

Crowd-powered Semantic Interaction for Text Analytics

Yali Bian, Tianyi Li, Ji Wang, Kurt Luther, Chris North

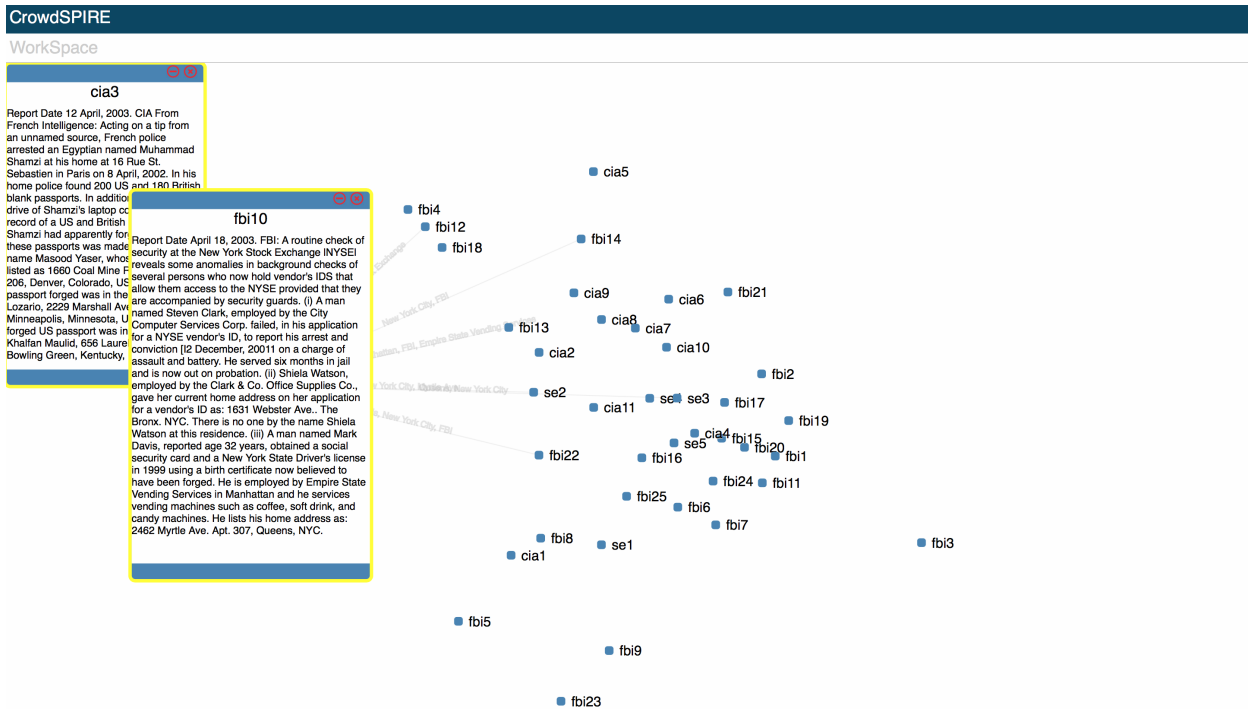


Fig. 1. Crowd-powered Semantic Interaction for Text Analytics

Abstract— Visual analytics could help users explore and gain insight from dataset through interactive visualization and underlying analytics models. However, making sense of large text dataset is still challenging in many domains. For example, an intelligence analyst might need to find a coordinated terrorist assault in three US cities from one hundred of documents in a limited time. Existing visual analytics techniques only assist with low-level tasks, such as finding related documents based on shared entities, which still require significant effort on the part of users.

To support the sensemaking process in visual analytics, we present the concept of crowd-powered semantic interaction, where semantic interactions can be used assign micro-tasks to crowds. This model could help users to steer crowd-workers to carry out sensemaking tasks implicitly even they are novice on crowdsourcing. The completed output could be used to update the visualization appropriately that foster the related foraging and synthesis parts for users. To demonstrate this model, we introduce CrowdSPIRE, a visual text analytics prototype that converts user interactions on documents into micro-tasks called "Connect the Dots". For example, the user drags two texts together, which triggers the assignment of certain sensemaking tasks to crowdworkers automatically, in parallel with the experts own investigation. When the tasks finished, outputs could be used to update current visualization appropriately without distracting user's attention from the current thinking.

Index Terms—Visual analytics, Semantic Interaction, Crowdsourcing, Sensemaking, Crowd-powered Interface.

1 INTRODUCTION

Sensemaking large amount of unstructured text is urgently needed but challenging today. If we can find ways to make sense of this big data, the possibilities for learning more about ourselves and how to improve the world we live in are almost boundless. For example, detecting and preventing a terrorist attack based on intelligence reports. Responding

to the challenging, there emerged the field of visual analytics[illuminate the path] that combines the powerful approaches—information visualization and data mining—to create a new class of sensemaking tools enabling new kinds of exploration and insights. (Example)

However, it remains time-consuming and onerous, and existing support tools still have a long way to go. Machine learning techniques could help find clusters and summarize documents efficiently, but currently, they cannot generate questions, hypotheses, or conclusions based on data that are more subtle than what an algorithm has been programmed to recognize. What's more, visualization tools amplify the cognitive abilities of their users, but many users could only access to low-level information on documents, which requiring significant effort on the part of users to get the deep insight of data with all kinds of external knowledge. For example, analysts cannot find the relationship

- Yali Bian, Kurt Luther, Chris North are with Virginia Tech. E-mail: [yali, north]@vt.edu.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxx/TVCG.201x.xxxxxxx

between AMTRAK #19 with a terrorist attack if they don't consult the AMTRAK schedules on the Internet to see that Train #19 is in fact called "The Crescent."

Crowdsourcing presents new opportunities to deal with this issue by augmenting the cognitive work of individual analysts, providing more insightful analysis than automated approaches and scaling better than traditional work. Crowdsourcing was originally used for simple, independent tasks that leverage innate human abilities like transcribing text, identifying images, and categorizing or labeling items[. Recently, researchers have begun to investigate how crowdsourcing can be applied to complex sensemaking tasks, like creating a taxonomy of items or performing a bottom-up analysis of a large corpus of qualitative data. These efforts show promise for how crowds might assist an individual analyst with a difficult sensemaking problem.

Though crowdsourcing is a powerful method to enhance users sense-making process, the integration of crowdsourcing into visual analytics could be tough work. Analysts who are non-expert in crowdsourcing might have a hard time to design and assign tasks to crowds. Even they are experts on human computation; they still have to put their time and energy in designing appropriate HITs and putting it to a crowdsourcing platform, which will distract analysts from their current investigation. [More details?]

Semantic interaction has been approved to be a good way to help users focus on their cognition of interesting elements on visualization, at the same time steering underlying models implicitly. Through basic interactions like dragging documents together, highlight important sentences, analysts could steer computational analytical models implicitly, instead of mastering the model first and changing the low-level input parameters of algorithms directly. Semantic interaction provides a bridge between analysts who are non-expert on machine learning algorithms and underlying models.

With semantic interaction, an expert analyst could guide the machine learning algorithms by directly manipulating the visualization layout. Our vision is that the expert interactions will also guide crowdsourcing micro-tasks to do more sensemaking works that machine learning algorithms not good at. We propose the crowd-powered semantic interaction model by coupling visual analytics with crowdsourcing through semantic interaction, and therefore, coupling analysts not only with the ability to parse huge quantities of documents but also the power to make sense of complex tasks effectively.

Designing visual analytics systems with crowdsourcing involves new challenges that we should take into consideration. Visual analytics system should map interactions to appropriate relevant sensemaking-tasks semantically. For example, it is not a good idea to map interaction of dragging two documents away from each other to micro-tasks that find their similarities. Also, coordinating and aggregating different kinds of tasks to effectively contribute to sensemaking visual feedback is another problem. Since different sub-tasks could have different granularity, information context, correctness, and efficiency.

To address such challenges, we present the crowd-powered semantic interactions to support the combination of visual analytics with human computations. Our key contributions in this paper are as follows:

(1) We propose the crowd-powered semantic interaction model that improve users' sensemaking process through crowdsourcing and formalize this model in the form of an updated visualization pipeline enhanced with crowdsourcing.

(2) We formalize the mapping from interaction to micro-tasks as sensemaking allocation strategies based on the users reasoning.

(3) We classify current existing sensemaking crowdsourcing tasks for text analytics based on their granularity and efficiency and explore the solutions to integrate them into real-time visual analytics system.

(4) To demonstrate crowd-powered semantic interaction, we present CrowdSPIRE, a visual analytics prototype based on ForceSPIRE. We present a usage scenario to demonstrate how crowd-powered semantic interactions to be used to help users' complex sensemaking tasks.

2 RELATED WORK

The key idea of crowd-powered semantic interaction is enhancing visual analytics's ability to leverage users' cognition for complex sensemak-

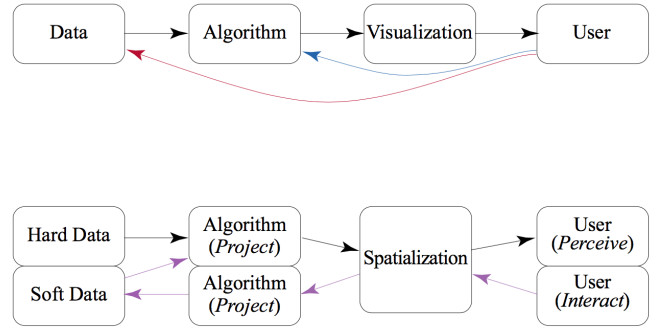


Fig. 2. (top) The basic version of the visualization pipeline. Interaction can be performed directly on the algorithm (blue arrow) or the data (red arrow). (bottom) Our modified version of the pipeline for semantic interaction, where the user interacts within the spatial metaphor (purple arrow). The image is from [8]

ing tasks with crowdsourcing through semantic interaction. Two key aspects are involved in this model: semantic interaction for visual analytics, crowdsourcing sensemaking tasks and crowd-powered interface.

2.1 Semantic Interaction

Visual analytics [15] combines the powerful approaches information visualization, and data mining together that creates a new class of sense-making tools enabling new kinds of exploration and insights. Usually, visual analytics systems provide visualized parameters such as controller bars [10] to help users directly steer lower-level computational models. For that kind of applications, the analyst needs to an expert in the underlying model based on their input parameters.

Semantic interaction makes visual analytic systems available for people who are not familiar with underlying mining method. Instead of providing interactions for controlling low-level input parameters of the algorithms, visual analytics with semantic interactions provides high-level interactions, which could be recast into low-level inputs through machine learning algorithms that attempt to recognize the reasoning process. Fig. 2 illustrates this the difference between basic visual analytics system and visual analytics system with semantic interaction, where the spatialization is treated as a medium through which the user can perceive information and gain insight, as well as interact and perform his analysis.

Semantic interaction successfully recognized analysts reasoning processes and relieved users from the need to organize many supporting documents or read many irrelevant documents [9]. We now recognize the opportunity to apply semantic interaction techniques to enable analysts to not only direct computational algorithms but also to manage a large force of crowd workers. Also, since semantic interaction recognizes opportunities for supporting subtasks and relevant information, it could also be used to narrow down context that the micro-tasks needed should assigned to crowd workers.

2.2 Crowdsourcing

Crowdsourcing [12] is a new and emerging research field that could help accomplish complex tasks with which computers typically struggle, such as image labeling, language translation. Nowadays, online crowdsourcing marketplaces like Amazon Mechanical Turk [1] where distributed groups of people complete small amounts of work (micro-tasks) for money make the use of human intelligence to perform tasks much available. Crowdsourcing has even been embedded in software back-ends and user interfaces to provides complex services, like, answering to visual questions [5].

2.2.1 Crowdsourced sensemaking tasks

Current research on crowdsourcing has changed from simple, independent tasks like images labeling [17] to more complex and creative tasks, like planning a vacation [18]. What's more, researchers start to

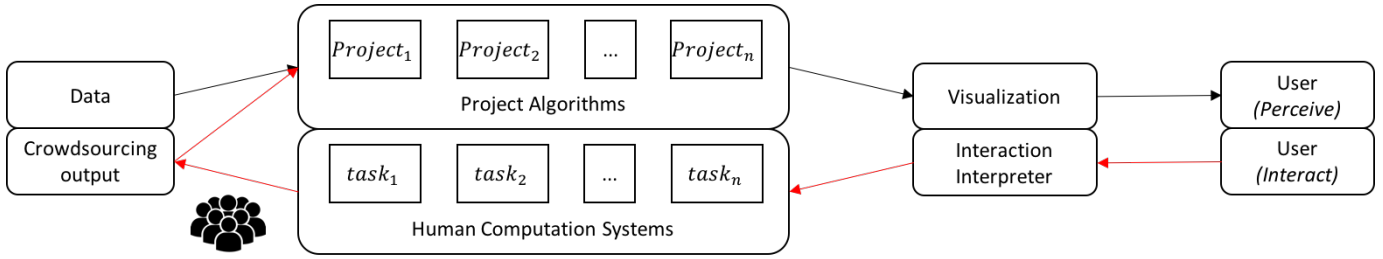


Fig. 3. crowdsourcing based semantic interaction visualization pipeline. Once the user perceives the visualization, they can choose to interact in it. This interaction feedback is interpreted as requests to the human computation system, which could assign several micro-tasks to crowds. The project algorithms could update visualization based on crowdsourcing outputs through their types, along with original data.

investigate how to apply crowdsourcing to complex sensemaking tasks, like creating a taxonomy of items or performing a bottom-up analysis of a large corpus of qualitative data, often with algorithms or workflows that decompose large tasks into smaller ones that can be completed in parallel.

Sensemaking [14] is a difficult process that involves tasks: finding relevant information, relating heterogeneous concepts, forming hypotheses, justifying arguments, reconciling conflicting or missing data, making inferences and complex reasoning [15]. There exist lots of crowdsourcing research explorations differ in their emphasis on different kinds of sensemaking tasks: Crowdnection [16], uses crowdsourcing to help connect high-level concepts through documents or raw texts. Cascade [7] produces crowdsourced taxonomies of hierarchical data sets by letting crowd works generate and later select, multiple categories per item. Frenzy [6] is a collaborative session organizer that classify papers into sessions based on their metadata. Crowdlines [13] explores how crowd works can synthesize information from diverse sources gathered online to produce a useful overview.

Visual analytics systems with crowd-powered semantic interaction could make full use of those great crowdsourcing research explorations on sensemaking as the underlying models that help mining complex documents and provides those crowd-processed information to users.

2.2.2 Crowd-powered System

With crowdsourcing in the computation to be universal, embedding human computation into applications to help carry out complex tasks that automatic computation is not skilled could be a good solution. And those kinds of applications with both automatic and human computation are called crowd-powered system.

However, crowdsourcing tasks that are usually time-consuming, making it difficult to incorporate and take advantage of crowds in real word applications. Lots of advanced crowdsourcing algorithms have been proposed to make crowd-powered systems easier to implement in real time.

VizWiz [4] provides a nearly real-time application to answer to visual questions by predicting and posting possible jobs to several crowd workers before users request directly. Soylent [2] is a crowd-powered word processor that divide writing tasks into small pieces that could be processed by crowd-workers parallelly.

Adrenaline [3] is a camera with crowdsourcing where crowds could help pick the best moment photo from a video. It could response to user's input within seconds based on the recruiting strategies to use synchronous crowd workers and dynamically adapt tasks. Chorus [11] is a conversational assistant that enables real-time interactions based on collaborative reasoning, dynamic scoring system and curated memory system.

Using these techniques, we could prototype out interactive crowd-powered visual analytics system that can provider users interactions feedbacks in real time and extends them to complex sensemaking tasks where crowds are directed by mechanisms other than explicit user requests.

3 CROWD-POWERED SEMANTIC INTERACTIONS

To leverage users from complex sensemaking tasks, we propose the crowd-powered semantic interaction. As shown on Fig. 3, visual analytics systems with crowd-powered semantic interaction could detect users' current sensemaking tasks based on their interactions; then complex tasks could be decomposed into sub-tasks and assigned to crowd workers; finally, recombine sub-task outputs and convey it via visual feedback to analysts.

We present an updated visual text analytics pipeline to reflect our crowd-powered semantic interaction model. At first, analysts will get an overview of original data based on underlying map and layout visualization algorithm. The user then could perceive data through their corresponding visual elements and explore interesting documents through searching terms and make sense of texts through interactions: such as text highlighting and documents movement.

Though the model of semantic interaction, interactions could be interpreted to user's current analytical reasoning with specific contexts. To be more specific, based on the controlled visual object, current visualization contexts, and interaction, we could determine user's sensemaking task in details. Then the specified sensemaking task could be projected to the crowdsourcing task that could be easily implemented. Not all sensemaking tasks will be directly projected to crowdsourcing HITs because the sensemaking task is too complex. For this situation, tasks should be decomposed into sub-tasks. For example, the sensemaking task corresponds to current interaction is compare two documents and find related documents based on two documents' similarities. We need at first, divide this task into finds connections between two documents, and finding relevant documents based on connections. When tasks finished nearly real-time, the crowdsourcing outputs combined with original data could again project to current visualization based on map and layout algorithms.

For example, an expert working in intelligence analysis has lots of reports to analyze in our visual analytics system with crowd-powered semantic interaction. The expert begins sifting through and grouping related documents. He puts three documents about "New York Stock Exchange" and "C-4" together to form a cluster. Meanwhile, the system creates sensemaking task of finding related documents. The system recruits crowd workers to suggest potentially documents that belong to the groups. Then the visualization updated based on the crowdsourcing results: there are three more documents emerged on the cluster: even those documents don't contain the keyword "New York Stock Exchange" or "C-4", but they do contain "NYSE" and "Explosive." After analyse this cluster, the expert find enough evidence that a terrorist group plan to bomb out the NYSE building.

Possible extension of this pipeline is combining automatic computation together with human computation to support users with sensemaking and computation power. Moreover, human computation and automatic computation could also provide modifications for each other. For example, system could dynamically control micro-tasks, based on removing irrelevant documents through data mining methods.

Two key techniques are involved in crowd-powered interaction: mapping interactions to sensemaking tasks based on task allocation strategy and updating visualization model based on crowds output. In the fol-

Table 1. Sensemaking context in each Semantic Interaction

Semantic Interaction	Associate Analytic Reasoning	Crowdsourced Subtasks	
		Foraging Task	Synthesis Task
Document Movement	Similarity/Dissimilarity Create spatial construct (e.g. cluster, timeline, list, etc.) Test hypothesis, see how document fits in region		Rearrange the spatial workspace to reflect the users organizational schema.
Text Highlighting	Mark importance of phrase (collection of entities) Augment visual appearance of document for reference	Retrieves documents matching the highlighted text from dataset	Find related entities and rearrange the document layout based on highlighted text
Pinning Document	Give semantic meaning to space/layout		Rearrange the spatial workspace to reflect the users organizational schema.
Annotation, Sticky Note	Put semantic information in workspace, within document context	Retrieve documents matching the annotation	Layout current documents on workspace based on annotations, find more links between this document based on the annotation
Open document	Change ease of visually referencing information (e.g., full detail = more important = easy to reference)	Retrieves more documents related to this document	Find more relationships between this document and other documents on current workspace. Re-layout the neighbors of selected document
Minimize document	Change ease of visually referencing information (e.g., full detail = more important = easy to reference)	Retrieves more documents related to this document	Find more relationships between this document and other documents on current workspace. Re-layout the neighbors of selected document
Search Terms	Expressive search for entity	Retrieve more documents matching searching terms	Find related terms have the meaning on current context, Re-layout documents on current workspace based on searching terms
Overlapping documents	Expressive search for entity	Retrieve more documents that connected with those two documents	Find relationships between overlapped documents and compare their contents, Re-layout documents on current workspace based on the overlapped documents locations

lowing two subsections, we discuss them in detail and explain how these techniques to improve users' sensemaking loop.

3.1 Task allocation strategy

Since semantic interactions right now are mainly on spatiation. We, design sensemaking tasks main on their relationships. For each interaction, semantic interaction could help us understand each of their reasoning, which we could define several sensemaking tasks. And each sensemaking tasks could be modularized and decomposed, to several sub-tasks so that crowdsourcing could be used to perform them. Different phrase have different contexts, Current senmaking tasks. For example, when searching new words, at the foraging loop, the context is the whole dataset. If highlight a setence on an opened document, the main context is current documents on the screen.

create multiple alternative competing hypotheses based on the given data find additional relevant information that supports the created hypotheses. dual search, occurs between these two loops in which analysts must simultaneously create multiple alternative competing hypotheses based on the given data and, at the same time, find additional relevant information that supports the created hypotheses.

Based on different sensemaking tasts and their context, we could design to the task based on task allocation strategy. For each of those sensemaking task, we could design several kind of sensemaking tasks to get a clear results. What's more, for each interaction, there is a context, that describe current sensemaking issues. How the sensemaking tasks identified in that study could be modularized and decomposed, so

that either automated techniques, or crowdsourcing, could be used to perform them. The goal of this study is to generate automated or crowd-powered alternatives to individual analysts performing these tasks, towards the goal of augmenting analysts with a suite of support tools. To reach this goal, we will perform a series of experiments to identify the benefits, drawbacks, and tradeoffs of using individuals, automated techniques like data mining, and crowdsourcing for each component of the sensemaking process.

The details of each subtask experimental procedure will vary depending on the details of the subtask alternatives; for example, foraging subtasks imply different participant goals than synthesis subtasks. However, in general, our goal is to devise and compare individual analyst performance and subjective experiences to those of crowdworkers, as well as a state-of-the-art automated approach. Crowdworker tasks and interfaces will be designed to assume minimal worker skill and expertise, to be verified using pre-survey screenings.

We convert the translates the sensemaking tasks into real problems like the distance between documents, which is a unified task purpose to do tasks and update the visual interface.

Spatializations are frequently employed to aid sensemaking (foraging and synthesis) of unstructured text documents [2, 21, 30, 33, 34]. Large, high-resolution displays in particular have been found beneficial in affording a large, flexible workspace that allows users to externalize knowledge and create semantic schemas.

For different interactions, the difference is that we get different contexts, and informations for this task. Also, each interaction, we

could reason this interaction from two kind of sensemaking phrase: based on the associate analytic reasoning, we list the task like this:

Two sensemaking phrase provide different functions, the foraging part is responsible for gathering relevant information based on current user's intention which could be expressed through interactions with the visualization. Even for the foraging crowdsourcing tasks, different interactions, could assign different level of foraging tasks: for example, text highlighting could trigger.

Right now semantic interactions are mainly based on spatiation, so the purpose of sensemaking tasks could also be the layout of statiation.

Users could make sense of dataset in details through interaction like searching. For example, ForceSPIRE[] provides a list of various forms of semantic interaction, including how each can be used within the analytic process of investigating textual information spatially. is list is likely incomplete, but serves as a starting point to introduce how semantic interaction can be integrated into a users reasoning process. Each interaction corresponds to reasoning of users within the analytic process. Corresponding model updates are performed to steer the crowd-tasks assign system. Based on different resoning of each interaction, that users use when they want to find more details about interested information about certain word, document, compare between two documets, or a cluster of documets. they want to narrow down the prolem.

**** Semantic interaction and their context** For each semantic interaction, they have a object to specifid, which narrow down their thinks to (Search for relations/in shobox) When they open document, As corresponded to past semantic interaction when they INTERPRETING interactions with ASSOCIATED ANALYTICAL REASONING In interpreting the interaction, the system determines the analytical reasoning associated with the interactions and updates the model accordingly. From previous findings [5], categories of user interaction can be associated with specific forms of analytical reasoning (see Table 4.1). It is essentially the models task to determine why, in terms of the data, the interaction occurred, and how that information can be used to augment and adjust the analytic models of the system to help the users task. e goal is to calculate, based on the data, what information is consistent with the captured interaction. For instance, we can associate text highlighting with adding importance to the text being highlighted. Since, each interaction have a reasoning when they do some interaction at specific level visual elemenets on visualization. Their sensemaking phrase could have a implicit specified context, based on [5]. Right now we list the implicit context as follows:

The context users used for sensemaking are as follows:

With diferent context:

As shown on tabel Table 1, when user drag two document together, this means on the context of whole documets shown on screen(dataset), they think those tow documets are similar, we should assign crowd that based on those two documents, we need find more related documents, and why those two documets are

Levels of documents could be mapped to different kinds of crowd tasks.

Why we assign crowd in this a way, because of sensemakings There are two very important factors that steer the crowd taks assignment part: sensemaking context, associate analytic reasoning.

For the associate analytic part, we could understand why two documents are related. For the associate analytic part, we could get more knowledge on this sensemaking context.

The crowd could help users find high level concepts from documents, Using crowdsourcing to help booster the sensemaking process through two part: foraging and synthesis.

Crowd Synthesis: make clusters.

As such crowd-workflows become complex, researchers must identify the level of crowd-supervision needed for optimal output. Intead of let crowdsourcing design experts to assign tasks to crowds, we use semantic interactions to latently

HIT Design based on Semantic Interactions

For different semantic interactions, we could assign different tasks. input: how to use semantic interaction as input to direct the crowd tasks

Table 2. Sensemaking tasks in different levels

Interaction	Task Schemas Level	Task demo
Keywords		
Minimize document	B	
Close document	A, B	
Annotation	A	
Search	A, B	
Highlight	A, B	
Overlap documents	A, B, C	
Cluster documents	D	

can create several kinds of micro-tasks for each SI, some quick, some slow (simulated in this paper with Tianyi data?)

Drags two docs together 1) e.g. when expert drags 2 docs together:

a) find entities that connect the 2 docs (quick) b) label semantic-level connections between the 2 docs (quick) -¿ text that can be used c) find related docs (slow) i) must compare to every other doc? ii) or use (a) and (b) to reduce the search set? context slice?

3.2 Integrate Crowds into Workspace

synchronous tasks How to integrating microtasks into a more collaborative, unstructured interface embodied in Frenzy and other forms of crowdware Human computation Open Document Search Keywords Clusters

To integrate the crowds into visual analytics, we need to merge all the sub tasks into formatted and combine the modularized subtasks in to a comprehensive sensemaking loop. To doing that, we should general the combination into two things: tasks levels (sub-tasks), and tasks time complexity.

considering how to recombine the modularized subtasks identified in the previous studies into a comprehensive, revised sensemaking loop, and to implement a software prototype based on this revised process. This effort implies a modification of the traditional sensemaking loop that accounts for the modularized components developed in Study 2. We plan a series of experiments leading to the design of effective workflows and task allocation strategies that allow individuals, crowds, and computation to synergistically perform complex sensemaking tasks, while minimizing bottlenecks and redundancies.

3.2.1 Crowds tasks in different levels

Crowds sensemaking tasks, based on different phrase, could be divided into two functions: foraging phrase: find more related documents on datasets. Find documents not based on whether they are entities related but also semantically related.

synthesis phrase: synthesize information with diverse or unknown schemas:

For the first phrase, we could translate the tasks output into distances or orders that which one close to each other.

For synthesis sub-tasks: we could find their schemas on three different levels based on different interactions.

Other level of relationships: compare the similarity between two overlapped documents, based on their contents, if has connections between each other, based on shared entities. or has same high level concepts. For this kinds of tasks, we needs store their outputs in three levels: entity links documents similariitys. clusters.

Entity level tasks

How to tranverse outs puts into workspaces. For each kind of crowd-sourcing tasks, we could integrate the results into workspaces.

Also, micro-task are not independent, micro-tasks could designed for each other automatically, for example, linked dots could be used as inputs or constraints for another subtasks.

Right now, crowdsourcing tasks for sensemaking could be classified into three levels: entities, documents (VizWiz: Nearly Real-time Answers to Visual Questions. find contents on documents, or edits ,

Frenzy: Collaborative Data Organization for Creating Conference Sessions.), clusters (Crowd Synthesis: Extracting Categories and Clusters from Complex Data.,). For each kinds of crowdsourcing tasks, each kind of crowdsourcing could build on each other. Lots of details on

Find entities that connect the 2 docs (quick) For the entities level, we could be used to be used as the input to other two level inputs. Also could used to provides as the inputs for automatic computation models.

For document level, we could find more related documents, or find similarity or dissimilarity between small number of documents (usually less than five documents). Find related documents, or remove unrelated documents. Directly to the documents,

For cluster level tasks, we could map the layout to the workspace directly, as a distance function.

label semantic-level connections between the 2 docs (quick) - ζ text that can be used find related docs (slow)

output: how to use crowd output in response to semantic interaction in the visualization can use crowd results in visualization (e.g. distance function for Force Directed layout) can use crowd results in further algorithmic processing (e.g. search) dynamic output, streaming from crowds

3.2.2 Crowds tasks in different time complexity

an individual analyst can be supported in real time by crowds or computation.

Real-time crowdsourcing systems have been developed to assist individuals with sensemaking tasks [5,6,26,27]. We seek to augment the capabilities of these analysts, regardless of their expertise, allowing them to work faster accomplish more than they could unaided.

real time For synthesis phrase tasks, main on a small number of documents,

For most foraging phrase tasks, since there are lots of documents to compile, automatic models could help to find the most possibly related documents, to let crowds to find. It might still be an time-assuming works (more than five mining): streaming batching results

for those crowdsourcing tasks, there are two approaches to managing latency in crowd-powered interfaces: streaming, batching results. Even if the crowdsourcing tasks are too slow to fit into the human in the loop (react less than ten seconds, the crowdsourcing could still be store to knowledge base, as an external knowledge for latter use, for example, analyst went to the similar situation, like overlap two document again, instead of assign a new task, the system could get crowdsourcing outputs immediately)

If current context changed, store data for latter use.

4 CROWDSPIRE

StartSPIRE. CrowdSPIRE (Crowd-powered Spatial Paradigm for Information Retrieval and Exploration) is a visual analytics tool prototype that implements crowd-powered semantic interaction technique: like ForceSPIRE, a semantic interaction visual analytics tool prototype for exploring unstructured text documents. CrowdSPIRE and ForceSPIRE share a flexible spatial workspace (driven by a modified force-directed layout and several semantic interactions. However, with different models on the background to help calculate the layouts of documents. Instead of using machine learning models, CrowdSPIRE use human computation to help calculate the distance between documents and update the layout of workspace.

Right now, CrowdSPIRE integrate the "Connect the dots" tasks, which let crowds labels related entities in documents in a context slice and helps crowd workers with the micro-task of creating and labeling connections between entities extracted from the text. Through related entities in each document, we could calculate their TF-IDF similarity. When doing overlapping two documents. This system extends upon previous work to integrate relevance-based retrieval and layout models, provides richer visual encodings, and adds to the semantic interactions leveraged.

StarSPIRE dynamically adjusts how many data points are displayed by using heuristic-based relevance metrics.

Difference from basic pipeline.

4.1 Visual Encodings

Within the spatial workspace, document nodes are visually encoded to relate their relevance to the users high dimensional understanding of the data [Figure 5]. Node size and saturation are encoded to reflect how closely a document matches the entities the user has deemed important. Node size and saturation are calculated by summing all of the entity weights in a document, ranking these values, and sorting them into quartiles. Quartiles were chosen instead of absolute ranking to optimize the node drawing process, minimizing the number of calculations and changes required with each user interaction. This was done to promote a quick interaction-feedback loop. These encodings give the illusion of a third dimension in the workspace where more important documents are in the foreground while less important documents fade into the background. However, unlike a true three-dimensional layout, document nodes cannot overlap each other, preventing occlusion. Additionally, StarSPIRE provides visual cues for navigating the workspace. Node color is used to indicate search term matches. Instead of showing all links between all documents, StarSPIRE restricts the edges shown to those connected to the selected node. Entities shared between documents are labelled on the edge, but are restricted to the top four entities, determined by their importance weights. All nodes are labelled with their documents titles in order to allow for easier navigation in the space and to allow users to track a specific node's movement throughout the space. Each node's outline color is used to denote its read or unread status in order to allow analysts to see which documents they have read and closed.

Within each document, search terms are identified and the text color is changed to allow the terms to stand out for easier identification. These encodings were identified and/or adjusted through an informal usability requirements analysis of StarSPIRE.

4.2 Crowd-powered Document Overlapping

CrowdSPIRE implement all the interactions on ForceSPIRE, users could explore the whole datasets based on document movement, text highlighting, pining document, annotation sticky note, open document, minimize document and overlapping documents. However, to make the system simple and easy to evaluate, we only combine document overlapping interaction with crowdsourcing tasks.

To evaluate the crowd-powered semantic interaction model, we only combined the document overlapping interaction with crowdsourcing tasks. At first, we have the distance between documents based on algorithms models.

We two or more documents overlapped each other, the semantic interaction will trigger the task allocation strategy design a connect the dots crowdsourcing, based on overlapped documents.

We define D as the set of overlapped documents, for each To carry out the 'connect the dots' task that help synthesis the overlapped documents: The task allocation strategy procedure that automatically assign current overlapping interactions to task:

- (1) Pick m documents d_1, d_j from D ($i \neq j$), for all the d_i, d_j .
- (2) Generate a Hit to MTurk that show m documents: (2) Show d_i, d_j to k workers on a visualization view sub-task, which requires workers connect the related entities if they are related.
- (3) For each link, the worker should input the certain, and how are these dots related. Document used to make this connection

For example, if three documents d_1 , and d_2 and d_3 are overlapped to each other, one of the micro-tasks is on Figure 1: $d_1, d_2, d_1, d_3 \dots$ will be formed to give tasks to different.

based two documents, the connect the dots will publish an Hit on MTurk

4.3 Integrate "Connect the Dots" into Workspace

To make full use of the 'Connect the Dots' tasks as a real time services, we prototype the task, before the overlapping interactions. For example, instead of design the hit, after the semantic interaction, we pre-assigned the task, and store outputs to the database, when the overlapping interaction be implements, we retrieve this context, as it is. The related

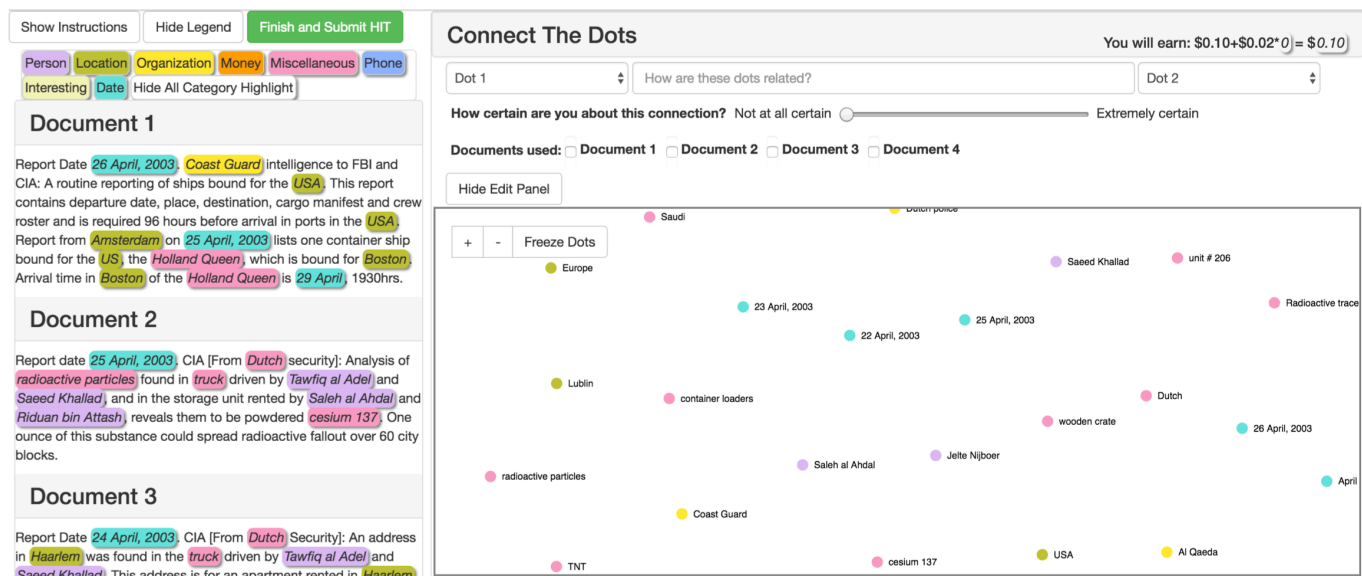


Fig. 4. The Connect the Dots web application interface

nodes could also be mapped to distances. based on Based on algorithms. Also, mapping the distance functions based on related entities. Pre-store current storage to mimic the real time crowdsourcing.

5 USAGE SCENARIO

In this section, we will demonstrate how CrowdSPIRE helps an analyst to make sense of complex dataset. We use The sign of the Crescent dataset[ref] which contains 41 fictional intelligence reports regarding a coordinated terrorist plot in three US cities. We will evaluate the effectiveness of CrowdSPIRE, we make a comparison of crowd-powered visual analytics (CrowdSPIRE) and algorithm-only based visual analytics (ForceSPIRE). Also, to test that the crowdsourcing tasks outputs correctness that we compare the layout of documents based on crowd-sourcing with the gold standard solution/and the most correct layouts we get when users make sense of documents.

In this section, we walk through a text analytics scenario to demonstrate the how crowdsourcing supports Comparison of crowd-enhanced version (CrowdSPIRE) with algorithm-only version(ForceSPIRE).

We prototype a web-version ForceSPIRE at the same version. The different between ForceSPIRE and CrowdSPIRE is that when overlapping documents together: ForceSPIRE will trigger the underlying machine learning algorithm to help calculatet the distance between documents. CrowdSPIRE will trigger the crowd micro-tasks allcation algorithm to assign "Connect the Dots" micro-tasks, and then use the output to calculate the document distance.

So if we let the a user explore the documents through this two documents, with the same interaction except overlapping documents, they will get similar layout. Through the layout of overlapping documents interaction, we could analyse current layout that could help user make sense of documents.

For example, if we drag document together, the different layout on two techniques could help use analyse the effectiveness of CrowdSPIRE.

To demonstrate StarSPIREs functionality, we used the VAST 2007 Challenge Dataset (Blue Iguanodon) [17]. Because StarSPIRE is currently designed to operate on unstructured text documents only, we omitted all images and spreadsheets from the dataset, resulting in approximately 1,500 text files. Blog entries that were included in the data were converted into text files, one for each blog entry. Preliminary entity extraction was done on the dataset. The challenge task is an open-ended sensemaking task to investigate unexpected activities concerning wildlife law enforcement, endangered species issues, and ecoterrorism [17]. We present the following usage scenario to demon-

strate how StarSPIRE can leverage the MSI technique. The user began with a search for chinchilla. This was unsurprising, because the dataset contained a directory titled Chinchillas. She read through several documents, arranging them in the display based on document similarity. The user then began highlighting information regarding chinchillas, which branched into additional endangered species. This loosely structured analysis continued until the user read a document concerning a musical artist owning an extremely large number of exotic animals whose actions did not seem to match his words regarding animal conservation. The analyst denoted this as suspicious and began investigating it further. This investigation was driven through highlighting the artists name and the name of his animal sanctuary, which imported many documents onto the display, some of which had a large node size. The analyst opened the largest new nodes first. [Figure 8] shows the evolution of the users spatial organization schemas through the sensemaking task. Clusters of documents were moved around the screen and a mixture of visual encodings and document proximity motivated the choice of documents to investigate next. Furthermore, it can be seen that the user initially executed two searches to obtain some initial documents, but then opted for other multi-scale semantic interaction techniques to obtain new documents (e.g. highlighting, linking documents denoted by the purple bars, and annotating documents). Document annotations were used to record hypotheses and insights (e.g. rBert is rBear? and rBear might have monkeypox). In the later stages of analysis, searches were used primarily to label the space, serving as reminders of which documents concerns which persons or topics. However, they were also used to ensure that important information or documents had not been overlooked.

Once the user identified suspicious activity regarding a large exotic animal reservation, it became apparent that many documents were interconnected via several subplots. As her understanding of the dataset evolved, so did her spatial representation. For example, two documents that were initially considered not quite relevant, but interesting enough to not minimize concerning an outbreak of a disease were initially placed in the upper right hand corner of the display. After realizing that the owner of the large exotic animal sanctuary had contracted the same disease, she moved the two documents down next to the exotic animal sanctuary documents. Highlights, document annotations, and document linking were primarily used to obtain new documents in the workspace. Searches were executed to check for additional information on important persons, but also used to label the spatial workspace. After approximately ninety minutes of analyzing the data, the user concluded that she had a sufficient understanding of the plot and

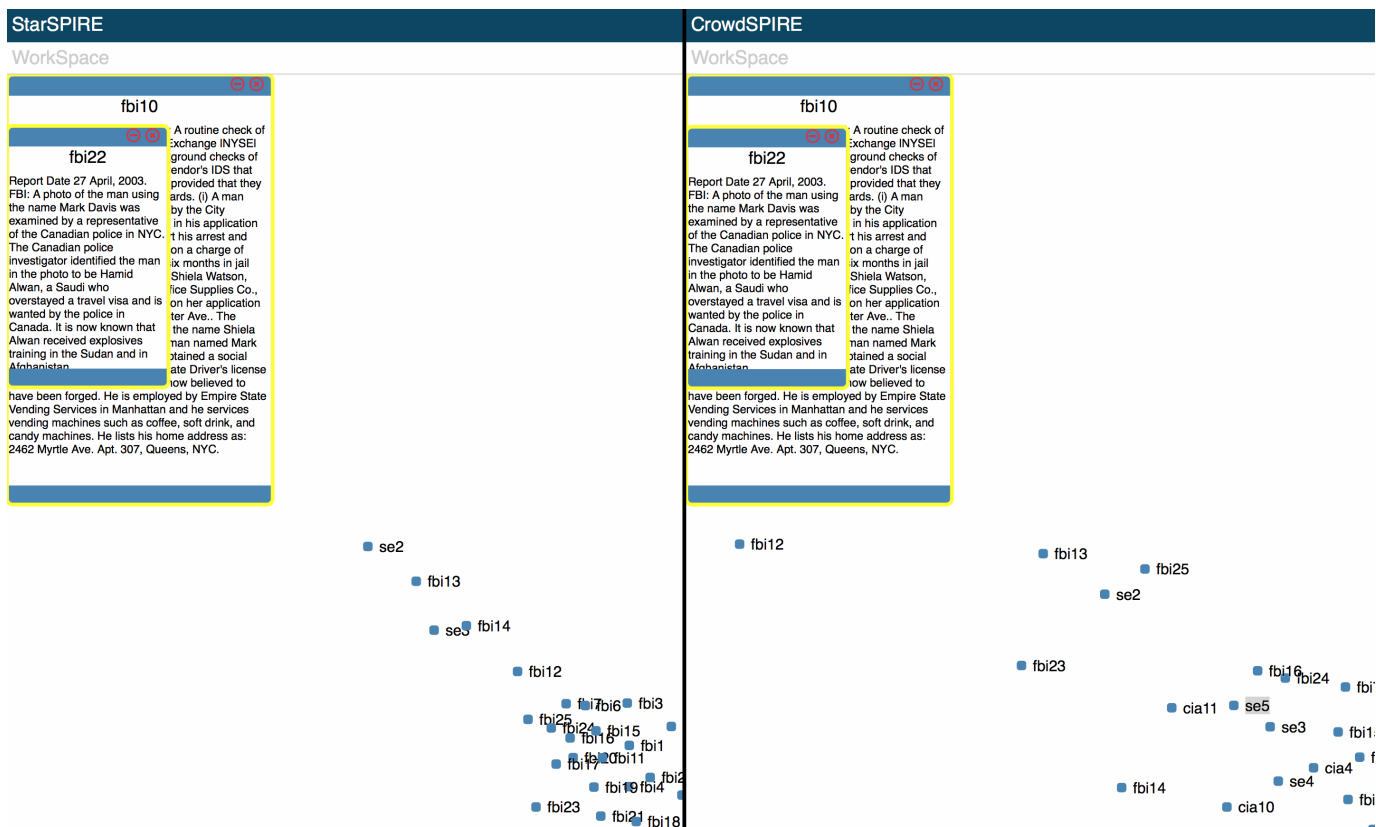


Fig. 5. A process of finding a major threat plot with key steps. (1): Based on A. Ramazi, finding that there are two similar bundles and two cells. (2): One name and two bundles are highlighted when hovering the mouse over B. Dhaliwal. (3): Three names and three bundles are highlighted when exploring F. Goba. (4) Referring to the four connected groups of useful entities for hypothesis generation.

subplots in the data. The users results were compared with the known ground truth solution. The user correctly identified four out of five subplots in the data. The user added 145 documents to the workspace, which is 1047 documents were opened and 33 remained open at the conclusion of the sensemaking session. The user made eight searches, four document annotations, and 21 highlights. 45 documents were added through searches, whereas the remaining 100 documents were added through other multi-scale semantic interactions (e.g. highlight, annotate, document proximity). Out of 26 documents relevant to the final solution, the user had added 18 of them to the workspace. Six of these 18 documents were added through an explicit search, while twelve were added through implicit multi-scale semantic interactions. Therefore, the documents that originated from multi-scale semantic interactions were similar in quality to those that originated from explicit searches from the user. Out of approximately 1,500 documents, 47 were read. Thus, the analyst was able to construct 80 While the results of this usage scenario appear promising, further work is required to evaluate the performance of MSI techniques as compared to existing SI techniques.

We get the conclusion that:

1. If Crowds could help remove noises from when sensemaking, to find most imports trifles/dots. 2. If Crowds could help group close dots together. 3. If Crowds could provides external knowledge that not list on documents.

1. Crowds could help remove the noise, that are irrelevant details, dots, or trifles, even they are closed related to important documents. three assignments contained little or no "noise" in the form of irrelevant details, dots, or trifles. I embedded in an array of irrelevant dots that exist in intelligence reports that will lead the students nowhere as far as the hypothesis that is suggested by the relevant dots.

2. Crowds could be helpful for simultaneous and coordinated terrorist activities involving three actions planned for 3. Provides knowledge that not included in documents. most important element of imaginative and productive intelligence analysis in real life.

have included a variety of irrelevant or distracter items. In short, skillful and thorough intelligence analysis requires that you carefully find out about what some dot or trifle is telling you. Such knowledge is not always, perhaps only rarely, revealed in the reports in which these dots or trifles are given to you. different persons will generate different hypotheses from the same body of evidence.

produces different insight? better insight??? compare to Gold Standard Solution beyond simple keywords, semantics similarities compare to previous user study cluster results?

How the assigned tasks is good to current tasks. Comparison of crowd-enhanced version with algorithm-only version produces different insight? better insight??? compare to Gold Standard Solution beyond simple keywords, semantics similarities compare to previous user study cluster results? Finally we find that .

6 CONCLUSION AND FUTURE WORK

We present crowd-powered semantic interaction model, a visual analytics model, which help analysts make sense of documents quickly through the help of crowdsourcing. In this model, we use semantic interaction to enable users to steer the crowdsourcing assignment implicitly instead of require users of domain knowledge for crowdsourcing. In addition, we provide several different ways to help combine the crowd outputs back into the workspace (visual interface) appropriately based on current visualization layout. Moreover, this model also takes the combination of human computation and automatic computation. That automatic computation methods, like, machine learning could be used to find more related data that can be used in micro-tasks. The outputs from crowds, could help update the layout of workspace directly, or indirectly through providing more knowledge for automatic computation methods. For example, more semantic links between documents or entities could help improve the correctness of machine learning algorithms, when doing clustering. With a usage scenario, we demonstrate how crowdsourcing can potentially support an analyst to explore complex sensemaking tasks. However, there are still three challenges that need further explorations. Current version of CrowdSPIRE only

contains the basic needed components of: Visualization, Analytic model and Crowd part "Connect the Dots". More works needed to be done on this parts:

C1: Find the most appropriate crowdsourcing tasks for current visual analytic system. In CrowdSPIRE, we combine the visual analytic system with machine learning algorithms, and basic crowdsourcing tasks "Connect the Dots". The connection between semantic interactions with "Connect the Dots" shows us crowdsourcing could provides help that machines are inadequacy to do right now. However, there are lots of crowdsourcing tasks that we could do for each interaction, which we must perform lots of experiments to find the best solution.

C2: Ways to integrate crowdsourcing output into workspace. Current ways of integrate crowdsourcing into workspace, is pre-performing all the needed micro-tasks, and store the outputs permanently. When some interactions trigger some specific micro-tasks, we just search for related micro-tasks results directly from database, instead of carry out a real-time HIT. Several other ways of carry out tasks should be considered: how to carry out real-time hits instead of the pre-stored results. More integrating ways needed to be explored, to find the best strategy to combine crowdsourcing and visual analytics.

C3: Comparison between crowdsourcing and machine learning algorithms. CrowdSPIRE give a simple demo on how could we combine crowdsourcing together with machine learning algorithms to help analysts' sensemaking. However, we must doing more research on finding which part is good at what kinds of tasks. So the semantic interaction could decide assign what kind of tasks to crowds and other tasks to machine learning algorithms.

REFERENCES

- [1] Amazon Mechanical Turk.
- [2] M. S. Bernstein, G. Little, R. C. Miller, B. Hartmann, M. S. Ackerman, D. R. Karger, D. Crowell, and K. Panovich. Soyent: a word processor with a crowd inside. *Proceedings of the 23rd annual ACM symposium on User interface software and technology*, pp. 313–322, 2010. doi: 10.1145/1866029.1866078
- [3] M. S. M. Bernstein, J. Brandt, R. C. Miller, and D. R. Karger. Crowds in Two Seconds: Enabling Realtime Crowd-powered Interfaces. *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, pp. 33–42, 2011. doi: 10.1145/2047196.2047201
- [4] J. P. Bigham, C. Jayant, H. Ji, G. Little, A. Miller, R. C. Miller, R. Miller, A. Tatarowicz, B. White, S. White, and T. Yeh. VizWiz: Nearly Real-Time Answers to Visual Questions. *UIST*, pp. 333–342, 2010. doi: 10.1145/1866029.1866080
- [5] Bigham, Jeffrey P, Jayant, Chandrika, Ji, Hanjie, Little, Greg, Miller, Andrew, Miller, Robert C, Miller, Robin, Tatarowicz, Aubrey, White, Brandyn, White, Samuel, and Yeh, Tom. VizWiz - nearly real-time answers to visual questions. *UIST*, p. 1, 2010.
- [6] L. B. Chilton, J. Kim, P. André, F. Cordeiro, J. A. Landay, D. S. Weld, S. P. Dow, R. C. Miller, and H. Zhang. Frenzy: Collaborative Data Organization for Creating Conference Sessions. *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, pp. 1255–1264, 2014. doi: 10.1145/2556288.2557375
- [7] L. B. Chilton, G. Little, D. Edge, D. S. Weld, and J. A. Landay. Cascade: Crowdsourcing taxonomy creation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1999–2008, 2013. doi: 10.1145/2470654.2466265
- [8] A. Endert. Semantic Interaction for Visual Analytics: Inferring Analytical Reasoning for Model Steering. *Synthesis Lectures on Visualization*, 4(2):1–99, Sept. 2016.
- [9] A. Endert, P. Fiaux, and C. North. Semantic Interaction for Sensemaking - Inferring Analytical Reasoning for Model Steering. *IEEE Trans. Vis. Comput. Graph.*, 2012.
- [10] D. H. Jeong, C. Ziemkiewicz, B. Fisher, W. Ribarsky, and R. Chang. iPCA: An Interactive System for PCA-based Visual Analytics. *Computer Graphics Forum*, 28(3):767–774, June 2009.
- [11] W. S. Lasecki, R. Wesley, J. Nichols, A. Kulkarni, J. F. Allen, and J. Bigham. Chorus: A Crowd-powered Conversational Assistant. *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, pp. 151–162, 2013. doi: 10.1145/2501988.2502057
- [12] E. Law and L. von Ahn. Human Computation. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 5(3):1–121, July 2011.

- [13] K. Luther, N. Hahn, S. P. Dow, and A. Kittur. Crowdlines: Supporting Synthesis of Diverse Information Sources through Crowdsourced Outlines. *Third AAAI Conference on Human Computation and Crowdsourcing*, pp. 110–119, 2015. doi: 10.1093/cybsec/tyv008
- [14] P. Pirolli and S. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. *Proceedings of International Conference on Intelligence Analysis*, 2005:2–4, 2005. doi: 10.1007/s13398-014-0173-7.2
- [15] J. Thomas and K. Cook. Illuminating the parth: the reserach and development agenda for Visual analytics. Technical Report 2, 2005. doi: 10.3389/fmicb.2011.00006
- [16] N.-C. Wang. Crowdnection - Connecting High-level Concepts with Historical Documents via Crowdsourcing. *CHI Extended Abstracts*, pp. 146–151, 2016.
- [17] P. Welinder and P. Perona. Online crowdsourcing: rating annotators and obtaining cost-effective labels. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pp. 25–32. IEEE, 2010.
- [18] H. Zhang, E. Law, R. C. Miller, K. Z. Gajos, D. C. Parkes, and E. Horvitz. Human Computation Tasks with Global Constraints. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*, pp. 217–226, 2012. doi: 10.1145/2207676.2207708