes colored by co-occurrence with 2 selected ontology terms
Topic Overview
Topic color saturated by connection to 2 selected ontology terms
Topic Details
family

ID: 1 **User:** 4 **Votes:** 130 **Comments:** 12 **Category:** diet

Question: We should compare individuals who manage their disease with medication and those who manage their disease with popular diets in the IBD community, such as SCD, FODMAPS, paleo, etc.

Description: One of the great questions in the IBD community is, understandably, about food. Some people are able to manage their disease with with diet alone, but many take medication. So, what's the difference? Why do particular meds work for some, and particular diets work for others? I propose comparing individuals who manage their condition with diet vs. those who manage their condition with medication, with the goal of figuring out whether it's genetics, the microbiome, or some other factor that makes a particular strategy effective for an individual. Ideally the "diet" and "med" groups would be as similar as possible (same disease in same location, similar initial clinical courses, same objective markers of inflammation, etc), and we'd want two groups of patients who have disease objectively "under control." This could impact every patient with IBD and better guide treatment decisions.

risk factor | treatment method | physiological factor | medication | genetic makeup | IBD course | post-diagnosis time period | lifestyle factor | diet | microbiome | management | designer or special diet | eating habit

ID: 16 **User:** 168 **Votes:** 19 **Comments:** 1 **Category:** genetics

Question: Are there sub-types of UC? IBD has been broken into UC and Crohn's which helped with treatment. What other sub-categories are there? Could explain why certain meds work for some and not others, and shed light on question of genetics?

risk factor | treatment method | medication | physiological factor | genetic makeup | family genetic history

ID: 17 **User:** 212 **Votes:** 75 **Comments:** 3 **Category:** genetics

Question: I propose a genetics-based investigation that explores why some drugs work for some people but not others.

Description: Many IBD patients go through years of pain and suffering because they haven't found "their drug" yet. This immense suffering could be reduced if we could determine which individuals respond to which drugs and WHY, as determined by their environment and DNA.

Comments

1. User: 168 I posted a similar question - but I'd like to see more than just genetics. I'd like to see more categorizations of IBD, beyond just UC and Crohn's.

2. User: 480 I think the study being conducted through 23andMe is going to address this question! If you're in the US, you can participate (and also get a free 23andMe test if you qualify) here: https://www.23andme.com/ibd/ If you already have a 23andMe profile, you can still participate in the research here: https://www.23andme.com/you/survey/take/ibd_main/ibd_background/

Figure 5: Example use case.

6 Example Scenario

Figure 5 illustrates how a user might explore the forum data. The user is interested in "genetic makeup," and searches for that term in the ontology search box, which highlights the term in red (Figure 5-a). The user selects the "genetic makeup" node, which highlights the co-occurrences with that term in the other ontology nodes. The user notices that "medication" has a relatively high co-occurrence with "alternative therapies," and adds "medication" to the term selection. The user then re-runs the t-SNE in the topic overview to cluster topics primarily by these selected terms (Figure 5-b). The user notices a cluster of three topic glyphs, including one very large node (indicating a popular topic with many comments), and so selects those three nodes for closer inspection in the topic details view, which the user filters to show only the selected topics (Figure 5-c). Various other exploratory work flows are also enabled by the tool, based on the research focus of the user.

7 Feedback and Future Work

After presenting the CCFA Explorer tool to members of the IBD Partners team, we received useful feedback that will help inform our future work. In general they thought that the tool was a useful way to explore the CCFA forum data, and made it possible to quickly identify major themes and popular research topics, however they felt that some of the features may be too complex for more naïve users. Two themes in particular that were identified were, (1) the utility of a patient-facing interface to help forum users find similar patients and more easily identify research topics relevant to them, and (2) a researcher-facing interface to help researchers in specific areas quickly identify information related to their research area and generate summaries of relevant information that can be easily presented to stakeholders. To this end, we intend to refine our tool in various way. For example, the ontology visualization, while effective at showing the overall structure of the ontology and highlighting relationships with the ontology terms, is not very well suited for navigation to find ontology terms of interest. We therefore plan to redesign our ontology visualization to make navigation easier, while incorporating some of our current work in interactive highlighting. We also plan to explore the use of text summarization techniques to include in a summary panel that will present an infographic-like view of any currently selected terms/topics. Another important line of future research will involve exploring automatic and semi-automatic methods to analyze the forum text and classify topics based on existing ontology terms, or by expanding the current ontology. In addition, we will conduct fmore ormal evaluations of the tool for usability.

8 Conclusion

We have presented an interactive visualization tool that enables users to explore patient-generated research questions from a forum for individuals suffering from Crohn's disease and colitis. We described the development of an ontology created from the forum text to enable more effective visualization an exploration of the data. To our knowledge, this is the first such ontology incorporating concepts of how patients actually talk about their own conditions. Using linked views that automatically highlight relationships between selected ontology terms and research topics, the user can gain insights into concepts of importance to the forum participants. Future work will further refine the tool for specific user populations, such as patients, or researchers with a specific focus.

References

[1] Chung AE, Sandler RS, Long MD, Ahrens S, Burris JL, Martin CF, et al. Harnessing person-generated health data to accelerate patient-centered outcomes research: the Crohns and Colitis Foundation of America PCORnet Patient Powered Research Network (CCFA Partners). Journal of the American Medical Informatics Association. 2016 May;23(3):485–490.

[2] Harpring P. Introduction to Controlled Vocabularies: Terminology for Art, Architecture, and Other Cultural Works. Getty Publications; 2010.

[3] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature Genetics. 2000 May;25(1):25–29.

[4] The Gene Ontology Consortium. Expansion of the Gene Ontology knowledgebase and resources. Nucleic Acids Research. 2017 Jan;45(D1):D331–D338.

[5] Khler S, Vasilevsky NA, Engelstad M, Foster E, McMurry J, Aym S, et al. The Human Phenotype Ontology in 2017. Nucleic Acids Research. 2017 Jan;45(D1):D865–D876.

[6] Rubin DL, Shah NH, Noy NF. Biomedical ontologies: a functional perspective. Briefings in Bioinformatics. 2008 Jan;9(1):75–90.

[7] Bodenreider O. Biomedical ontologies in action: role in knowledge management, data integration and decision support. Yearbook of Medical Informatics. 2008;p. 67–79.

[8] Shneiderman B, Plaisant C, Hesse BW. Improving Healthcare with Interactive Visualization. Computer. 2013 May;46(5):58–66.

[9] Gotz D, Borland D. Data-driven healthcare: Challenges and opportunities for interactive visualization. IEEE Computer Graphics and Applications. 2016 May;36(3):90–96.

[10] Shneiderman B. The eyes have it: a task by data type taxonomy for information visualizations. In: Proceedings 1996 IEEE Symposium on Visual Languages; 1996. p. 336–343.

[11] Stasko J, Gorg C, Liu Z, Singhal K. Jigsaw: Supporting Investigative Analysis through Interactive Visualization. In: 2007 IEEE Symposium on Visual Analytics Science and Technology; 2007. p. 131–138.

[12] Shneiderman B. Tree Visualization with Tree-maps: 2-d Space-filling Approach. ACM Trans Graph. 1992 Jan;11(1):92–99.

[13] Kruskal JB, Landwehr JM. Icicle plots: Better displays for hierarchical clustering. The American Statistician. 1983;37(2):162–168.

[14] Reingold EM, Tilford JS. Tidier Drawings of Trees. IEEE Trans Softw Eng. 1981 Mar;7(2):223–228.

[15] Kamdar MR, Walk S, Tudorache T, Musen MA. Analyzing user interactions with biomedical ontologies: A visual perspective. Journal of Web Semantics. 2018 Mar;49:16–30.

[16] Wildemuth BM. Applications of Social Research Methods to Questions in Information and Library Science. Westport, Conn: Libraries Unlimited; 2009.

[17] Musen MA. The Protege Project: A Look Back and a Look Forward. AI matters. 2015 Jun;1(4):4–12.

[18] van der Maaten LJP, Hinton GE. Visualizing High-Dimensional Data Using t-SNE. Journal of Machine Learning Research. 2008;9:2579–2605.