Collaborative Sensemaking in Online Knowledge Groups

Pengyi Zhang, Zhe Wang, and Yuyu Yang
Department of Information Management
Peking University
Motivation

“This information seems relevant, but what does it mean?”

Finding information → Making sense of it

Beyond information retrieval, supporting users in making sense of the information found is the next frontier in information research.
What is Sensemaking

• Sensemaking is the task of creating an understanding of a problem or task so that further actions may be taken in an informed manner. (Stefik, 1999)
  - An important part of sensemaking involves making clear the interrelated concepts and their relationships in a problem or task space

• Sensemaking is particularly important in
  - New situations
  - Less structured tasks
  - Collaborative settings
Individual sensemaking

Interrelated concepts and their connections

Task/ Problem
Collaborative sensemaking

Interactions

Negotiation

Feedback

Interactions
Sensemaking Elements

Processes

Searching

Making sense

Activities

Identification of gaps

Search
- Exploratory search
- Focused search for data

Building structure

Instantiating structure

Consuming instantiated structure

Mechanisms

Data-driven, inductive
- Key item extraction
- Schema induction
- Generalization

Structure-driven, logic-driven
- Definition
- Specification
- Elimination
- Explanation
- Inference

Other
- Comparison
- Analogy
- Classification
- Semantic fit
- Socratic dialogues

Outcomes

Accretion

Tuning

Restructuring

01/19/2017
Building and Instantiating Structure

- **Building structure**: creating structural representations of knowledge, for example:
  - Organizing concepts and relationships in a concept map

- **Instantiating structure**: fitting data into the structure created, for example:
  - Linking facts to a concept in a concept map

(Russell, 1993)
Research Question

• “How do people build knowledge structures and instantiate the structures with data collectively in online knowledge groups?”
  • What activities are involved and how they proceed in sequence?
  • How are conflicts, inconsistency, and abnormality handled?
  • What are the roles and division of labor in coordinating group sensemaking?
# Research Design

<table>
<thead>
<tr>
<th></th>
<th><strong>Study I</strong>: Topic hierarchy construction on a social QA site</th>
<th><strong>Study II</strong>: Topic discussion in an online course group wiki</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sensemaking focus</strong></td>
<td>Structure building</td>
<td>Structure building and instantiation</td>
</tr>
<tr>
<td><strong>Task</strong></td>
<td>Constructing a topic hierarchy</td>
<td>Writing a report about certain topic</td>
</tr>
<tr>
<td><strong>User group</strong></td>
<td>Online QA community (zhihu.com)</td>
<td>Student study group</td>
</tr>
<tr>
<td><strong>Group size</strong></td>
<td>Medium to large</td>
<td>Small</td>
</tr>
<tr>
<td><strong>Nature of data collected</strong></td>
<td>Quantitative log data, including:</td>
<td>Qualitative data including:</td>
</tr>
<tr>
<td></td>
<td>- Topic hierarchies of selected areas</td>
<td>- Offline discussion recordings</td>
</tr>
<tr>
<td></td>
<td>- Topic history log recording:</td>
<td>- Wiki editing history</td>
</tr>
<tr>
<td></td>
<td>- addition, deletion, modification of relationships, username, and a time stamp</td>
<td>- Final report</td>
</tr>
<tr>
<td></td>
<td>- Topic information (questions, answers, related user activities)</td>
<td>- Other communications including emails, IM messages, etc.</td>
</tr>
<tr>
<td></td>
<td>- User profiles</td>
<td></td>
</tr>
</tbody>
</table>
STUDY I: Social Creation of a Topic Hierarchy

- Disciplines (5)
  - Social sciences (23)
    - Politics (8)
      - Political sociology (13)
      - Elitism (0)
      - Social class (13)
      - View all
    - Institution (5)
    - View all
  - Journalism and Communication (2)
    - Communication (5)
    - Journalism (9)
  - Sociology (26)
    - Political sociology (13)
    - Ethnosociology (7)
    - View all
  - Economics (50)
  - View all

- Natural sciences (9)
  - Chemistry (36)
  - Environmental science (1)
  - Biology (33)
  - View all

- Formal sciences (4)
  - Logic (4)
  - Theoretical computer science (7)
  - Statistics (32)
  - Math (59)

- Humanities (16)
  - Arts (2)
  - Filmology (0)
  - View all

- Technology and applied sciences (31)
  - Optics (14)
  - Medical science (35)
  - View all

Part of the Zhihu Topic Hierarchy (Translated from Chinese)
## Types of Structure Building Activity

<table>
<thead>
<tr>
<th>Concept edits</th>
<th>Creating a concept</th>
<th>Creating topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revising a concept</td>
<td>Editing topic description</td>
<td>Editing topic name</td>
</tr>
<tr>
<td>Editing synonym relationship</td>
<td>Adding topic alias</td>
<td>Removing topic alias</td>
</tr>
<tr>
<td>Editing hierachal relationship</td>
<td>Adding parent topic</td>
<td>Removing parent topic</td>
</tr>
<tr>
<td>Editing weak relationship</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#LogID</th>
<th>Username</th>
<th>Action</th>
<th>Topic 1</th>
<th>Topic 2</th>
<th>Date</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>378</td>
<td>胡**</td>
<td>Creating topic</td>
<td>Mobile Internet</td>
<td></td>
<td>2010/12/20</td>
<td>8:25:44</td>
</tr>
<tr>
<td>91672</td>
<td>V**</td>
<td>Editing topic description</td>
<td>Mobile Internet</td>
<td>Mobile Internet</td>
<td>2011/03/18</td>
<td>13:22:14</td>
</tr>
<tr>
<td>1091333</td>
<td>奚**</td>
<td>Adding topic alias</td>
<td>Mobile Internet</td>
<td>Wireless Internet</td>
<td>2011/09/08</td>
<td>17:31:28</td>
</tr>
<tr>
<td>4203756</td>
<td>陈**</td>
<td>Adding parent topic</td>
<td>Mobile Internet</td>
<td>Windows Phone 7</td>
<td>2012/09/18</td>
<td>9:50:46</td>
</tr>
<tr>
<td>4203758</td>
<td>陈**</td>
<td>Removing parent topic</td>
<td>Mobile Internet</td>
<td>Windows Phone 7</td>
<td>2012/09/18</td>
<td>9:50:51</td>
</tr>
<tr>
<td>4279564</td>
<td>张**</td>
<td>Editing topic picture</td>
<td>Mobile Internet</td>
<td></td>
<td>2012/09/25</td>
<td>20:58:20</td>
</tr>
</tbody>
</table>

Data set: 10,055 topics under “disciplines”; 80,927 edits by 8,694 users.
Types of Structure Building
Concept vs. Relationship Edits

![Graph showing the changing percentage of edits over time. The x-axis represents dates from 11/1/2010 to 11/1/2015, and the y-axis represents the percentage of edits. The graph uses different colors to represent different types of edits: Editing Hierarchal relationship, Editing synonym relationship, Revising a concept, and Creating a concept.]
## Conflicts in constructing the structure

<table>
<thead>
<tr>
<th># of relationships</th>
<th>28,283</th>
</tr>
</thead>
<tbody>
<tr>
<td>No conflicts</td>
<td></td>
</tr>
<tr>
<td>Edited once</td>
<td>19261 (68.1%)</td>
</tr>
<tr>
<td>Edited more than once by the same person</td>
<td>2263 (8%)</td>
</tr>
<tr>
<td>Conflicts</td>
<td>6,759 (23.9%)</td>
</tr>
<tr>
<td>Removed</td>
<td>6123 (90.6%)</td>
</tr>
<tr>
<td>Kept</td>
<td>636 (9.4%)</td>
</tr>
</tbody>
</table>

No conflicts edited once.
User participation and roles

(a) distribution of users' edits;

(b) distribution of user edited topics
STUDY II: Small Group Collaboration

Sensemaking

- Information seeking
- Individual outline
- Group discussion
- Writing report collaboratively

2 weeks

6 small groups
6 topics
36 group discussion recordings
36 reports on wiki
Analysis of discussion recordings

<table>
<thead>
<tr>
<th>Group</th>
<th># of Participants</th>
<th>Length (min)</th>
<th># of Sub-sessions</th>
<th># of Participants</th>
<th>Length (min)</th>
<th># of Sub-sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>1</td>
<td>11</td>
<td>94</td>
<td>5</td>
<td>10</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>12</td>
<td>64</td>
<td>5</td>
<td>11</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>11</td>
<td>69</td>
<td>5</td>
<td>11</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>12</td>
<td>71</td>
<td>5</td>
<td>12</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>10</td>
<td>62</td>
<td>5</td>
<td>10</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>10</td>
<td>47</td>
<td>5</td>
<td>10</td>
<td>19</td>
</tr>
<tr>
<td>Average</td>
<td>11</td>
<td>68</td>
<td>5</td>
<td>11</td>
<td>34</td>
<td>4</td>
</tr>
</tbody>
</table>

**Topic 1:** FRBR  
**Topic 2:** Folksonomy

<table>
<thead>
<tr>
<th>Label</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Questioning: Asking for information; challenging someone's opinion</td>
</tr>
<tr>
<td>A2</td>
<td>Statement: Expressing opinions or attitudes towards facts or opinions</td>
</tr>
<tr>
<td>A3</td>
<td>Summarizing: Summarizing the content of previous discussion</td>
</tr>
<tr>
<td>A4</td>
<td>Illustrating: Providing examples in order to make opinions understandable or valid</td>
</tr>
<tr>
<td>A5</td>
<td>Verification: Making sure a former statement is true or correct</td>
</tr>
<tr>
<td>A6</td>
<td>Clarification: Giving further description or explanation to avoid confusion</td>
</tr>
<tr>
<td>A7</td>
<td>Explanation: Explaining reason or cause of something</td>
</tr>
<tr>
<td>A8</td>
<td>Definition: Defining a concept using technical terms</td>
</tr>
<tr>
<td>A9</td>
<td>Judgment-Evaluation: Judging or evaluating the quality of someone’s statement or opinion</td>
</tr>
<tr>
<td>A10</td>
<td>Refuse to participate: Refusing to take part in sensemaking process when prompted</td>
</tr>
<tr>
<td>A11</td>
<td>Prompt: Moving the discussion forward by encouraging someone to talk</td>
</tr>
<tr>
<td>A12</td>
<td>Search-Encounter: Assisting the process of sensemaking by finding data or proof</td>
</tr>
<tr>
<td>A13</td>
<td>Digress: Speaking about something different from the main subject discussed</td>
</tr>
</tbody>
</table>
Sensemaking Activity Patterns

Topic 1

01/19/2017
• Verification and Clarification always appear in pairs
  • A1: You mean it has no vocabulary control? [Verification]
  • A4: Yes, exactly. So I mean, without vocabulary control these tags are useless, which will result in inefficiency of the folksonomy. [Clarification]

• Questioning, Statement, Verification, Clarification and Summarizing often take place in sequence
  • C1: What pieces could be classified into new work or new expression in this graph? And why? [Questioning]
  • C7: I think a novel and a film with subtitles are new works, and the others are new expressions. [Statement]
  • C1: Really? [Verification]
  • C7: Changes in authorship and content is my standard. [Clarification]
  • C12: Well, so actually the novel is a new work while the translation version of a novel is the same work. [Summarizing]
## Conflicts and agreement

<table>
<thead>
<tr>
<th>Activity</th>
<th>Agreement</th>
<th>Conflict</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Statement</td>
<td>79</td>
<td>61%</td>
<td>50</td>
</tr>
<tr>
<td>Questioning</td>
<td>26</td>
<td>68%</td>
<td>12</td>
</tr>
<tr>
<td>Illustrating</td>
<td>16</td>
<td>76%</td>
<td>5</td>
</tr>
<tr>
<td>Summarizing</td>
<td>14</td>
<td>93%</td>
<td>1</td>
</tr>
<tr>
<td>Prompt</td>
<td>9</td>
<td>64%</td>
<td>5</td>
</tr>
<tr>
<td>Judgment-Evaluation</td>
<td>8</td>
<td>100%</td>
<td>0</td>
</tr>
<tr>
<td>Verification</td>
<td>7</td>
<td>54%</td>
<td>6</td>
</tr>
<tr>
<td>Clarification</td>
<td>5</td>
<td>45%</td>
<td>6</td>
</tr>
<tr>
<td>Explanation</td>
<td>5</td>
<td>56%</td>
<td>4</td>
</tr>
<tr>
<td>Search-Encounter</td>
<td>2</td>
<td>100%</td>
<td>0</td>
</tr>
<tr>
<td>Digress</td>
<td>1</td>
<td>50%</td>
<td>1</td>
</tr>
<tr>
<td>Definition</td>
<td>1</td>
<td>100%</td>
<td>0</td>
</tr>
<tr>
<td>Refuse to participate</td>
<td>0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>173</td>
<td></td>
<td>90</td>
</tr>
</tbody>
</table>
Dealing with conflicts

• Agreement => Reinforcement
  • “Of course, that’s what I thought.”

• Compensatory => Acceptance
  • “I did not know that... It is good to know.”

• In-conflict-with => sometimes persuasion, but more often doubts and confusion
  • I am not sure about this. Let’s check with the teacher/TA
  • This is confusing. Let’s come back to it later.
Summary

• Process:
  • At the beginning: of a collaborative sensemaking process to build structures, more attention is to identify the concepts involved in the task/problem, whereas later on focus shifts to examining relationships
  • Smaller groups have different patterns of combination of sensemaking activities that moves the process along
  • Some activities are more likely to proceed in sequence (for example, statement-verification)

• Dealing with conflicts:
  • In zhihu, Conflicts are mostly like to result in removal of a relationship under discussion (90.6%)
  • In smaller groups, there are more opportunity to solve disagreement but very often it lead to confusion or failure in sensemaking
Implications

• In collaborative information seeking and sensemaking tasks:
  • It is important to be able to share “senses made” (knowledge structure instantiated with data) among group members in addition to share information found

• For collaborative systems:
  • The ability to show and communicate about the differences and disagreements might lead to more successful sensemaking
Ongoing and future directions

• Examining user roles in both studies
  • STUDY I:
    • domain experts vs. classification experts
  • STUDY II: (looking at both discussion and writing)
    • Content:
      • Building structure (theorists)
      • Instantiating structure
    • Process: facilitators
    • Leader vs. follower
Thank you!

- This research is supported by the OCLC/ALISE Research Grant. Thank you for this support.

- Pengyi Zhang:
  - Email: pengyi@pku.edu.cn
  - Website: pengyi.kvlab.org