Ontology-based interactive visualization of patient-generated research questions

David Borland  RENCI, UNC-Chapel Hill
Laura Christopherson  RENCI, UNC-Chapel Hill
Charles Schmitt  NIEHS
Overview

• **Background**
  – Crohn’s disease & colitis
  – IBD Partners discussion forum

• **Ontology creation**

• **CCFA Explorer**
  – Ontology visualization
  – Research topic visualization
    • Overview
    • Details

• **Example scenario**

• **Feedback and future work**
Crohn’s Disease and Colitis

• **Crohn’s disease**
  – Inflammatory bowel disease (IBD)
  – Chronic
  – Diverse set of symptoms
    • Diarrhea
    • Inflammation (gut and other body parts)
    • Fatigue
    • Abdominal pain
    • Weight loss
    • Etc.

• **Colitis**
  – Inflammation of the inner lining of the colon
  – Commonly co-occurs with Crohn’s disease
Crohn’s Disease and Colitis

• **Crohn’s disease**
  – Inflammatory bowel disease (IBD)
  – Chronic
  – Diverse set of symptoms
    • Diarrhea
    • Inflammation (gut and other body parts)
    • Fatigue
    • Abdominal pain
    • Weight loss
    • Etc.

  No known cure

  Certain therapies can help treat *symptoms*

  Treatment largely consists of *disease management*

• **Colitis**
  – Inflammation of the inner lining of the colon
  – Commonly co-occurs with Crohn’s disease
IBD Partners
(formerly CCFA Partners)

• Patient-powered online research network
  – Crohn’s and Colitis Foundation
  – UNC-Chapel Hill School of Medicine

• Discussion forum
  – Patients propose and discuss research ideas
  – Researchers engage in discussion

Ontology-based interactive visualization of patient-generated research questions: Background
Discussion Forum Data

Research topic
Research question
Description
User ID, Votes, Category

Comment
User ID

Comment
User ID

Ontology-based interactive visualization of patient-generated research questions: Background
Discussion Forum Data

Forum snapshot:

97 research topics
121 total comments
400 unique users
1246 total votes
Example topic post:

**Question:**
Nicotine has shown to be effective for UC [ulcerative colitis] in some individuals, both prior- and nonsmokers. 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.
Discussion Forum Data

One of 9 predefined categories:
- diet
- medications
- procedures and testing
- environment
- alternative therapies
- lifestyle
- genetics
- exercise
- other
Discussion Forum Data

Goals:

- Identify common themes
- Prioritize patient-generated research questions

Ontology-based interactive visualization of patient-generated research questions: Background
Initial Approach

Phrases were extracted from forum text (i.e., from questions, descriptions, and comments). Those phrases with >2 appearances were used to create the initial visualizations.

<table>
<thead>
<tr>
<th>Words</th>
<th>Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>ulcerative colitis</td>
<td>20</td>
</tr>
<tr>
<td>inflammatory bowel</td>
<td>13</td>
</tr>
<tr>
<td>controlled trial</td>
<td>10</td>
</tr>
<tr>
<td>inflammatory bowel disease</td>
<td>10</td>
</tr>
<tr>
<td>bowel disease</td>
<td>10</td>
</tr>
<tr>
<td>top priority</td>
<td>10</td>
</tr>
<tr>
<td>periodontal disease</td>
<td>9</td>
</tr>
<tr>
<td>disease activity</td>
<td>8</td>
</tr>
<tr>
<td>vitamin d</td>
<td>7</td>
</tr>
<tr>
<td>ibd management</td>
<td>5</td>
</tr>
<tr>
<td>other auto-immune</td>
<td>4</td>
</tr>
<tr>
<td>disease</td>
<td>4</td>
</tr>
<tr>
<td>small bowel</td>
<td>4</td>
</tr>
<tr>
<td>combination</td>
<td>4</td>
</tr>
</tbody>
</table>

This didn't turn out to be as informative as we'd hoped. The frequency of words and/or word phrases didn't successfully capture the 'aboutness' of the conversation on the CCFA forum.
Initial Approach

Phrases were extracted from forum text (i.e., from questions, descriptions, and comments). Those phrases with >2 appearances were used to create the initial visualizations.

<table>
<thead>
<tr>
<th>Words</th>
<th>Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>ulcerative colitis</td>
<td>20</td>
</tr>
<tr>
<td>inflammatory bowel</td>
<td>13</td>
</tr>
<tr>
<td>controlled trial</td>
<td>10</td>
</tr>
<tr>
<td>inflammatory bowel disease</td>
<td>10</td>
</tr>
<tr>
<td>bowel disease</td>
<td>10</td>
</tr>
<tr>
<td>top priority</td>
<td>10</td>
</tr>
<tr>
<td>periodontal disease</td>
<td>9</td>
</tr>
<tr>
<td>disease activity</td>
<td>8</td>
</tr>
<tr>
<td>vitamin d</td>
<td>7</td>
</tr>
<tr>
<td>ibd management</td>
<td>5</td>
</tr>
<tr>
<td>other auto-immune</td>
<td>4</td>
</tr>
<tr>
<td>disease management</td>
<td>4</td>
</tr>
<tr>
<td>small bowel</td>
<td>3</td>
</tr>
<tr>
<td>combination</td>
<td>3</td>
</tr>
</tbody>
</table>

This didn’t turn out to be as informative as we’d hoped. The frequency of words and/or word phrases didn’t successfully capture the ‘aboutness’ of the conversation on the CCFA forum.
Ontology Creation

- **Ontology**
  - Controlled vocabulary
  - Represent knowledge about a domain of interest
  - Enable multiple types of relationships
    - “is a”
    - “treats”
    - Etc.
Ontology-based interactive visualization of patient-generated research questions: **Ontology creation**
Ontology-based interactive visualization of patient-generated research questions: **Ontology creation**

**Manifest Content vs. Latent Content**

The appearance of a given word in a text

Conceptual and cannot be directly observed
Content Analysis

Manifest Content vs. Latent Content

The appearance of a given word in a text

Conceptual and cannot be directly observed

E.g., “…the level of research anxiety present in user narratives about their experiences at the library.”

Content Analysis

Manifest Content vs. Latent Content

The appearance of a given word in a text

Conceptual and cannot be directly observed

E.g., “...the level of research anxiety present in user narratives about their experiences at the library.”

“I am so anxious”

“My heart won’t stop beating”

“I wish I could relax”

Content Analysis

Ontology-based interactive visualization of patient-generated research questions: Ontology creation
Content Analysis

“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.”

The Ontology

Protégé

Borrows some terms from pre-existing ontologies:

Ontology of Adverse Events, Disease Ontology

<table>
<thead>
<tr>
<th>REPRESENTATIVE CLASSES</th>
<th>POSTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>comorbidity</td>
<td>11</td>
</tr>
<tr>
<td>diagnosis/monitoring method</td>
<td>7</td>
</tr>
<tr>
<td>IBD course</td>
<td>39</td>
</tr>
<tr>
<td>pre-diagnosis time period</td>
<td>1</td>
</tr>
<tr>
<td>diagnosis event</td>
<td>5</td>
</tr>
<tr>
<td>post-diagnosis time period</td>
<td>31</td>
</tr>
<tr>
<td>quality of life</td>
<td>8</td>
</tr>
<tr>
<td>risk factor</td>
<td>58</td>
</tr>
<tr>
<td>demographic factor</td>
<td>7</td>
</tr>
<tr>
<td>environmental factor</td>
<td>18</td>
</tr>
<tr>
<td>lifestyle factor</td>
<td>20</td>
</tr>
<tr>
<td>physiological factor</td>
<td>28</td>
</tr>
<tr>
<td>psychological factor</td>
<td>5</td>
</tr>
<tr>
<td>symptom</td>
<td>36</td>
</tr>
<tr>
<td>gastrointestinal manifestation</td>
<td>12</td>
</tr>
<tr>
<td>extra-gastrointestinal manifestation</td>
<td>3</td>
</tr>
<tr>
<td>treatment method</td>
<td>50</td>
</tr>
<tr>
<td>alternative therapy</td>
<td>7</td>
</tr>
<tr>
<td>holistic treatment</td>
<td>12</td>
</tr>
<tr>
<td>medication</td>
<td>28</td>
</tr>
<tr>
<td>surgery</td>
<td>13</td>
</tr>
</tbody>
</table>

Ontology-based interactive visualization of patient-generated research questions: Ontology creation 20
The Ontology

Ontology-based interactive visualization of patient-generated research questions:

Protégé

337 classes
7 top-level classes

OWL→ ROBOT→ OBO Graph (JSON)

Borrows some terms from pre-existing ontologies:
Ontology of Adverse Events, Disease Ontology

<table>
<thead>
<tr>
<th>REPRESENTATIVE CLASSES</th>
<th>POSTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>comorbidity</td>
<td>11</td>
</tr>
<tr>
<td>diagnosis/monitoring method</td>
<td>7</td>
</tr>
<tr>
<td>IBD course</td>
<td>39</td>
</tr>
<tr>
<td>pre-diagnosis time period</td>
<td>1</td>
</tr>
<tr>
<td>diagnosis event</td>
<td>5</td>
</tr>
<tr>
<td>post-diagnosis time period</td>
<td>31</td>
</tr>
<tr>
<td>quality of life</td>
<td>8</td>
</tr>
<tr>
<td>risk factor</td>
<td>58</td>
</tr>
<tr>
<td>demographic factor</td>
<td>7</td>
</tr>
<tr>
<td>environmental factor</td>
<td>18</td>
</tr>
<tr>
<td>lifestyle factor</td>
<td>20</td>
</tr>
<tr>
<td>physiological factor</td>
<td>28</td>
</tr>
<tr>
<td>psychological factor</td>
<td>5</td>
</tr>
<tr>
<td>symptom</td>
<td>36</td>
</tr>
<tr>
<td>gastrointestinal manifestation</td>
<td>12</td>
</tr>
<tr>
<td>extra-gastrointestinal manifestations</td>
<td>3</td>
</tr>
<tr>
<td>treatment method</td>
<td>50</td>
</tr>
<tr>
<td>alternative therapy</td>
<td>7</td>
</tr>
<tr>
<td>holistic treatment</td>
<td>12</td>
</tr>
<tr>
<td>medication</td>
<td>28</td>
</tr>
<tr>
<td>surgery</td>
<td>13</td>
</tr>
</tbody>
</table>
The Ontology

Borrows some terms from pre-existing ontologies: Ontology of Adverse Events, Disease Ontology

Ontology-based interactive visualization of patient-generated research questions: Ontology creation

- comorbidity: 11 posts
- diagnosis/monitoring method: 7 posts
- IBD course: 39 posts
- pre-diagnosis time period: 1 post
- diagnosis event: 5 posts
- post-diagnosis time period: 31 posts
- quality of life: 8 posts
- risk factor: 58 posts
- demographic factor: 7 posts
- environmental factor: 18 posts
- lifestyle factor: 20 posts
- physiological factor: 28 posts
- psychological factor: 5 posts
- symptom: 36 posts
- gastrointestinal manifestation: 12 posts
- extra-gastrointestinal manifestation: 3 posts
- treatment method: 50 posts
- alternative therapy: 7 posts
- holistic treatment: 12 posts
- medication: 28 posts
- surgery: 13 posts
CCFA Explorer

Ontology Visualization

Topic Overview

Topic Details

3 linked views

D3
React/Flux
Bootstrap
Ontology Visualization

Ontology-based interactive visualization of patient-generated research questions: CCFA Explorer
Ontology Visualization

• **Force-directed network**
  - **Node size**
    • Number of topics labeled with term
    • Labeled with child implies labeled with parent
  - **Links**
    • Class relationships
    • Currently only “is a”
Ontology-based interactive visualization of patient-generated research questions: CCFA Explorer
Topic Overview

• **t-SNE Layout**
  – Based on labeled ontology terms
    • Clusters of topics with similar ontology terms

• **Glyph**
  – Size
    • Number of comments
  – Border width
    • Number of votes
  – Color
    • Category

t-SNE Modifications

1. Force-directed layout of overlapping glyphs

t-SNE Modifications

2. Differential weighting of higher-level (left) vs. lower level (right) ontology terms
t-SNE Modifications


3. Emphasizing selected ontology terms for layout
Ontology-based interactive visualization of patient-generated research questions: CCFA Explorer
Topic Details

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

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

**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.

Tags: treatment method, IBD course, symptom, surgery, diagnosis/monitoring method, diagnosis event, pre-diagnosis time period
**Topic Details**

<table>
<thead>
<tr>
<th>ID: 83</th>
<th>User: 2560</th>
<th>Votes: 3</th>
<th>Comments: 3</th>
<th>Category: procedures and testing</th>
</tr>
</thead>
</table>

**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?
Topic Details

- **List**
  - Sort
    - Topic ID
    - User ID
    - Votes
    - Comments
    - Category
  - Filter
    - Selected topics
    - Selected ontology terms
  - Highlight
    - Text from search box in red
Interactive Selection and Highlighting

- **Selection**
  - Ontology terms
  - Research topics
  - Selection in any view updates all views

- **Relationships**
  - Between ontology terms
    - Number of topics in common
    - Union for multiple
  - Between research topics
    - Number of terms in common
    - Union for multiple
  - Between terms and topics
    - Topic labeled with term or not
    - Sum for multiple
Interactive Selection and Highlighting

- **Selection**
  - Ontology terms
  - Research topics
  - Selection in any view updates all views

- **Relationships**
  - Co-occurrence
    - Between ontology terms
    - Number of topics in common
    - Union for multiple
  - Association
    - Between research topics
    - Number of terms in common
    - Union for multiple
  - Connection
    - Between terms and topics
    - Topic labeled with term or not
    - Sum for multiple
Interactive Selection and Highlighting

- **Selection**
  - Ontology terms
  - Research topics
  - Selection in any view updates all views
- **Relationships**
  - Co-occurrence
    - Between ontology terms
    - Number of topics in common
    - Union for multiple
  - Association
    - Between research topics
    - Number of terms in common
    - Union for multiple
  - Connection
    - Between terms and topics
    - Topic labeled with term or not
    - Sum for multiple
Interactive Selection and Highlighting

• Selection
  – Ontology terms
  – Research topics
  – Selection in any view updates all views

• Relationships
  – Co-occurrence
    • Between ontology terms
    • Number of topics in common
    • Union for multiple
  – Association
    • Between research topics
    • Number of terms in common
    • Union for multiple
  – Connection
    • Between terms and topics
    • Topic labeled with term or not
    • Sum for multiple

Select Term 3 and Term 4
Interactive Selection and Highlighting

- **Selection**
  - Ontology terms
  - Research topics
  - Selection in any view updates all views

- **Relationships**
  - Co-occurrence
    - Between *ontology terms*
    - Number of topics in common
    - Union for multiple
  - Association
    - Between *research topics*
    - Number of terms in common
    - Union for multiple
  - Connection
    - Between *terms and topics*
    - Topic labeled with term or not
    - Sum for multiple

Select *Term 3* and *Term 4*

Co-occurrence:
- Term 1: 1
- Term 2: 2
- Term 3: 2
- Term 4: 1
Interactive Selection and Highlighting

• **Selection**
  – Ontology terms
  – Research topics
  – Selection in any view updates all views

• **Relationships**
  – Co-occurrence
    • Between *ontology terms*
    • Number of topics in common
    • Union for multiple
  – Association
    • Between *research topics*
    • Number of terms in common
    • Union for multiple
  – Connection
    • Between *terms and topics*
    • Topic labeled with term or not
    • Sum for multiple

Select *Term 3* and *Term 4*

- **Co-occurrence:**
  - Term 1: 1
  - Term 2: 2
  - Term 3: 2
  - Term 4: 1

- **Connection:**
  - Topic 1: 1
  - Topic 2: 0
  - Topic 3: 2
Interactive Selection and Highlighting

- **Selection**
  - Ontology terms
  - Research topics
  - Selection in any view updates all views

- **Relationships**
  - Co-occurrence
    - Between ontology terms
    - Number of topics in common
    - Union for multiple
  - Association
    - Between research topics
    - Number of terms in common
    - Union for multiple
  - Connection
    - Between terms and topics
    - Topic labeled with term or not
    - Sum for multiple

Select *Topic 2 and Topic 3*
Interactive Selection and Highlighting

- **Selection**
  - Ontology terms
  - Research topics
  - Selection in any view updates all views

- **Relationships**
  - Co-occurrence
    - Between *ontology terms*
    - Number of topics in common
    - Union for multiple
  - Association
    - Between *research topics*
    - Number of terms in common
    - Union for multiple
  - Connection
    - Between *terms and topics*
    - Topic labeled with term or not
    - Sum for multiple

Select *Topic 2* and *Topic 3*

**Association:**
- Topic 1: 3
- Topic 2: 1
- Topic 3: 3
Interactive Selection and Highlighting

- **Selection**
  - Ontology terms
  - Research topics
  - Selection in any view updates all views

- **Relationships**
  - Co-occurrence
    - Between ontology terms
    - Number of topics in common
    - Union for multiple
  - Association
    - Between research topics
    - Number of terms in common
    - Union for multiple
  - Connection
    - Between terms and topics
    - Topic labeled with term or not
    - Sum for multiple

Select *Topic 2* and *Topic 3*

**Association:**
- Topic 1: 3
- Topic 2: 1
- Topic 3: 3

**Connection:**
- Term 1: 1
- Term 2: 1
- Term 3: 1
- Term 4: 1
Ontology Highlighting

Node interior size and color proportional to co-occurrence/connection

Labels shown if > 25%

Ontology-based interactive visualization of patient-generated research questions: CCFA Explorer
Topic Overview Highlighting

Glyph saturation proportional to association/connection

drug

Ontology-based interactive visualization of patient-generated research questions: CCFA Explorer
Topic Overview Highlighting

Glyph saturation proportional to association/connection

Ontology-based interactive visualization of patient-generated research questions: CCFA Explorer
Topic Overview Highlighting

Glyph saturation proportional to association/connection

Ontology-based interactive visualization of patient-generated research questions: CCFA Explorer
Ontology-based interactive visualization of patient-generated research questions: CCFA Explorer

Ontology terms colored as in ontology network visualization
Example Scenario

- User searches for *genetic makeup*
  - **Highlighted in red**
- User selects *genetic makeup*
  - Notices *medication* co-occurs with *genetic makeup*
- User select *medication* also
Example Scenario

- User re-runs t-SNE
  - Notices cluster of three topics, including one very large glyph
- User selects these three topics
Example Scenario

- User filters topic details to show only these topics for in-depth inspection/comparison
Feedback and Future Work

• **Demonstration to IBD Partners**
  - Generally positive feedback
    • Useful way to explore the forum data
    • Quickly identify major themes and popular research topics
    • Some features may be too complex for more naïve users
  - Two particular themes identified
    • Patient-facing interface
      – Find similar patients
      – More easily identify research topics relevant to them
    • Researcher-facing interface
      – Quickly identify relevant information in their research area
      – Quickly generate summaries of relevant information
      – Easily presented to stakeholders
Feedback and Future Work

- **Interface**
  - Redesign ontology visualization
    - Improve navigation
  - Explore automatic text summarization
    - Include in summary panel for selected terms/topics

- **Data processing**
  - Explore ways to automate/semi-automate topic classification
  - Enable ontology editing/expansion

- **User evaluations**
Feedback and Future Work

• Interface
  – Redesign ontology visualization
  – Improve navigation
  – Explore automatic text summarization
    • Include in summary panel for selected terms/topics

• Data processing
  – Explore ways to automate/semi-automate topic classification
  – Enable ontology editing/expansion
Feedback and Future Work

- **Interface**
  - Redesign ontology visualization
    - Improve navigation
  - Explore automatic text summarization
    - Include in summary panel for selected terms/topics

- **Data processing**
  - Explore ways to automate/semi-automate topic classification
  - Enable ontology editing/expansion

- **User evaluations**
Questions?

Thanks to IBD Partners for their help with this work!