Designing Explanation Interfaces for Transparency and Beyond

Chun-Hua Tsai
University of Pittsburgh
Pittsburgh, USA
cht77@pitt.edu

Peter Brusilovsky
University of Pittsburgh
Pittsburgh, USA
peterb@pitt.edu

ABSTRACT

In this work-in-progress paper, we presented a participatory process of designing explanation interfaces for a social recommender system with multiple explanatory goals. We went through four stages to identify the key components of the recommendation model, expert mental model, user mental model, and target mental model. We reported the results of an online survey of current system users (N=14) and a controlled user study with a group of target users (N=15). Based on the findings, we proposed five set of explanation interfaces for five recommendation models (N=25) and discussed the user preference of the interface prototypes.

CCS CONCEPTS

• **Information systems → Recommender systems**; • **Human-centered computing → HCI design and evaluation methods**.

KEYWORDS

Social Recommendation; Explanation; Mental Model; User Interface

1 INTRODUCTION

Enhancing explainability in recommender systems has drawn more and more attention in the field of Human-Computer Interaction (HCI). Further, the newly initiated European Union's General Data Protection Regulation (GDPR) required the owner of any data-driven application to maintain a "right to the explanation" of algorithmic decisions [1], which urging to gain *transparency* in all existing intelligent systems. Self-explainable recommender systems have been proved to gain user perception on system transparency [17], trust [13] and accepting the system suggestions [7]. Instead of the offline performance improvements, more and more researches focused on the works of evaluating the system from the *user experience*, i.e., what is the user perception on the explanation interfaces?

Explaining recommendations (i.e., enhancing the system explainability) can achieve different *explanatory goals* which help users to make a better decision or persuading them to accept the suggestions from a system [14, 16]. We followed the seven explanatory goals that proposed by Tintarev and Masthoff [17]: *Transparency*, *Scrutability*, *Trust*, *Persuasiveness*, *Effectiveness*, *Efficiency*, and *Satisfaction*. Since it is hard to have a single explanation interface that achieves all these goals equally well, the designer needs to make a trade-off while choosing or designing the form of interface [17]. For instance, an interactive interface can be adapted to increase the user trust and satisfaction but may prolong the decision and explore process while using the system (i.e., lead to decreasing of efficiency) [19].

Over the past few years, several approaches have been discussed to enhance the explainability in the recommender systems. The approaches can be summarized by different styles, reasoning models, paradigms and information [2]. 1) *Styles:* Kouki et al. [8] conducted an online user survey to explore the user preference in nine explanation styles. They found *Venn diagrams* outperformed all other visual and text-based interfaces. 2) *Reasoning Models:* Vig et al. [24] used tags to explain the recommended item and the user's profile. The approach emphasized the factor of why a specific recommendation is plausible, instead of revealing the process of recommendation or data. *3) Paradigms:* Herlocker et al. [5] presented a model for explanations based on the user's conceptual model of the collaborative-based recommendation process. The result of the evaluation indicates two interfaces - "Histogram with grouping" and "Presenting past performance" - improved the acceptance of recommendations. 4) *Information:* Pu and Chen [13] proposed explanations tailored to the user and recommendation, i.e., although one recommendation is not the most popular one, the explanation would justify the recommendation by providing the reasons.

Although many approaches have been proposed to enhance the recommender explainability, bringing explanation interfaces to an existing recommender system is still a challenging task. More recently Eiband et al. [1] suggested a different approach to improve *user mental model (UMM)* while bringing transparency (explanations) to a recommender system. The model described the process of a user builds an internal conceptualization of the system or interface along with user-system interactions, i.e., building the knowledge of how to interact with the system. If the model is misguided or opaque, the users will face difficulties in predicting or interpreting the system [1]. Hence, the researchers suggested to improve the *mental model*, so the users can gain *awareness* while using the system as well as the explanation interfaces.

In this work-in-progress paper, we presented a stage-based participatory process [1] for integrating seven exploratory goals into real-world hybrid social recommender system. First, we introduced the *Expert Mental Model* to summarize the key components of each recommendation feature. Second, we conducted an online survey to identify the *User Mental Model* of seven explanatory goals from the current system users. Third, we did a user study with card-sorting and semi-interview to determine the user's *Target User Model.* Fourth, we proposed a total of 25 explanation interfaces for five recommendation features and compared the user perceptions across designs.
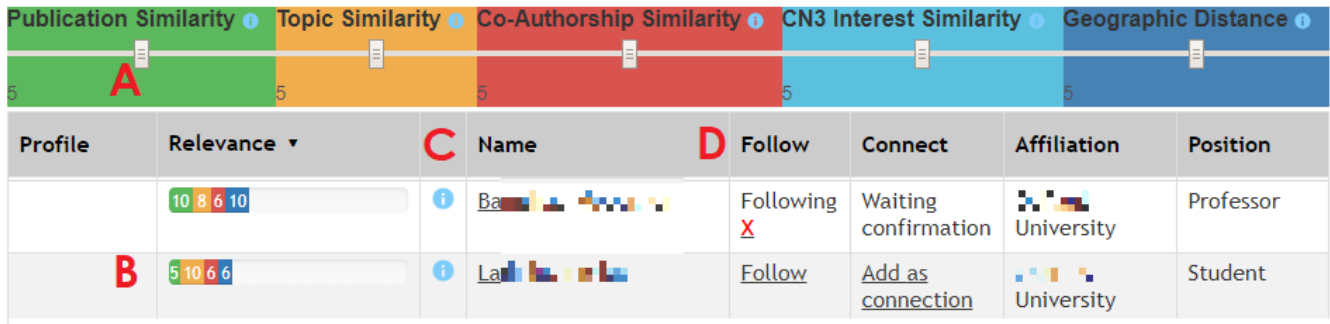
Figure 1: Relevance Tuner+: (A) relevance sliders; (B) stackable score bar; (C) explanation icon; (D) user profiles. The interface supports the user-driven exploration of recommended items in Section A and inspects the fusion in Section B. The user can further inspect the explanation model by clicking Section C, and more profile detail is presented in section D. Our goal is to provide an explanation interfaces for each explanation model. (The scholar names have been pixelated for privacy protection)

2 BACKGROUND

We adopted the stage-based participatory framework from Eiband et al. [1], which intends to answer two key questions while designing the explainable user interface (UI): a) What to Explain? And b) How to explain? The process can be summarized in four stages. *1) Expert Mental Model:* What can be explained? We defined an expert as the recommender system developer. *2) User Mental Model:* What is the user mental model of the system based on its current UI? The model should be built through the current recommender system users. *3) Target Mental Model:* Which key components of the algorithm do users want to be made explainable in the UI? The *target user* is the users who are new to the system. 4) *Iterative Prototyping:* How can the target mental model be reached through UI design. The key is to measure if the proposed explanation interfaces achieved the explanatory goals.

In this work, we aimed to enhance the explainability in a conference support system - *Conference Navigator 3 (CN3)*. The system has been used to support more than 45 conferences at the time of writing this paper and has data on approximately 7,045 articles presented at these conferences; 13,055 authors; 7,407 attendees; 32,461 bookmarks; and 1,565 social connections. Our work was informed by the results of a controlled user study where we explored an earlier version of the social recommender interface *Relevance Tuner* [19] (shown in Figure 1). It was a controllable interface for the user to fuse weightings of multiple recommendation models and to inspect the explanations.

A total of five recommendation models were introduced in this study: *1) Publication Similarity:* the degree of cosine similarity of users' publication text. *2) Topic Similarity:* the overlap of research interests (using topic modeling). *3) Co-Authorship Similarity:* the degree of connection, based on a shared network of co-authors. *4) Interest Similarity:* the number of papers co-bookmarked, as well as the authors co-followed. *5) Geographic Distance:* a measurement of the geographic distance between affiliations. Based on the stage-based participatory framework, we went through the same four stages for each recommendation model to identify the user-preferred user interface design. We aimed to design explanation interfaces for each recommendation model with multiple exploratory goals.

3 FIRST STAGE: EXPERT MENTAL MODEL

Instead of interactive recommender [7, 23], we attached an *explanation icon* next to each social recommendation. The users have a choice of requesting the explanations while exploring or browsing the recommendations. We adopted a hybrid explanation approach [8, 12], which mixed multiple visualizations to explain the details of the recommendation model. We would like to let the users understand both a) the mutual relationship (similarity) between him/herself and the recommended scholar and b) the key component in each recommendation model. We then discussed the *Expert Mental Model* through the system developing process of the five recommendation models.

1) Publication Similarity: The similarity was determined by the degree of text similarity between two scholars' publications using cosine similarity. We applied tf-idf to create the vector with a word frequency upper bound of 0.5 and a lower bound of 0.01 to eliminate both common and rarely used words. In this model, the key components were the *terms* of the paper title and abstract as well as its *term frequency*.

2) Topic Similarity: This similarity was determined by matching research interests using topic modeling. We used latent Dirichlet allocation (LDA) to attribute collected terms from publications to one of the topics. We chose 30 topics to build the topic model for all scholars. Based on the model, we then calculated the topic similarity between any two scholars. The key components were the *research topics* and the *topical words* of each research topic [25].

3) Co-Authorship Similarity: This similarity approximated the network distance between the source and recommended users. For each pair of the scholar, we tried to find six possible paths for connecting them, based on their coauthorship relationships. The network distance is determined by the average distance of the six paths. The key components were the *coauthors* (as nodes), *coauthorship* (as edges) and the *distance of connection the two scholars*.

4) CN3 Interest Similarity: This similarity was determined by the number of co-bookmarked conference papers and co-connected authors in the experimental social system (CN3). We simply used the number of shared items as the CN3 interest similarity. The key component is the shared *conference papers* and authors.

5) Geographic Distance: This similarity was a measurement of the geographic distance between attendees. We retrieved longitude and latitude data based on attendees' affiliation information. We used the Haversine formula to compute the geographic distance between scholars. The key components are the *geographic distance* and *affiliation information of the scholars.*

4 SECOND STAGE: USER MENTAL MODEL

As a first step towards understanding the design factors of explanatory interfaces, we deployed a survey through a social recommender system, Conference Navigator [18], and analyzed data from the respondents. We targeted the users who had created an account and interacted with the system in their previous conference attendance (at least using the system for one conference). The survey was initiated by sending an invitation to the qualified users in December 2017. We sent out 89 letters to the conference attendees of UMAP/HT 2016, and a total of 14 participants (7 female) replied to create the pool of participants for the user study. The participants were from 13 different countries; their ages ranged from 20 to 40 (M=31.36, SE=5.04). We did an online survey to collect necessary demographic information and self-reflection about how to design an explanation function in seven explanatory goals [17].

The proposed questions were: *How can an explanation function help you to perceive system* **1) Transparency** - explain how the system works? **2) Scrutability** - allow you to tell the system it is wrong? **3) Trust** - increase your confidence in the system? **4) Persuasiveness** - convince you to explore or to follow new friends? **5) Effectiveness** - help you make good decisions? **6) Efficiency** - help you to make decisions faster? **7) Satisfaction** - make using the system fun and useful? We asked the participants to answer each question in 50-100 words, in particular reflecting the explanatory goals of the social recommendation. The data was published in [20].

1) Transparency: 71% of respondents pointed out the *reasons* of generated social recommendation that help them to perceive higher system transparency, i.e., the personalized explanation, the linkage and data sources, reasoning method and understandability. We then summarized the feedback into five factors: *1) The visualization presents the similarity between my interest and the recommended person. 2) The visualization presents the relationship between the recommended person and me. 3) The visualization presents where did the data were retrieved. 4) The visualization presents more in-depth information on how the score amounts up. 5) The visualization allows me to see the connections between people and understand how they are connected.*

2) Scrutability: Half of the respondents mentioned they needed "inspectable details" to figure out the wrong recommendation. 35% of respondents suggested the mechanism of accepting user feedback on improving wrong recommendations, such as a space to submit user ratings or yes/no options. 14% of respondents preferred a dynamic exploration process to determine the recommendation quality. We then summarized the feedback into four factors: *6) The visualization allows me to understand whether the recommendation is good or not. 7) The visualization presents the data for making the recommendations. 8) The visualization allows me to compare and decide whether the system is correct or wrong. 9) The visualization allows me to explore and then determine the recommendation quality.*

3) Trust: 28% of respondents mentioned that they trusted the system more when they perceived the benefits of using the system. 35% of respondents preferred to trust a system with reliable and informative explanations, more detailed information or understandable. 35% of respondents mentioned they trust a system with transparency or passed their verification. We then summarized the feedback into three factors: *10) The visualization presents a convincing explanation to justify the recommendation. 11) The visualization presents the components (e.g., algorithm) that influenced the recommendation. 5) The visualization allows me to see the connections between people and understand how they are connected.*

4) Persuasiveness: Half of the respondents mentioned the explanation of social familiarity would persuade them to explore novel social connections; namely, when shown social context details or shared interests. 21% of respondents indicated that an informative interface could boost the exploration of new friendship. 28% of respondents preferred a design that inspired curiosity, implicit relationships. We then summarized the feedback into three factors: *12) The visualization shows me the shared interests, i.e., why my interests are aligned with the recommended person. 13) The visualization has a friendly, easy-to-use interface. 14) The visualization inspired my curiosity (to discover more information).*

5) Effectiveness: 64% of respondents mentioned that the aspects of social recommendation relevance helped them to make a good decision. The aspect included explaining the recommendation process, understandable or more informative. 28% of respondents suggested a reminder that a historical or successful decision could help them to make a good decision, i.e., a previously-made user decision and success stories. We then summarized the feedback into three factors: *15) The visualization presents the recommendation process. 5) The visualization allows me to see the connections between people and understand how they are connected. 11) The visualization presents the components (e.g., algorithm) that influenced the recommendation.*

6) Efficiency: 28% of respondents mentioned that a proper highlighting of the recommendation helped to make the decision faster. For example, they are emphasizing the relatedness, identifying the top recommendations or providing success stories. 28% of respondents preferred a tune-able or visualized interface to accelerate the decision process, such as tuning the recommendation features, visualizing the recommendations. However, the explanations may not always be useful. 21% of respondents argued that the explanation would prolong the decision process instead of speeding it up: the user may need to take extra time to examine the explanations. We then summarized the feedback into two factors: *16) The visualization presents highlighted items/information that is strongly related to me. 17) The visualization presents aggregated, non-obvious relations to me.*

7) Satisfaction: The feedback on how an explanation can help the user satisfy the system was varied. Three aspects received an equal 7% of respondents' preferences. That is, users preferred to view the feedback from the community, shown the historical interaction record and provided a personalized explanation. Two aspects received an equal 14% of respondents' preference; i.e., a focus on a friendly user interface and saved decision time. 21% of respondents reported a higher satisfaction on using the explanation as a "small talk topic", i.e., as an initial conversation in a conference. 28% of

Table 1: The card-sorting results of the third stage.

	Very Important	Less Important	Not Important	Not Relevant
Factor 1	11	1	3	0
Factor 2	9	5	1	0
Factor 3	0	2	10	3
Factor 4	1	8	3	3
Factor 5	5	4	6	0
Factor 6	7	6	2	0
Factor 7	3	2	9	1
Factor 8	4	3	3	5
Factor 9	7	2	4	2
Factor 10	3	9	2	1
Factor 11	0	6	6	3
Factor 12	4	6	5	0
Factor 13	13	2	0	0
Factor 14	0	13	2	0
Factor 15	4	7	3	1
Factor 16	10	5	0	0
Factor 17	3	6	3	3
Factor 18	1	5	5	4
Factor 19	1	10	3	1

respondents preferred an interactive interface for perceiving the system to be fun, e.g., a controllable interface. We then summarized the feedback into four factors: *18) The visualization presents the feedback from other users, i.e., I can see how others rated the recommended person. 19) The visualization allows me to tell why does this system recommend the person to me. 1) The visualization presents the similarity between my interest and the recommended person. 13) The visualization is a friendly, easy-to-use interface.*

Based on the result of the online survey, we concluded a total of 19 factors in the second stage of building the user mental model.

(1) The visualization presents the similarity between my interest and the recommended person.
(2) The visualization presents the relationship between the recommended person and me.
(3) The visualization presents where the data was retrieved.
(4) The visualization presents more in-depth information on how the scores sum up.
(5) The visualization allows me to see the connections between people and understand how they are connected.
(6) The visualization allows me to understand whether the recommendation is good or not.
(7) The visualization presents the data for making the recommendations.
(8) The visualization allows me to compare and decide whether the system is correct or wrong.
(9) The visualization allows me to explore and then determine the recommendation quality.
(10) The visualization presents a convincing explanation to justify the recommendation.
(11) The visualization presents the components (e.g., algorithm) that influenced the recommendation.

(12) The visualization shows me the shared interests, i.e., why my interests are aligned with the recommended person.
(13) The visualization has a friendly, easy-to-use interface
(14) The visualization inspired my curiosity (to discover more information).
(15) The visualization presents the recommendation process clearly.
(16) The visualization presents highlighted items/information that is strongly related to me.
(17) The visualization presents aggregated, non-obvious relations to me.
(18) The visualization presents feedback from other users, i.e., I can see how others rated a recommended person.
(19) The visualization allows me to tell why does this system recommend the person to me.

We also found some factors across different exploratory goals. For example, Factor 1 were shared by the exploratory goal of *Transparency* and *Satisfaction*. Factor 5 were shared by *Transparency*, *Trust* and *Effectiveness*. Factor 11 was shared by *Trust* and *Effectiveness*. Factor 13 was shared by *Persuasiveness* and *Satisfaction*.

5 THIRD STAGE: TARGET MENTAL MODEL

In this stage, we conducted a controlled lab study for creating the *Target Mental Model*. The model is used to identify the key components of the recommendation model that the users might want to be explainable in the UI. Since the goal is to identify the information need for new users, we specifically selected subjects who never used the CN3 system. A total of 15 (6 female) participants (N=15) were recruited for this study. They are first, or second-year graduate students (major in information sciences) at the University of Pittsburgh with age ranged from 20 to 30 (M=25.73, SE=2.89). All participants had no previous experience of using the CN system. Each participant received USD$20 compensation and signed an informed consent form.

We asked the subjects to complete a card-sorting task about their preference for the 19 factors we identified in the second stage. We started by presenting the CN3 system (shown in Figure 1) to the subjects and introducing the five recommendation models through the *Expert Mental Model*. After the tutorial, the subjects were asked to do a closed card-sorting that assigns cards into four predefined groups. The four groups are 1) very important; 2) less important; 3) not important and 4) not relevant.

The survey result is reported in Table 1. We found that for the target users, factor 1, 13, 16 outperformed other factors: more than ten subjects assigned the three factors into the "very important" group. The factor 2, 6, 10, 12, 14, 15 and 19 formed the secondary preference group with at least 10 subject assigning them into "very important" or "less important" groups. The subjects least preferred factor were 3, 7, 11, 18 with at least nine subjects assigning these factors into "not important" or "not relevant" groups.

Based on the card-sorting result, we found the user preferred an explainable UI is presenting the similarity between his/her interests and the recommended person (F1). The UI should be friendly and easy-to-use (F13) as well as highlighted the items or information that is strongly related to the user (F16). Besides, some factors are also liked by the subjects. For instance, the UI is presenting the mutual relationship (F2), shared interests (F12) and recommendation

Table 2: The card-sorting results of the fourth stage.

	R1	R2	R3	R4	R5	Not Applicable	Total Votes
E1-1	19	25	21	19	44	22	150
E1-2	23	37	17	30	26	17	150
E1-3	7	16	42	44	19	22	150
E1-4	**76**	32	27	2	0	13	150
E1-5	19	31	33	28	20	19	150
E2-1	12	8	14	21	60	35	150
E2-2	6	2	9	73	36	24	150
E2-3	24	78	28	7	2	11	150
E2-4	**86**	31	13	11	0	9	150
E2-5	13	21	70	14	11	21	150
E3-1	13	5	9	18	69	36	150
E3-2	37	26	17	36	20	14	150
E3-3	32	38	29	28	11	12	150
E3-4	**45**	41	37	11	0	16	150
E3-5	15	32	41	36	11	15	150
E4-1	8	11	6	31	64	30	150
E4-2	17	61	48	16	2	6	150
E4-3	49	41	41	11	3	5	150
E4-4	**64**	28	41	7	1	9	150
E4-5	8	5	6	65	46	20	150
E5-1	20	7	13	24	55	31	150
E5-2	16	22	6	45	36	25	150
E5-3	**42**	16	44	11	6	31	150
E5-4	15	49	36	18	4	28	150
E5-5	40	35	26	20	3	26	150

process (F15). The UI should also allow the user to understand (F6) and justify (F10) the quality of recommendation as well as inspired the curiosity of exploration (F14) and recommendation process (F19). Interestingly, we also found the users were less interested in a UI of presenting the data source (F3) and raw data (F7) as well as the detail of algorithm (F11) and the recommendation feedback from the other users in the same community (F18).

Hence, we decide to filter out the factors that were less preferred by the subjects. We choose to keep the factors with more than ten votes in the groups of "Very Important" and "Less Important", which are F1, F2, F6, F10, F12, F13, F14, F15, F16, F19, the chosen factors were highlighted in red color in Table 1 We can project the factors back to the original explanatory goals. The mentioned percentage of each exploration goal is listed as below: Transparency (40%, 2 out of 5), Scrutability (0%, 0 out of 4), Trust (33%, 1 out of 3), Persuasiveness (67%, 2 out of 3), Effectiveness (33%, 1 out of 3), Efficiency (50%, 1 out of 2) and Satisfaction, (75%, 3 out of 4). That is, the *Target Mental model* was built through the exploratory goal of (rank from high to low importance) Satisfaction, Persuasiveness, Efficiency, Transparency, Trust, and Effectiveness.

6 FOURTH STAGE: ITERATIVE PROTOTYPING

The fourth stage: interactive prototyping was performed within the same user study as the third stage. After the card-sorting task, we asked the subject to identify the chosen ten factors across some UI prototypes. A total of 25 interfaces (five interfaces for each recommendation model) were exposed in this stage. We used a within-subject design, i.e., all participants required to do a card-sorting task. In each session, the participants were asked to sort the given five interfaces into groups 1 to 5 (1: Strongly Agree, 5: Strongly Disagree), in each exploratory factor. If one interface is not contributing to the factor, the participant can mark it as irrelevant (not applicable). We continued with a semi-interview after the subject completed each session to collect the qualitative feedback.

There were a total of five card-sorting sessions for all five recommendation model. At the beginning of each session, we introduced the recommendation model through the *Expert Mental Model*, i.e., tell the participant how the similarity is calculated and what data were adopted in this process, to make sure the subject understands the details of each recommendation model. After that, we provided five interface printouts, a paper sheet with a table contains 19 exploratory factors and a pen - the subjects were expected to write down rankings on the paper sheet. All subjects took around 80-100 minutes to complete the study.

6.1 Explaining Publication Similarity

The key component of publication similarity is *terms* and *term frequency* of the publication as well as its mutual relationship (i.e., the common terms) between two scholars. We presented four visual interface prototypes (shown in Figure 2) for explaining publication similarity and one text-based interface (E1-1), which simply says "You and [the scholar] have common words in [W1], [W2], [W3]."

6.1.1 E1-2: Two-way Bar Chart. The bar chart is a common approach in analyzing the text mining outcome [15] using a histogram of terms and term frequency. We extended the design to a two-way bar chart to show the mutual relationship of two scholars' publication terms and term frequency, i.e., one scholar in positive and the other scholar on a negative scale. The design is shown in Figure 2a.

6.1.2 E1-3: Word Clouds. Word cloud is a common design in explaining text similarity [18]. We adopted the word cloud design from [26], which presented the term in the cloud and the term frequency by the font size. This interface provided two word clouds (one for each scholar) so the user can perceive the mutual relationship. The design is shown in Figure 2b.

6.1.3 E1-4: Venn Word Cloud. Venn diagram was recognized as an effective hybrid explanation interface by Kouki et al. [8]. This interface could be considered as a combination of a word cloud and a Venn diagram [22], which presents term frequency using the font size. The unique terms of each scholar are shown in a different color (green and blue) while the common terms are presented in the middle, with red color, for determining the mutual relationship. The design is shown in Figure 2c.

6.1.4 E1-5: Interactive Word Cloud. A word cloud can be interactive. We extend the idea from [18] and used Zoomdata Wordcloud [27],

(a) E1-2: Two-way Bar Chart

(b) E1-3: Word Clouds

(c) E1-4: Venn Word Cloud

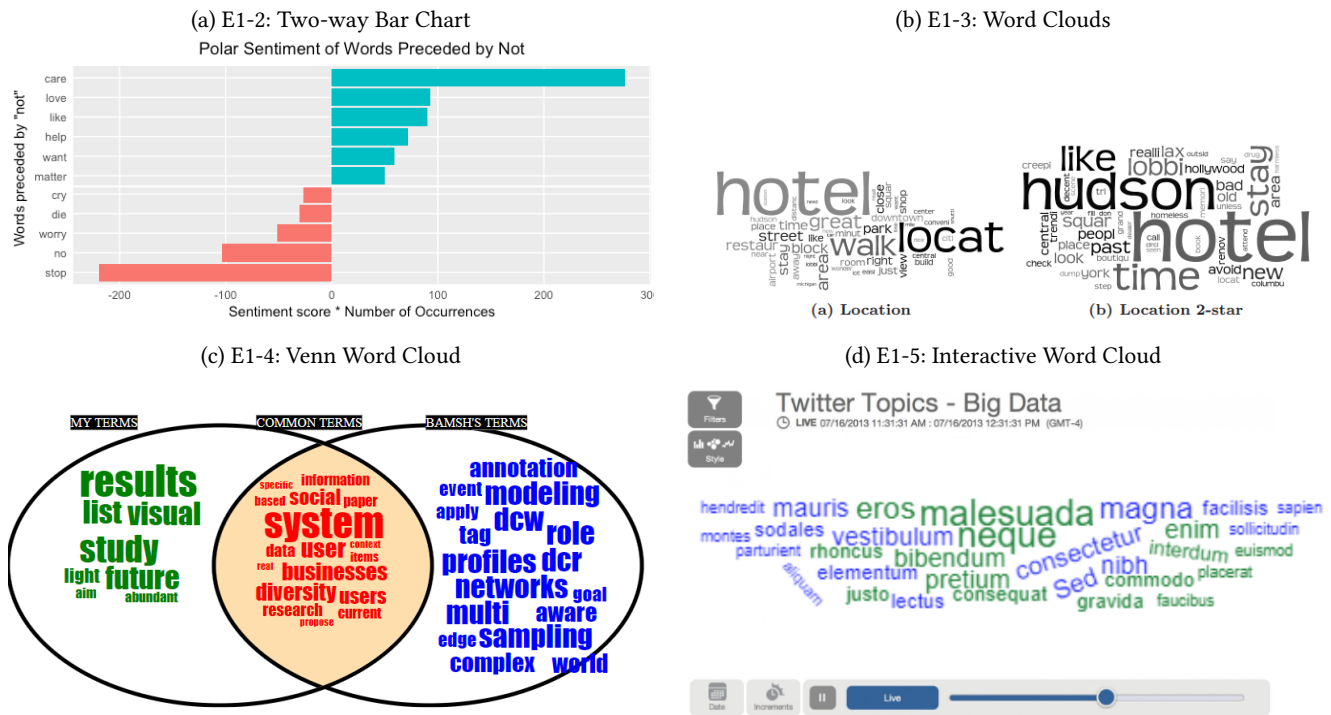(d) E1-5: Interactive Word Cloud

Figure 2: The interfaces used to explain the *Publication Similarity* in the fourth stage.

which follows the common approach to visualize term frequency with the font size. The font color was selected to distinguish the scholars' terms, i.e., different term color for each scholar. A slider was attached to the bottom of the interface that provides real-time interactive functionality to increase or decrease the number of terms in the word cloud. The design was shown in Figure 2d.

6.1.5 Results. The card-sorting result was presented in Table 2. We found the *E1-4 Venn Word Cloud* was preferred by the participants, received 76 votes in Rank 1, which was outperformed other four interfaces. According to the post-session interview, 13 subjects agreed E1-4 is the best interface versus the other four interfaces. The supporting reasons can be summarized as 1) the Venn diagram provided common terms in the middle, which highlighted the common terms and shared relationship; 2) it is useful to show non-overlapping terms on the sides (N=5) and 3) the design is simple, easy to understand and require less time to process (N=3). Two subjects mentioned they preferred E1-2 the most due to histograms gives them the "concrete numbers" for "calculating" the similarity, which was harder when using word clouds.

6.2 Explaining Topic Similarity

The key component of topic similarity is *research topics* and *topical words* of the scholar as well as its mutual relationship (i.e., the common research topics) between two scholars. We presented four visual interfaces prototypes (shown in Figure 3) and one text-based prototype for explaining the topic similarity. The text-based interface (E2-1) simply says "You and [the scholar] have common research topics on [T1], [T2], [T3]."

6.2.1 E2-2: Topical Words. This interface followed the approach of [10], which attempted to help users in interpreting the meaning of each topic by presented topical words in a table. We adopted the idea as *E2-2 Topical Words* that present the topical words in two multi-column tables (each column contains the top 10 words of each topic). The design is shown in Figure 3a.

6.2.2 E2-3: FLAME. This interface followed Wu and Ester [26], which adopted a bar chart and two word clouds in displaying the opinion mining result. In their design, each bar would be considered as a "sentiment"; then the user can interpret the model by the figure (for the *beta* value of topic) and table (for the topical words). We extended the idea as *E2-3: FLAME* that showed two sets of research topics (top 5) and the relevant topic words in two word clouds (one for each scholar). The design is shown in Figure 3b.

6.2.3 E2-4: Topical Radar. The *E2-4 Topical Radar* was used in Tsai and Peter [22]. The radar chart was presented in the left. We picked the top 5 topics (ranked by *beta* value from a total of 30 topics) of the user and compared them with the examined attendee through the overlay. A table with topical words was presented in the right so that the user can inspect the context of each research topic. The design is shown in Figure 3c.

6.2.4 E2-5: Topical Bars. We adopted several bar charts in this interface as *E2-5: Topical Bar*. The interface showed top three topics of two scholars (top row and the second row) and the topical information (top eight topical words in the y-axis and topic *beta* value in x-axis) using a bar chart with histograms. The design was shown in Figure 3d.

(a) E2-2: Topical Words

(b) E2-3: FLAME

(c) E2-4: Topical Radar

(d) E2-5: Topical Bar

Figure 3: The interfaces used to explain the *Topic Similarity* in the fourth stage.

6.2.5 Results. The card-sorting result was presented in Table 2. We found the *E2-4 Topical Radar* received 86 votes in Rank 1 outperforming all other interfaces. E2-3 ended up being second with most votes in the R2 group. According to the post-session interview, 13 subjects agreed E2-4 is the best interface among all examined interfaces. One subject preferred E2-3, and one subject suggested a mix of E2-3 and E2-4 as the best design. The supporting reasons for E2-4 can be summarized as 1) It is easy to see the relevance through the overlapping area from the Radar chart and the percentage numbers from the table (N=12). 2) It is informative to compare the shared research topics and topical words (N=9).

6.3 Explaining Co-Authorship Similarity

The key component of co-authorship similarity is *coauthors*, *coauthorship* and *distance of connections* of the scholars as well as its mutual relationship (i.e., the connecting path) between two scholars. We presents the five prototyping interfaces (shown in Figure 4, E3-1 presented in text below) for explaining publication similarity. In addition to four visualized interfaces, we also include one text-based interface (E3-1). That is, "You and [the scholar] have common co-authors, they are [A1], [A2], [A3]."

6.3.1 E3-2: Correlation Matrix. E3-2 Correlation matrix was inspired by Heckel et al. [4] that was used to present overlapping user-item co-clusters in a scalable and interpretable product recommendation model. We extended the interface to a user-to-user correlation matrix that the user can inspect the scholar co-authorship network. The design was shown in Figure 4(a).

6.3.2 E3-3: ForceAtlas2. E3-3: ForceAtlas2 was inspired by Garnett et al. [3] that presented Co-authorship graph of NiMCS and related research with both high and low-level network structure and information. Nodes and edges are representing authors and co-authorship, respectively. Graph layout uses the ForceAtlas2 algorithm [3]. Clusters are calculated via Louvain modularity and delineated by color. The frequency of co-authorship is calculated via Eigenvector centrality and represented by size. The design was shown in Figure 4(b).

6.3.3 E3-4: Strength Graph. E3-4 Strength Graph was inspired by Tsai and Brusilovsky [18] that tried to present the co-authorship network using D3plus network style [9]. Nodes and edges are representing authors and co-authorship, respectively. The edge thickness is the weighting of the coauthorship (number of co-worked papers). The node was assigned different color by their groups, i.e., the original scholar, target scholar and via scholars. The design was shown in Figure 4(c).

6.3.4 E3-5: Social Viz. The *E3-5 Social Viz* was used in [22]. There were six possible paths (one shortest and five alternatives). The user will be presented in the left with a yellow circle. The target user will be presented in the right with red color. The circle size represented the weighting of the scholar, which was determined by the appearing frequency in the six paths. For example, the scholar *Peter* is the only node that scholar *Chu* can reach scholar *Nav*, so the circle size was the largest one (size = 6). The design was shown in Figure 4(d).

6.3.5 Results. The card-sorting result was presented in Table 2. We found the *E3-4 Strength Graph* was preferred by the participants, received 45 votes in Rank 1. However, the votes were close with
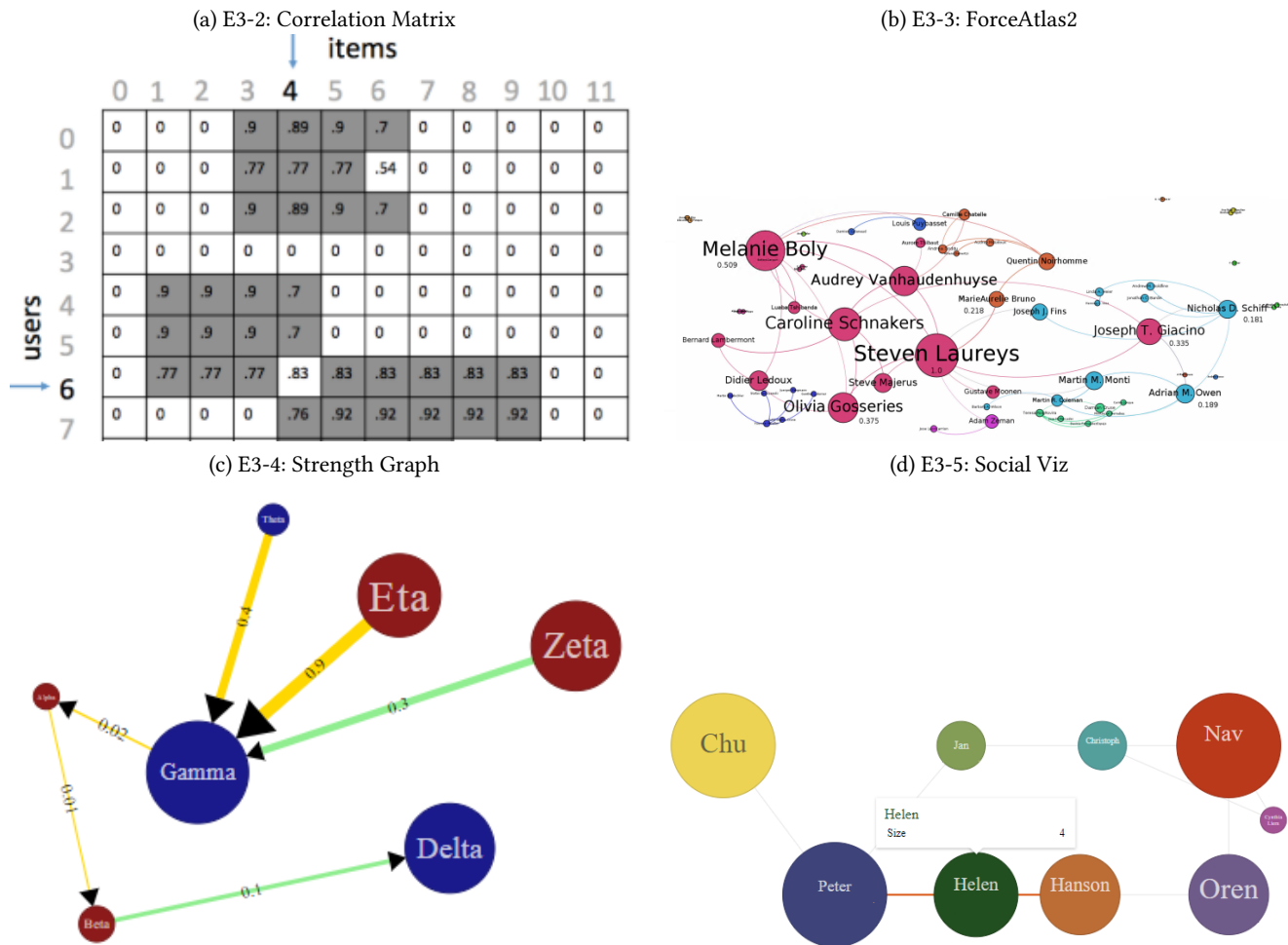
(a) E3-2: Correlation Matrix

(b) E3-3: ForceAtlas2

(c) E3-4: Strength Graph

(d) E3-5: Social Viz

Figure 4: The interfaces used to explain the *Co-Authorship Similarity* in the fourth stage.

E3-2 Correlation Matrix (37 votes) and *E3-3 ForceAtlas2* (32 votes). According to the post-session interview, four subjects agreed E3-4 is the best interface versus the other four interfaces. The supporting reasons were the interface highlighted the mutual relations and let the user can understand the path between two scholars. The arrow and edge thickness were also useful. Two subjects supported E3-2, they liked the correlation matrix provided a clear number and correlation information that easier for them to process. Three subjects supported E3-3, they preferred the interface provided a piece of high-level information by giving a "big picture". Also, E3-3 would be good to explore the coauthorship network beyond the connecting path, although the interface was reported to be too complicated as an explanation. Four subjects supported E3-5, they enjoy the simple, clear and "straightforward" connecting path as the explanation for coauthorship network.

6.4 Explaining CN3 Interest Similarity

The key component of CN3 interest similarity is *papers* and *authors* of the system bookmarking as well as its mutual relationship (i.e.,

the common terms) between two scholars. We presented the five prototyping interfaces (shown in Figure 5, E4-1 presented in the text below) for explaining publication similarity. In addition to four visualized interfaces, we also include one text-based interface (E4-1). That is, "You and [the scholar] have common bookmarking, they are [P1], [P2], [P3]."

6.4.1 E4-2: Similar Keywords. *E4-2 Similar Keywords* was proposed and deployed in Conference Navigator [11]. We extended the interface to explain shared bookmarks between two scholars. The interface represents the scholars in two sides and the common co-bookmarking items (e.g., the five common co-bookmark papers or authors) in the middle. A strong (solid line) or weak (dash line) tie will be used to connect the item was bookmarked by the one-side or two-sides. The design was shown in Figure 5(a).

6.4.2 E4-3: Tagsplanations. *E4-3 Tagsplanations* was proposed by Vig et al. [24]. The idea is to show both tag, user preference, and relevance that used to recommending movies. We extended the interface to explain the co-bookmarking information. In our design,
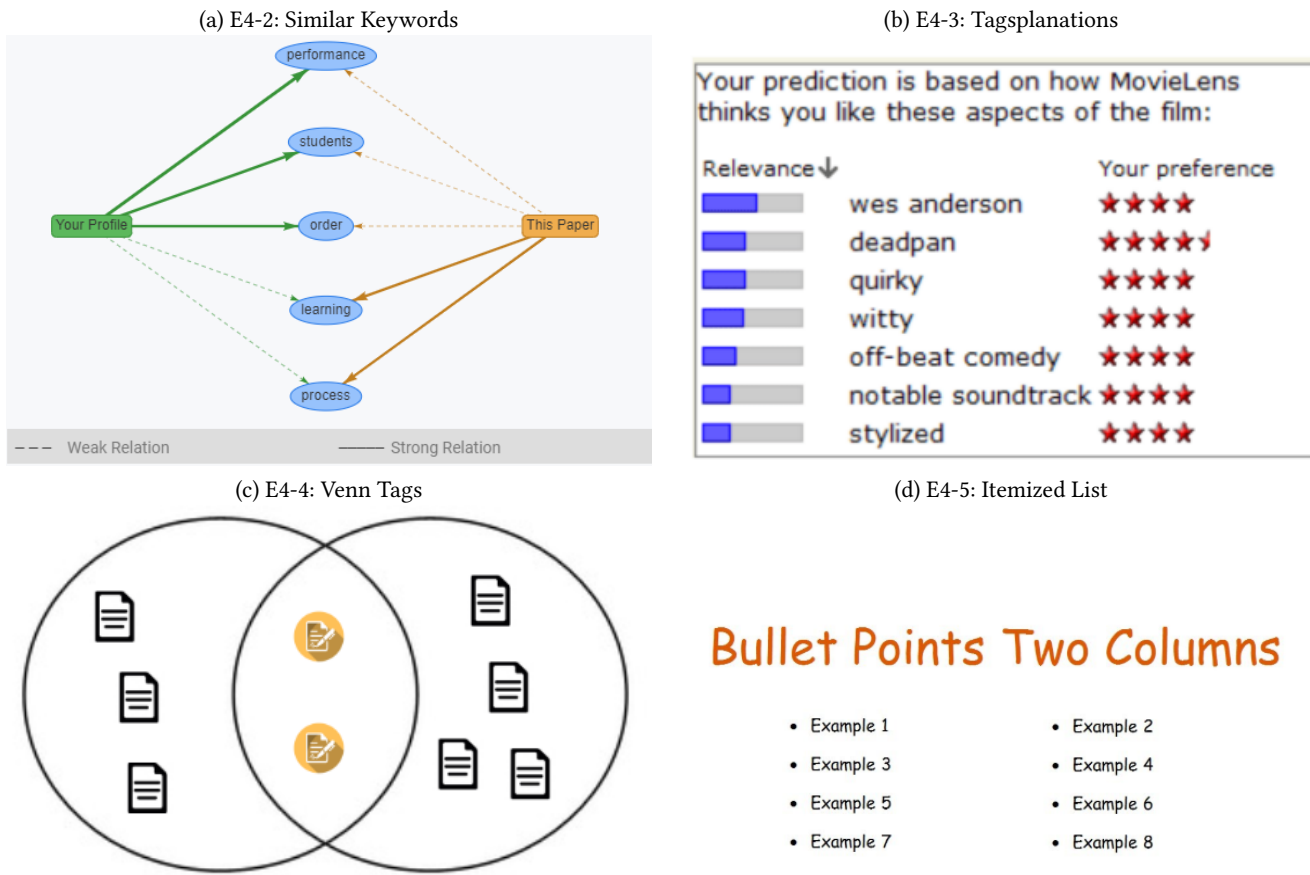
(a) E4-2: Similar Keywords

(b) E4-3: Tagsplanations

(c) E4-4: Venn Tags

(d) E4-5: Itemized List

Figure 5: The interfaces used to explain the *CN3 Interest Similarity* in the fourth stage.

the co-bookmarked item will be listed and ranked by its social popularity, i.e., how many users have followed/bookmarked the item? The design was shown in Figure 5(b).

6.4.3 E4-4: Venn Tags. The study of [8] has pointed out the user preferred the Venn diagram as an explanation in a recommender system. In the interface of *E4-4: Venn Tags*, we implemented the same idea with the bookmarked items. The idea is to present the bookmarked item, using an icon, in the Venn diagram. The two sides are the bookmarked item belong to one party. The co-bookmarked or co-followed item will be placed in the middle. The users can hover the icon for detail information, i.e., paper title or author name. The design was shown in Figure 5(c).

6.4.4 E4-5: Itemized List. An itemized list has been adopted to explain the bookmark in [21]. We proposed *E4-5: Itemized List* that presented the bookmarked or followed items in two lists. The design was shown in Figure 5(d).

6.4.5 Results. The card-sorting result was presented in Table 2. We found the *E4-4 Venn Tags* was preferred by the participants, received 64 votes in Rank 1, which was outperformed all other four interfaces. *E4-4 Venn Tags* was also be favored by the subject, which received 49 votes. According to the post-session interview,

eight subjects agreed E4-4 is the best interface versus the other four interfaces. The supporting reasons can be summarized as 1) the Venn diagram is more familiar or clear than other interfaces (N=4); 2) The Venn diagram is simple and easy to understand (N=4). Three subjects mentioned they preferred E4-3 the most due to the interface provide extra attribution, don't need to hover for detail and easy-to-use.

6.5 Explaining Geographic Similarity

The key component of geographic similarity is *location* and *distance* of the two scholars as well as their mutual relationship (i.e., the geographic distance). We presented the five prototyping interfaces (shown in Figure 6, E5-1 presented in the text below) for explaining the geographic similarity. In addition to four visualized interfaces, we also include one text-based interface (E5-1). That is, "From [Institution A] to [sample]'s affiliation ([Institution B]) = N miles."

6.5.1 E5-2: Earth Style. Using Google Map [6] for explaining geographic distance in a social recommender system has been discussed in Tsai and Brusilovsky [21]. We extended the interface to a different style. In *E5-2 Earth Style*, we "zoom out" the map to an earth surface and place the two connected icons (with geographic distance) on the map. The design was shown in Figure 6(a).
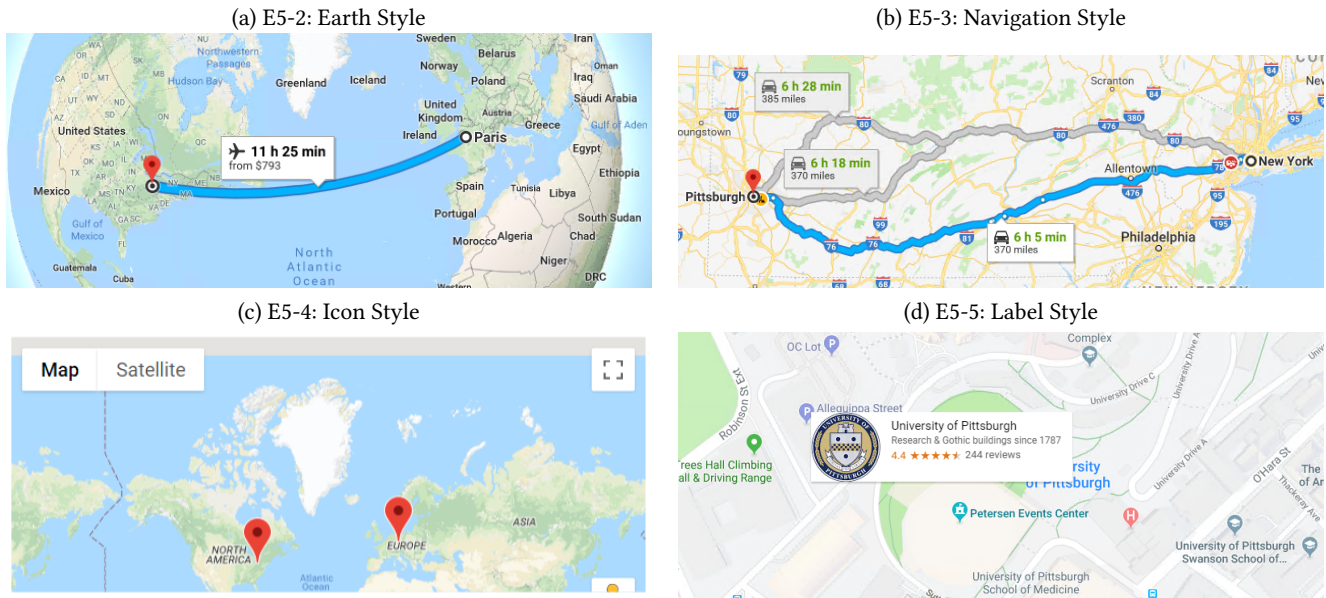
(a) E5-2: Earth Style

(b) E5-3: Navigation Style

(c) E5-4: Icon Style

(d) E5-5: Label Style

Figure 6: The interfaces used to explain the *Geography Similarity* in the fourth stage.

6.5.2 E5-3: Navigation Style. E5-3 *Navigation Style* followed the same Google Map API (shown in E5-2), but presented navigation between the two locations, either by car or flight. To be noted, the *transportation time*, i.e., the fly or driving time in E5-2 or E5-3, did not be considered in the recommendation model. The design was shown in Figure 6(b).

6.5.3 E5-4: Icon Style. E5-4 *Icon Style* followed the same Google Map API (shown in E5-2), but presented two icons on the map without any navigation information. The users can hover to see the detail affiliation, but the geographic distance information was not presented. The design was shown in Figure 6(c).

6.5.4 E5-5: Label Style. E5-4 *Label Style* followed the same Google Map API (shown in E5-2), but presented two labels on the map without any navigation information. The users can see the detail affiliation profile through the floating label without extra clicking or hovering interactions. The design was shown in Figure 6(d).

6.5.5 Results. The card-sorting result was presented in Table 2. We found the *E5-3 Navigation Style* was preferred by the participants, received 42 votes in Rank 1. However, the votes are close with *E3-5 Label Style* (40 votes). According to the post-session interview, six subjects agreed E5-3 is the best interface versus the other four interfaces. But there were three subjects particularly mentioned the navigation function was irrelevant in explaining or exploring the social recommendations. The supporting reasons of E5-3 can be summarized as 1) The map is informative (N=2). 2) It is useful to see navigation (N=5). Three subjects mentioned they preferred E5-5 the most due to the label contains affiliation information that they can understand the affiliation without extra actions. Although there is no geographic distance information, one subject pointed out he will realize the distance after knowing the affiliation title.

7 DISCUSSION AND CONCLUSIONS

In this work-in-progress paper, we presented a participatory process of bringing explanation interfaces to a social recommender system. We proposed four stages in responding to the challenge questions in identifying the key components of explanation models and mental models. In the first stage, we discussed the *Expert Mental Model* by discussing the key components (based on the similarity algorithm) of each recommendation model. In the second stage, we reported an online survey of current system users (N=14) and identified 19 explanatory goals as the *User Mental Model*. In the third stage, we reported the card-sorting results of a controlled user study (N=15) that created the *Target Mental Model* through the target users' preference of the explanatory factors.

In the fourth stage, we proposed a total of 25 explanation interfaces for five recommendation models and reported the card-sorting and semi-interview result. We found, in general, the participants preferred visualization interfaces more than the text-based interface. Based on the study, we found *E1-4: Venn Word Cloud*, *E2-4: Topical Radar*, *E3-4: Strength Graph*, *E4-4: Venn Tags*, *E5-3: Navigation Style* were preferred by the study participants. We further discussed the top-rated and second-rated explanation interfaces and user feedback in each session. Based on the experiment results, we concluded the design guideline of bringing the explanation interface to a real-world social recommender system.

A further controlled study will be required to test if the proposed explanation interface can achieve the target mental as we identified in this paper. In our future works, we plan to implement the top-rated explanation interfaces and deploy those interfaces to the CN3 system. Moreover, we expect to provide the explanation interfaces with an information-seeking task, so we can analyze how and why does a user adopt the explanation interfaces in exploring the social recommendations.

REFERENCES

[1] Malin Eiband, Hanna Schneider, Mark Bilandzic, Julian Fazekas-Con, Mareike Haug, and Heinrich Hussmann. 2018. Bringing Transparency Design into Practice. In *23rd International Conference on Intelligent User Interfaces*. ACM, 211–223.

[2] Gerhard Friedrich and Markus Zanker. 2011. A taxonomy for generating explanations in recommender systems. *AI Magazine* 32, 3 (2011), 90–98.

[3] Alex Garnett, Grace Lee, and Judy Illes. 2013. Publication trends in neuroimaging of minimally conscious states. *PeerJ* 1 (2013), e155.

[4] Reinhard Heckel, Michail Vlachos, Thomas Parnell, and Celestine Dünner. 2017. Scalable and interpretable product recommendations via overlapping co-clustering. In *Data Engineering (ICDE), 2017 IEEE 33rd International Conference on*. IEEE, 1033–1044.

[5] Jonathan L Herlocker, Joseph A Konstan, and John Riedl. 2000. Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. ACM, 241–250.

[6] Google Inc. 2018. Google Maps Directions API. https://developers.google.com/maps/documentation/directions/intro

[7] Bart P. Knijnenburg, Svetlin Bostandjiev, John O'Donovan, and Alfred Kobsa. 2012. Inspectability and Control in Social Recommenders. In *6th ACM Conference on Recommender System*. 43–50.

[8] Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. 2017. User preferences for hybrid explanations. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*. ACM, 84–88.

[9] Lawrence. 2018. Customize D3plus network style. https://codepen.io/choznerol/pen/evaYyv

[10] Julian McAuley and Jure Leskovec. 2013. Hidden factors and hidden topics: understanding rating dimensions with review text. In *Proceedings of the 7th ACM conference on Recommender systems*. ACM, 165–172.

[11] Conference Navigator. 2018. Paper Tuner. http://halley.exp.sis.pitt.edu/cn3/portalindex.php

[12] Alexis Papadimitriou, Panagiotis Symeonidis, and Yannis Manolopoulos. 2012. A generalized taxonomy of explanations styles for traditional and social recommender systems. *Data Mining and Knowledge Discovery* 24, 3 (2012), 555–583.

[13] Pearl Pu and Li Chen. 2007. Trust-inspiring explanation interfaces for recommender systems. *Knowledge-Based Systems* 20, 6 (2007), 542–556.

[14] Amit Sharma and Dan Cosley. 2013. Do social explanations work?: studying and modeling the effects of social explanations in recommender systems. In *Proceedings of the 22nd international conference on World Wide Web*. ACM, 1133–1144.

[15] Julia Silge and David Robinson. 2016. tidytext: Text mining and analysis using tidy data principles in r. *The Journal of Open Source Software* 1, 3 (2016), 37.

[16] Nava Tintarev and Judith Masthoff. 2012. Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction* 22, 4-5 (1 Oct. 2012), 399–439.

[17] Nava Tintarev and Judith Masthoff. 2015. Explaining recommendations: Design and evaluation. In *Recommender systems handbook*. Springer, 353–382.

[18] Chun-Hua Tsai and Peter Brusilovsky. 2017. Providing Control and Transparency in a Social Recommender System for Academic Conferences. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*. ACM, 313–317.

[19] Chun-Hua Tsai and Peter Brusilovsky. 2018. Beyond the Ranked List: User-Driven Exploration and Diversification of Social Recommendation. In *23rd International Conference on Intelligent User Interfaces*. ACM, 239–250.

[20] Chun-Hua Tsai and Peter Brusilovsky. 2018. Explaining Social Recommendations to Casual Users: Design Principles and Opportunities. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*. ACM, 59.

[21] Chun-Hua Tsai and Peter Brusilovsky. 2019. Exploring Social Recommendations with Visual Diversity-Promoting Interfaces. *TiiS* 1, 1 (2019), 1–1.

[22] Chun-Hua Tsai and Brusilovsky Peter. 2019. Explaining Recommendations in an Interactive Hybrid Social Recommender. In *Proceedings of the 2019 Conference on Intelligent User Interface*. ACM, 1–12.

[23] Katrien Verbert, Denis Parra, Peter Brusilovsky, and Erik Duval. 2013. Visualizing recommendations to support exploration, transparency and controllability. In *Proceedings of the 2013 international conference on Intelligent user interfaces*. ACM, 351–362.

[24] Jesse Vig, Shilad Sen, and John Riedl. 2009. Tagsplanations: explaining recommendations using tags. In *Proceedings of the 14th international conference on Intelligent user interfaces*. ACM, 47–56.

[25] Yao Wu and Martin Ester. 2015. FLAME: A Probabilistic Model Combining Aspect Based Opinion Mining and Collaborative Filtering. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15)*. ACM, New York, NY, USA, 199–208. https://doi.org/10.1145/2684822.2685291

[26] Yao Wu and Martin Ester. 2015. Flame: A probabilistic model combining aspect based opinion mining and collaborative filtering. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*. ACM, 199–208.

[27] Zoomdata. 2018. Real-time Interactive Zoomdata Wordcloud. https://visual.ly/community/interactive-graphic/social-media/real-time-interactive-zoomdata-wordcloud