



Article

Reorganize Your Blogs: Supporting Blog Re-visitation with Natural Language Processing and Visualization

Shuo Niu ^{1,*}, D. Scott McCrickard ², Timothy L. Stelter ², Alan Dix ³ and G. Don Taylor ⁴

¹ Department of Mathematics and Computer Science, Clark University, Worcester, MA 01610, USA

² Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, USA; mccricks@cs.vt.edu (D.S.M.); tstelter@vt.edu (T.L.S.)

³ Computational Foundry, Swansea University, Bay Campus, Fabian Way, Swansea SA1 8EN, UK; alan@hcibook.com

⁴ Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24061, USA; don.taylor@vt.edu

* Correspondence: shniu@clarku.edu; Tel.: +1-540-998-8827

Received: 28 June 2019; Accepted: 19 September 2019; Published: 7 October 2019

Abstract: Temporally-connected personal blogs contain voluminous textual content, presenting challenges in re-visiting and reflecting on experiences. Other data repositories have benefited from natural language processing (NLP) and interactive visualizations (VIS) to support exploration, but little is known about how these techniques could be used with blogs to present experiences and support multimodal interaction with blogs, particularly for authors. This paper presents the effect of reorganization—reorganizing the large blog set with NLP and presenting abstract topics with VIS—to support novel re-visitation experiences to blogs. The BlogCloud tool, a blog re-visitation tool that reorganizes blog paragraphs around user-searched keywords, implements reorganization and similarity-based content grouping. Through a public use session with bloggers who wrote about extended hikes, we observed the effect of NLP-based reorganization in delivering novel re-visitation experiences. Findings suggest that the re-presented topics provide new reflection materials and re-visitation paths, enabling interaction with symbolic items in memory.

Keywords: blog; natural language processing; visualization; re-visitation

1. Introduction

Digital personal archives log life materials chronologically, necessitating the identification of novel user experiences in re-visiting them and generating new ideas [1–3]. Interaction designs supporting reading one’s own life-logs seek to regenerate prior experiences by encouraging meaningful interaction with photographs [1], videos [2], and geo-locations [4]. However, personal textual content such as blogs has long been under-studied in re-visitation [3] and generation of new perspectives about self [5]. The textual content takes time to read, especially when it grows massively large but minimally organized [6,7]. When bloggers need to re-visit and share stories from blogs, for example during themed conferences or meetings, the massive volume and textual nature of blogs makes it hard to understand core themes and capture an overview and connections.

People spend time crafting blogs, recording moments in life, and capturing essences of thoughts [8]. In 2017, blog authors on the WordPress crafted over 80 million new blog posts each month (a live look at activity across WordPress.com (September 2018): <https://wordpress.com/activity/>). Around 80% of the regular bloggers posted at least once per month, and over 90% of blogs had more than 500 words (Blogging Statistics and Trends: The 2017 Survey of 1000 Bloggers (September 2018); <https://www.orbitmedia.com/blog/blogging-statistics/>). Bloggers write about experiences over time, generating a temporally-ordered flow of experience [8,9]. Many blog authors have a desire

to reflect, but generally do not wish to re-read the entire collection [6]. Given the massive volume of content in many blogs, retrieving and organizing information about experiences cost effort of memory and time [10], while combination and reorganization of materials generates new ideas [11]. To support novel blog interaction, natural language processing (NLP) and visualization (VIS) provide opportunities to reorganize large blog content. Large touch displays can present the representations of personal contents dynamically to support multimodal interaction with blogs.

NLP and text visualization (VIS) offer new design materials to build multimodal experiences with large personal texts. This research reports our experiences inviting three long-term bloggers to re-visit their large personal blogs with *BlogCloud*, a system which implements blog **reorganization**—breaking the temporal order of blogs and recombining experience segments by natural language processing (NLP) techniques. Reorganization depicts a new design opportunity to utilize text processing techniques to support multimodal interaction with personal blogs, in contrast to their original forms of long textual information on web pages. To obtain an early understanding of user's reaction to the new form of personal content, we designed and implemented a blog re-visitation system, *BlogCloud*. We report our experiences taking a walk-through of the interaction in use context and report users' open-ended experience stories, which helped us collect meaningful feedback and identify new design opportunities with the reorganization [12,13]. We consider the blog re-visitation during themed public events such as workshops and exhibitions, to probe how the new re-visitation modality created by NLP and VIS could encourage re-discovery of one's own themed experiences. Concluded from our experience with *BlogCloud*, we identify that the *reorganized topics*—sets of related or similar content from personal blogs—offer new materials for reflection and new paths for re-visitation, enabling multimodal interaction with symbolic memories. This paper is organized as follow: we first review prior work on re-visitation and reflection on large personal content, with identification of key factors and theories in life-logging and re-visitation. Then, we introduce the opportunities to utilize natural language processing techniques to reorganize large personal blogs. Next, we introduce *BlogCloud*, a multimodal interactive system which operationalizes blog reorganization with natural language processing and visualization technologies. There after, we focus on our experiences with three long-term blog authors using *BlogCloud* during a themed public event, to describe new blog re-visitation manners with the reorganized content. Finally, our discussion and conclusion seek to suggest new design opportunities to enable interaction with the reorganized blog topics.

2. Related Work

Weblogs, or *blogs*, are an open online medium that allow people to journal their own experiences. Studies have investigated people's reasons for blogging, recognizing a wide spectrum of motivations (e.g., [6,8]). People blog to document life events, provide opinions, express emotions, articulate ideas, and form and maintain community forums [8]. Other studies also suggested self-presentation [14,15] and experience-sharing [16,17] could encourage blogging. Bloggers record their personal experiences by documenting remarkable items and activities, along with correlated opinions and emotions [8]. The life documentation and expression aspect of personal blogs makes them rich and valuable materials to reflect on the past and generate new ideas [1]. However, prior studies investigating the "reading" side of blogs noticed that many blog explorers tend to quickly skim through most recent posts [18,19]. Some readers appear to be casual and passive; they rarely reflect on knowledge gleaned, nor do they care about the freshness of the content [6]. These somewhat careless behaviors are reflected by posts appearing in reverse chronological order [8,20] that are updated over long periods of time [6] and are comprised of large volume of text and photos [21]. Following blogs is time-consuming, and it is hard to combine and reorganize related items in experiences [6,7,11]. During public events such as conferences, meetings, workshops, and exhibits, when people re-live and share themed experiences, the overwhelming texts make blogs a less effective material to combine and reorganize to deliver novel re-visitation experiences [11,22–24].

We focus on understanding whether the NLP-reorganized personal blogs could bring in novel experiences re-visiting the broken-and-recombined experiential items [11,25]. This paper provides a seminal understanding of how bloggers interact with a multimodal system augmented by NLP and VIS. Interactive systems encouraging re-visitation and reflection have been widely studied over the past decade [3,26]. One thread of studies explored people's daily practices in preserving reflective items, such as personal archives [27], mementos [28], and digital legacies [29]. Technologies for gathering and managing digital personal items, referred to as life-logging systems, are recommended for reminiscence and reflection [2,9,30,31]. With increasingly more personal content collected by technologies [9], another realm of research sought to understand people's interaction with digital past [1,32–35]. Fleck et al. identified re-visitation as a fundamental reflection activity that depends on the higher levels of reflection that require an explaining, connecting, and transforming experience [3].

Systems for experiencing re-visitation need to deal with the challenges of ever-growing, minimally-organized content [9,26,31]. Many prior studies describe re-visitation and reflection as common but complex psychological activities, in which people deliberate in-and-on actions and associate thoughts for new perspectives [25,26,36]. Ernst von Glasersfeld defined reflection on past experiences as an activity "to step out of the stream of direct experience, to re-present a chunk of it, and to look at it as though it were direct experience" [25]. People "segment and recombine the memorable parts at will" [25], to re-present as if a direct experience, and compare it with other experiences that are similar. Through reflective activities with personal content, people re-understand experiences and uncover new life patterns [26]. Other researchers have examined the roles that text, pictures, relationships, and randomness can have on the re-visitation experience [1,7,37]. Sellen et al. summarized this area of work with their "5Rs" (*recollecting, reminiscing, retrieving, reflecting, and remembering*) as potential design benefits of life-logging systems [30]. They put forth the research challenge to explore connections within and between these areas, which inspired our direction of study. Areas of focus in this research lie in identifying ways to trigger reminiscence, and to identify the connections and relationships between blogged items [3,32,34].

A promising way to address these issues with blog re-visitation is through natural language processing and visualization [4,35,38]. Studies suggest that aggregated personal content creates value in blog reading experiences [5,11,34,38]. Recent studies showed that visualization could raise re-living of personal events [4,31,39]. Technologies in natural language processing (NLP) and interactive information visualization make it possible to categorize and present similar textual items in an aggregated manner, and allow users to interactively suggest interests [40–42]. Using text-mining and content indexing could facilitate blog navigation, as prior studies found that blog readers would search information about specific interests, and prefer an overview to a wide variety of content [43–45]. Reordered and categorized content from one's own blogs may create unique multimodal interaction experiences, as studies interestingly found blog categorization widgets were voluntarily used by participants to navigate their own blogs [43,46]. Considering the textual format and large volume of blogs could hinder generating new ideas, this study examined using NLP and VIS to reorganize and present the blogs and design opportunities to augment multimodal life-log re-visitation.

3. Blog Reorganization for Re-Visitation

This study explored the *reorganization* of large personal blogs with text processing and visualization technologies, with the goal of understanding their roles in creating new modalities and interaction experiences from re-visitation. Natural language processing (NLP) and visualization (VIS) are promising approaches to break their original order and reorganize experiential items [3]. However, little is known about the opportunities of re-visiting one's own NLP-processed and VIS-presented long blogs, to engender new ideas about past experiences and open multimodal pathways for re-visitation. Our conceptual approach of *reorganization* with NLP reduces the effort to identify and retrieve related experiences. Exploring the computer reorganized contents and topics offers a high level view of the experiences. NLP and VIS support re-visitation by reducing the effort of organizing and exploring the

large blog volume. The derived topics may provide new perspectives from the massive data, which potentially increases the expressiveness of blogs and the enjoyment of re-visitation.

We examined two blog reorganization approaches to support re-visitation: NLP techniques break the original order of blogs; users can search keywords among blogs and re-visit blog sections that are similar to a particular experience (Figure 1). NLP techniques are applied on personal blogs to raise awareness of similar experiences, which match a symbolic word or relate to a particular section. NLP makes it technologically possible to search large blog texts and recognize relevant content by computing document similarity. Common NLP technologies provide methods to filter less meaningful words (i.e., stop words), identify keywords, vectorize textual data, and compute similarity. The processed blogs can be indexed to enable searching. NLP techniques can also automatically identify and group the blog content that are similar to a re-visited blog section. Contents that are identified from searching or grouped by NLP-calculated similarity constitute a new content group, which breaks the original order and reorganized into a new topic. A visualized topic contains similar or related experiences, which could illustrate the context of a past experience and offer a new interaction modal [31]. Interactive text visualizations can present inferential information of the topics, such as the number of similar segments and keywords of the topics, to inspire new perspectives and connect related experiences [7,11].

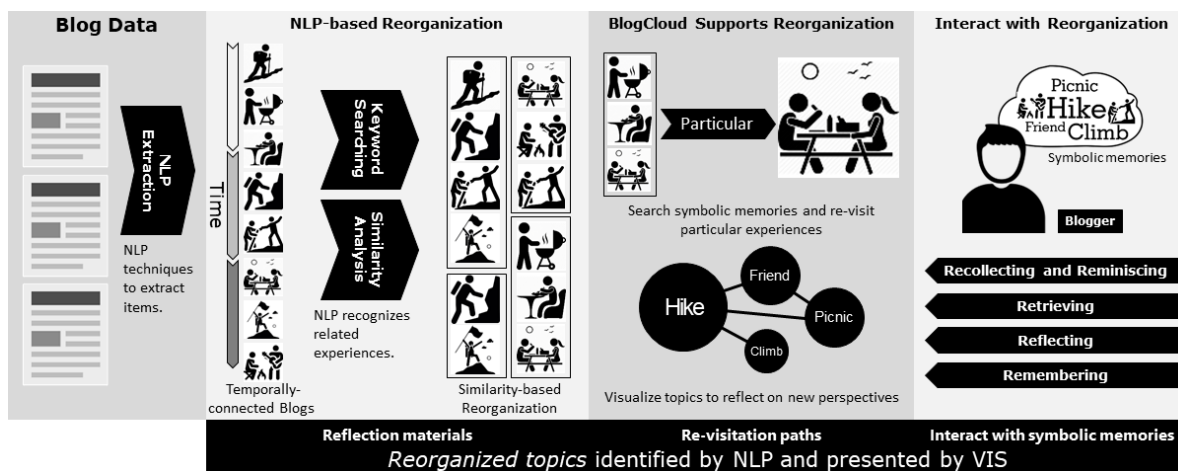


Figure 1. The conceptual description of *reorganization*. NLP breaks the original temporal order of blogs by keyword searching and similarity-based grouping. Bloggers can re-visit particular experiences and reflect on the visualized high-level topics. Reorganization leads to new manners to re-visit personal blogs.

4. BlogCloud: Blog Re-Visitation Tool

To explore the *reorganization* ability of natural language processing (NLP) and visualization (VIS) in supporting blog re-visitation, we designed and implemented *BlogCloud* (Figure 2). *BlogCloud* is an interactive blog exploration tool implemented on large displays, which allows the bloggers to search, read, and organize blog segments (the system is adapted from a prior system [47,48]). During public events such as a themed workshop, bloggers can walk up and use *BlogCloud* to search related keywords among their blog archives (Figure 2). After searching, retrieved blogs will be presented on a tabletop, while a word cloud visualization on the vertical display shows topic words, blog numbers, and sentiment information to support reflection on topics from the blogs. Users can iterate searching keywords and reflecting on the visualization to re-explore the blog collection.

4.1. *BlogCloud* on Large Interactive Displays

We considered re-visiting one's own personal blogs during public events, when bloggers interact with a walk-up-and-use tool [49] to reflect on the NLP- and VIS- augmented blogs. To explore

the multimodal design for this context, BlogCloud is implemented with large interactive displays: a multi-touch tabletop and a vertical display (Figure 2). We chose large interactive displays as the platform since they feature content manipulation with natural touch interaction [50]. Large displays also benefit reading activities by expanding visual and interaction space [51]. As blogs may contain multiple entries of the same topic, presenting related content in one interaction space affords comparison and reflection. Large displays allow users to allocate spaces to facilitate organization and comparing details [51]. With multi-display design, bloggers can shift between visiting particular blog entries and viewing high-level topics, with possible better perception to the NLP results.

4.2. System Implementation

although there are other ways to reorganize large blog collections, we chose to use keyword searching and similarity analysis as two common methods to explore the design benefits of reorganization. BlogCloud reorganizes blogs (Figure 1) and supports re-visitation by:

- Using natural language processing techniques to chunk the temporal blog stream and identify potentially important keywords.
- Allowing bloggers to denote symbolic memories and retrieve blogs by searching keywords.
- Supporting generation of related topics connected to the re-visited blog sections, by visualizing keywords and presenting inferential information of topics.

BlogCloud incorporates natural language processing (NLP) techniques to break the temporal stream of blogs, which allows searching and similarity-based grouping. To instantiate the concept, we chose three generic but common NLP techniques in text processing: keyword extraction, term-frequency based vectors and cosine-similarity clustering. Although other high-level NLP and machine learning techniques could be used, extracting key information and grouping documents reflect core functions of many NLP and machine learning based approaches.

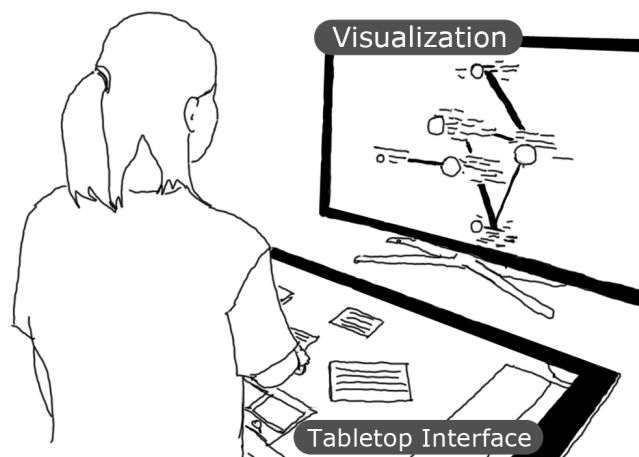


Figure 2. The BlogCloud system. The tabletop supports blog searching and viewing. The vertical display shows a visualization of the NLP-processed content.

A paragraph is used as the basic reflection section (see Figure 3) since paragraphs usually focus on a few interrelated experiential items that are lexically self-contained and map well for natural language processing. Text processing is running on a paragraph based collection. BlogCloud segments the words by spaces and punctuation. Stop-words (e.g., “a”, “an” and “the”) are eliminated from further processing. For each paragraph, nouns, verbs, adjectives and adverbs are extracted as keywords with coreNLP [52]. The weight of each keyword in a paragraph is calculated using term frequency-inverse document frequency (TF-IDF) [53], through which the importance of words are estimated. Segmentation and term weighting techniques enable the system to process words and prepare for blog reorganization. BlogCloud stores the weight of each keyword for each paragraph.

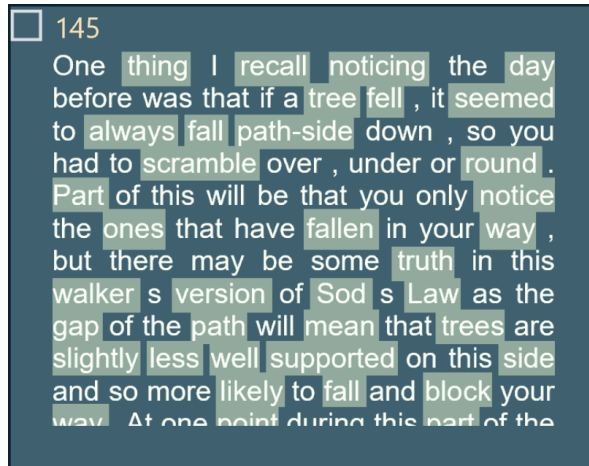


Figure 3. A paragraph card on tabletop interface. Words with background color are recognized keywords. This card illustrates an experience of mountain climbing (A1). Showing actual content allows re-visit experience details.

The text processing segments the blog stream, so that the users can search memories and read matched blog paragraphs (Figure 4). Grouping similar paragraphs makes a topic that is related to the re-visited blog section. The system first converts the entire blog set into a document-term matrix M . w_{ij} is the TF-IDF weight [53] of T_i in paragraph D_j . When one blog paragraph (VP) is viewed on a card, BlogCloud creates an array of terms D_V , which only contains the keywords that appear in VP . Then, the system creates a sub-matrix M' from M by selecting columns whose corresponding terms appear in D_V . The similarity comparison method used by BlogCloud is based on paragraph vectors and cosine similarity. For the re-visited paragraph (VP) and a paragraph (CP) that is being compared, the system decides similarity through two generic steps: (1) For VP , create a vector $\vec{V}_u = \langle w'_1, w'_1, \dots, w'_u \rangle$ where w'_1 to w'_u are the weights of the u unique keywords in D_V . The row c in M' is CP 's vector $\vec{V}_c = \langle w'_{1c}, w'_{2c}, \dots, w'_{uc} \rangle$. Since M' is extracted from D_V , \vec{V}_u and \vec{V}_c have the same size and each element represents the TF-IDF weights of the same word. (2) Calculate the cosine similarity [41] between \vec{V}_u and \vec{V}_c ; if the similarity score is above 0.5, the system determines CP is a paragraph that is similar to VP . All similar CP s form a *paragraph group*. For all paragraphs in this group, the system chooses the top m words with highest TF-IDF weights to generate the word cloud for VP . m varies between 10 and 20, with larger paragraph groups containing more words. To indicate the relatedness between a pair of paragraph groups, BlogCloud tracks the number of paragraphs that co-appear in both paragraph groups. The NLP methods and parameters are decided from multiple trials on a test blog set.

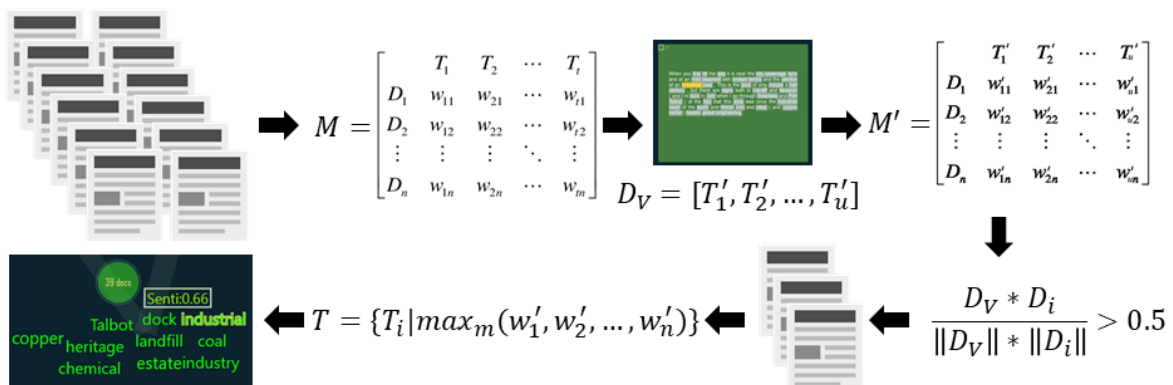


Figure 4. Blog processing pipeline to generate word cloud in the BlogCloud visualization.

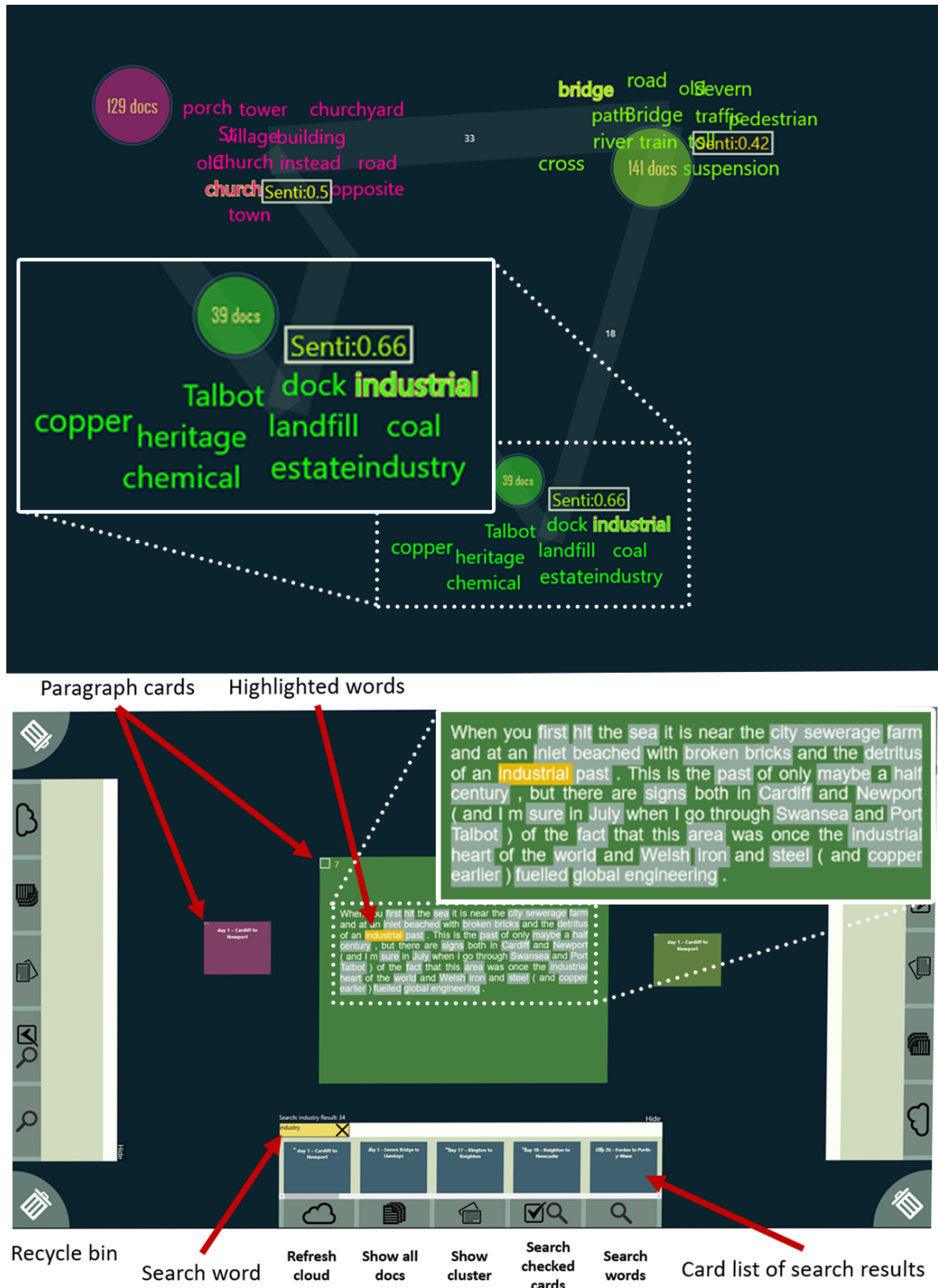


Figure 5. (Top) Connected word cloud visualizing the reorganization of blogs. The word cloud abstracts paragraph groups and connect topics to inspire new perspectives. (Bottom) Touch interface runs on the tabletop display. The three paragraph cards generate three word clouds in the vertical display. The card in middle is a paragraph about an industrial relic, and its corresponding word cloud is shown in the zoomed frame in the top figure.

The BlogCloud interface supports breaking the original time order of blogs through keyword searching and visualizing similar content, to present the *reorganization* of blogs. Users can search words with BlogCloud to retrieve a list of matched paragraphs. To support re-visiting particular experiences, digital cards presenting the paragraphs can be created and manipulated on the multi-touch tabletop (Figure 3). The cards can be moved, rotated and zoomed with multi-finger gestures. Zooming a card shows the actual blog paragraph content. Bloggers can search keywords in their memory to retrieve and read particular blog paragraphs. The search word is automatically highlighted on the paragraph card. When a paragraph card is placed on the tabletop, it triggers the NLP technique to identify other similar paragraphs to create a related topic (paragraph groups). A word cloud is visualized for each paragraph group created by the paragraph cards on the tabletop (see Figure 5). A word cloud contains four components about a topic: highlighted words, feature words, number of similar documents, and average sentiment values (The word cloud visualization in Figure 5). The highlighted words are the words being searched to retrieve the topic and presented with yellow borders. The size of the feature words in one word cloud is proportional to the weight of words in the paragraph group. A count of similar paragraphs in each paragraph group is presented in a circle. The connection between topics is represented by the number of paragraphs that co-appear between two paragraph groups. A strip with the number of associated blogs connects each cloud pair. The width of the strip is proportional to the number of blog paragraphs that exist in both paragraph groups. The summary of paragraph groups seeks to capture a high-level concept about different topics. The layout of word cloud and positioning of different topics are implemented by a force-directed graph. Sentiment scores, which reflect the feeling or tone of text and are calculated by Stanford NLP, range between 0 and 1, where 0 is entirely negative (red) and 1 is positive (green).

The tabletop interface contains a menu bar to use the searching function. *Search word* allows users to type a word or phrase with a virtual keyboard, displaying matched paragraphs in the scrollable paragraph card list (Figure 5, bottom). It allows bloggers to actively query symbolic words in their memory. With the menu bars on the sides of the tabletop interface, users can search for words, view results, and update the visualization (Figure 5, bottom). Paragraph cards in the card list can be dragged onto the table to read the actual blog content. *Show cluster* shows paragraph cards in each paragraph group. *Show all doc* loads all paragraph cards in the card list in chronological order. This button is added in case the users do not have anything to search for or want look at blogs in the original order. *Refresh cloud* button reloads the visualization.

5. Re-Visiting Personal Blogs

To evaluate how bloggers interact with BlogCloud and re-visit the NLP-reorganized content, we adopted the framework of “5Rs” identified by Sellen et al. [30]: *recollecting*, *reminiscing*, *retrieving*, *reflecting*, and *remembering*. We observed blog authors’ interaction with BlogCloud from each of the five re-visitation activities. Recollecting and reminiscing is when blog authors think back about experiences and re-live particular life events. Retrieving personal experiences depends on details in memory to find desired experiences for re-visitation. Reflecting on experiences requires examining patterns of past experiences and derive abstract representations. Remembering is a form of prospective memory, which is a memory for actions to be performed sometime in the future. In this case, remembering serves as cues to remember forgotten events. Re-visitation of large personal blogs needs to solve the high effort of organizing the content, the low expressiveness of unorganized text, and the enjoyment of reading large volume of data. We explored whether the NLP-reorganization could lead to new perspectives when exploring one’s own data, and support multimodal interaction experiences.

Grounded on the five re-visitation activities of life-logging systems, we sought to understand whether and how the NLP-based blog reorganization implemented by BlogCloud. BlogCloud is designed for bloggers to re-visit the computer reorganized life experiences with multimodal interaction. The system allows bloggers to search and read blogs, re-live specific life experiences and reminisce about emotional and sentimental traits. The searching function enables retrieving symbolic memories

by exploring different words among blogs. While reading the actual blogs, the visualization depicts the topics of the re-visited sections to support reflection. When multiple topics from the blogs are read together, the inferential information of topic sizes and connections between different topics afford reflection from a new perspective. When exploring a topic, bloggers could possibly encounter content that fades in memory and trigger remembering the experience. To examine the opportunities of natural language processing in supporting multimodal interaction with one's own large personal blogs, we report our experiences with long-distance hikers re-visiting their hiking blogs with BlogCloud.

6. Examining BlogCloud with "5 Rs"

In planning an examination of the NLP-based reorganization and how BlogCloud supports reading one's own large personal blogs, we recognized great value in gaining perspectives both from frequent bloggers and from interface and interaction experts. We identified three blog authors who also had relevant research skills in areas including human-computer interaction, user experience, ethnography, tech design, industrial systems engineering, and management systems to re-visit their blogs with BlogCloud (Figure 6). All of these blog authors have crafted a blog that describes a long-distance (over 1000 miles) hike that they undertook. Long-distance hiking was chosen as the blog theme because of its life significance and need for re-visitation: dedicated hikers prepare for months or even years for a long-distance endeavor, and they often blog, journal, or otherwise write about the hike's significance in their lives (e.g., Trail Journals <http://www.trailjournals.com/>, The Hiking Life <https://www.thehikinglife.com/> and [54,55]). During events when hikers are gathered, re-visiting and reflecting on such hiking experiences is worthwhile even years after the accomplishment, when memories may have faded and the enormity of the writings from the hike may appear even more daunting. Considering the challenges of measuring blog re-visitation through controlled lab studies (e.g., quantifying reflection and motivating sharing personal experiences) [56] and given the early stage of this research, we chose to engage in highly interactive and largely qualitative observation sessions with a small number of experts who maintain large personal blogs and also have relevant research knowledge. Observing blog authors' reflective activities and summarizing their feedback helped us obtain a descriptive understanding of the roles of NLP and VIS in multimodal interaction with blogs. The author sessions were conducted during an evening session at a hiking-themed workshop in which the authors were invited guests. During each session, blog authors were introduced to the system and encouraged to re-visit their extended hiking experiences, and share their hiking knowledge with others verbally so that the research team could understand their thought patterns. In each author session, we sought to observe how bloggers interact with the reorganized blogs presented by BlogCloud, with a focus on learning the following "5R" categories of re-visitation activities:

- How do bloggers **recollect** and **reminisce** about both high level ideas and specific experiences?
- How do bloggers interact with BlogCloud to **retrieve** symbolic memories from the reorganized experiences?
- How do bloggers perceive the topics derived from the re-visited blog paragraphs and **reflect** on the reorganized content?
- Whether bloggers' re-visitation to the reorganized blogs triggers **remembering** lost memories?

In the author session, each blog author spent 20–40 min using BlogCloud to explore his or her own hiking blogs. During the workshop, other attendees occasionally joined the conversation and engaged with the system. However, the blog author remained the primary user of the system throughout the session. A graduate student served as the exhibitor to introduce the system and explain operations. The qualitative data were collected through the conversation with blog authors while they were interacting with the system, and emails afterwards discussed feedback to the system.

Blog author 1 (A1) is a male HCI researcher and a dedicated hiker who keeps a personal hiking website with more than 100 hiking-related blog posts since 2012. In A1's online blog set, there are 97 blog posts with 3999 paragraphs and 230,159 words in total about a 1000-mile hike that he undertook.

Blog author 2 (A2) is a female HCI researcher, consultant, university instructor, and ethnographer who maintains a personal website that contains her hiking blog posts. During one 2500-mile hike, A2 recorded her hiking experiences and posted on her personal website. This blog set has 34 posts, which contain 467 paragraphs and 23,663 words. Blog author 3 (A3) is a male university professor and a section hiker (a hiker who breaks a long trail into sections). Starting in 2007, A3 started hiking sections of a 2000-mile trail, and wrote one post for each section he hiked. Cumulatively, he posted 213 blogs with 1904 paragraphs and 1,536,054 words.



Figure 6. A1 uses the BlogCloud to reflect on his own blogs. The visualization shows a word cloud of blog entries.

7. Findings

Blog authors acknowledged that it was a unique and valuable experience using BlogCloud to explore their own blog collection. All three blog authors noted they did not carefully organize the blogs and tweak the writing, and all had an unfulfilled wish to clean up and better organize their blogs when they have time (even though all authors completed their hikes several years prior to our session). They commented that the system helped them re-visit and organize thoughts about their blogs, which implies potential extended utility of the blog reorganization.

7.1. Recollecting and Reminiscing

Recollecting and reminiscing are interactions with life-logs, which lead to “re-living” specific experiences and reminiscing about the past experiences. During the interaction with BlogCloud, authors reminisced about particular experiences that connect to their memories and recollect experience topics from a non-temporal dimension.

7.1.1. Reminiscing about Particular Experiences

Through word searching, matched experiences are grouped by BlogCloud, avoiding skimming through the most-recent posts. During the author session, blog authors reminisced about particular experiences, meanwhile connecting them to the symbolic words that represent symbolic memories. When A1 searched “climb” (see Figure 7), he explored the first few results and picked a paragraph about a gradually narrowing cliff path. A1 reminisced about the experience on the cliff path by telling a story: “It was a path on the cliff. The first time when I passed it, it was this wide”, A1 opened his arms to show how wide it was. “But the second time when I was there it was barely my feet’s width. I have to pass it with my body clinging to the cliff and be careful.” The particular experience of passing a terrifying cliff path reorganized this kind of climb experience. A2 noted that reflecting on the word cloud surfaced a memory about her experiences: “the connections I got to explore and the way that the semantic connections

surfaced specific memories... it was a really cool experience to see words connected that acted as triggers to surface memories that I otherwise would not necessarily have been thinking about." Reading the visualization raised the reminiscence about particular experiences, which suggests that re-visiting a general experience topic turns into reminiscing about a particular one [25]. However, without the NLP reorganization, the experience might be neglected in the long blog stream, with fewer chances to encounter a particular experience and reconstruct that type of experience.



Figure 7. Word cloud of "climb" generated by A1.

7.1.2. Crossing the Time Dimension

We notice the broken-and-recombined experiential items bring forth a new perspectives to perceive blogs by recollecting multiple experiences across multiple days and merging them into a topic or story. A1 re-visited multiple paragraphs from different days after each search, trying to interpret whether the content matched his internal conceptualization of the tough nature of his experiences. A2 searched for the word "picnic", and after viewing the visualization (Figure 8) and several paragraph cards, she reminisced about two specific experiences on two different days: "[Interacting with the system reminded me of] a very specific memory about ... a nearby store that had non-backpack-friendly snacks available for purchase" and another time "... at a picnic table in time for the luxury of a place to sit-down while eating a less exciting rehydrated meal." BlogCloud helped A2 bring two experiences together, explaining how hiking food needs to be prepared for hikers. A3 noted he wanted to view content about different topics, commenting "It (BlogCloud) allows me to compare different things. This tool can help me organize my thought, like if I am thinking about writing a book about section hiking". The keywords in the visualization were abstracted high-level descriptions of a general idea, but the actual content comes from content at different time. Our blog authors appreciated viewing experiences across time, enabling recollection of similar items that are sparsely distributed.



Figure 8. Word cloud of "picnic" generated by A2.

7.2. Retrieving

BlogCloud uses user-defined keywords to retrieve and reorganize the blog content. In our observation of BlogCloud, we noticed the importance of the symbolic words in bloggers' memory. Based on blog authors' choice of words to describe memorable experiences, we identify the patterns of keywords usage when retrieving symbolic memories.

7.2.1. Active Experience Seeking

During the author session, blog authors were seeking familiar experiences with BlogCloud. Although BlogCloud allows blog viewing with the show-all button (a non-reorganized presentation), blog authors rarely selected paragraphs from their original order, using the search tool to search their own blogs, in contrast with the skimming and random reading observed in prior studies [6,18]. All three blog authors spontaneously identified multiple interesting words to interact with BlogCloud. All three authors could explain why they wanted to search for these words to retrieve interesting memories, which direct the interaction with BlogCloud. Rather than passively or randomly re-visiting the blog paragraphs, blog authors used BlogCloud in an active manner: they actively asked questions to the system to seek the reorganized content related to their memories.

7.2.2. Using Alternative Search Words

Blog authors used search words that contrasted common semantics. The NLP used in BlogCloud introduced bias interpreting author semantics in the context of their experiences. A1 wrote about a "bus shed" in his blogs, an uncommon term for observers. He noted *"it is my own geographical neologism...I was using it to mean something different as a parallel to a 'water shed'. It is not so much a single landmark as a generic kind of feature, but boundaries which are easy to walk across, but hard to return by public transport"*. Similarly, A2 used "picnic" in her blog, but in contrast to "picnic" in common sense. Her use of picnic referred to "non-backpack-friendly snacks" and "rehydrated food". Identifying unique semantics of words which contrast with ordinary meaning but importantly represent experiences is a challenge for natural language processing technique of systems which support re-visitation of large personal texts.

The authors attempted to identify alternatives to imprecise search words. When realizing their words were semantically different from ones found in blog paragraphs, blog authors sought alternative words to describe their experiences. After realizing results for the searched term "tear" was not related to sadness, A1 came up with another search word "cry" as an alternative. However, again, he noticed the word "cry" was mostly used to describe animal sounds. A3 searched the phrase "New Jersey" but noticed results only matching "new", so he searched "Jersey". These issues imply that blog authors had particular topics to explore that were not well retrieved, so they had to refine and adjust their search words or acknowledge failure for a topic. Text processing techniques should reflect alternative word meaning, and interactive visualization should guide bloggers for clarification when necessary.

7.3. Reflecting

BlogCloud derives topics from the re-visited paragraphs and conducts reorganization of related contents. It presents similar contents as topics in the visualization, to provide abstract representations of blogs and support reflection. During the author session, the topics identified from the entire blog collections led to alternative ways to re-visit hiking experiences.

7.3.1. Making Sense of Topics

BlogCloud offers a different view of experiences to trigger reflection. BlogCloud reorganize experiences by similarity. Information about experience topics can be extracted to inspire new perspectives. In the author sessions, blog authors undertook sensemaking activities with their own blogs. A1 found sensemaking important for reflection on personal content. He commented, *"although you have written the blog, it does not mean you fully grasp the big themes that run through it ... hence the*

need for sense-making tools.” A2 searched several words and reflected on the overview of multiple experiences. She commented “the connections were the interesting part, much more than just reflecting on some experience alone.”. The visualization of paragraph groups was used by A3 to make sense of symbols as a way to reflect on his own blog from a different perspective. A3 used the visualization to compare recurring experiences; he searched “difficult”, “rocky”, “steep”, and “hot”, creating four corresponding paragraph groups. After reading the connecting strips (Figure 9), A3 commented, “it makes sense to me since ‘steep’ is more related to ‘rocky’ and ‘difficult’ than ‘hot’”. When seeing “bike” had more documents than “scooter”, he commented “[the visualization] reflects that I used bike more than scooter for transportation”. When speculating that “cold” was mentioned more than “hot” in his blogs, he paused for a while and said, “‘cold’ is bigger than ‘hot’, hmm... maybe that is because it took me more time hiking the colder north part, than the warmer south.”. Making sense of the visualized blog reorganization leads to new perspectives and deeper reflections about the experiences.

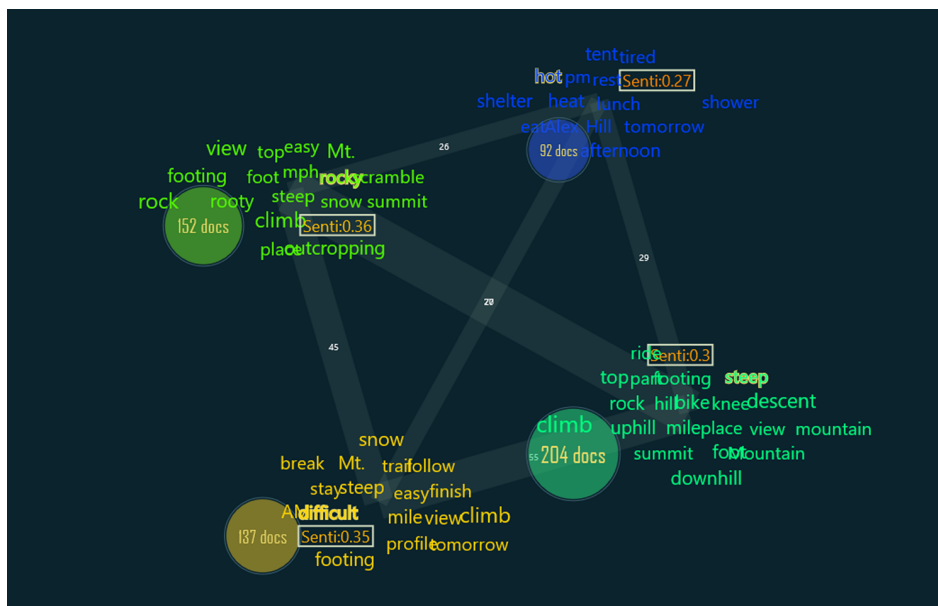


Figure 9. Four word clouds created by A3.

7.3.2. Connecting Related Topics

Blog authors associated related symbolic memories during reflection. Besides perceiving experience overviews and making sense of experience aspects, blog authors were also interested in delving into the connections between different experience topics. Seeing connections between word clouds of different topics helped A2 interpret related experiential items, noting: “it would be good if I could see the words between two groups, like on the table, or in the visualization, so I know how the words relate to each group, and know what things fall between [the two word clouds].” The connection from one topic to another could also be a flow of recalling consecutive experiences. When reading blog paragraphs about “bramble”, A1 remembered on that day there was one “bus-shed”, even though the paragraph was not directly describing it, and he explained his unique definition of “bus shed”. Connecting groups of reorganized experiential items sheds light on design improvements: besides showing abstracted information, the visualization should reflect how different experience topics are connected.

7.4. Remembering

Remembering in Sellen et al.’s definition emphasizes suggesting prospective events [30]. However, blogs store experiences along life and contain large volume of content, therefore memory might fail even with the blogs created by oneself.

7.4.1. Recovering Lost Memories

Experience documentation has been recognized as the most prevalent purpose for blogging [8,57]. However, considering the large number of blog posts and time passed since creation, experiential items might be lost in memory. During multiple rounds of searching and reading the visualization, blog authors have chances to explore content they did not remember well. We notice all three blog authors found words in the visualization or paragraphs on cards that they did not immediately recognize. However, blog authors did not skip these clues; they would collect more information to recover their lost memory. For example, when the visualization showed “dead” and “porpoise” as keywords in one paragraph group, one observer asked the author whether seeing a “dead porpoise” made A1 sad? A1 could not remember where he wrote about a porpoise, but a search revealed that he once thought the cracked wood looked like the “snout of a dead porpoise”. As another example, A3 searched “Jersey”, and several people’s names were displayed. Through reading the paragraphs related to the names, he recalled hikers he met in New Jersey. In addition, efforts to recover failed memories were also reflected through their suggestions for system improvement. A1 realized he could not find consecutive paragraphs with BlogCloud, so he recommended to add a “before” and “after” button to help transition gently from searching to reading. Similarly, A2 suggested an additional layer to show the whole blog entry. To recover lost memories, blog exploration tools should not only present the blog sections, but also give enough contextual information to help memory recovery.

8. Discussion

The extended author sessions demonstrate how the reorganization of blogs can support interaction with large blogs and leads to novel re-visitation experiences (as illustrated by Figure 1). With support from NLP and VIS in BlogCloud, blog authors interpreted experiences independent of time. Rooted in the findings, we discuss opportunities to identify and incorporate NLP and VIS in systems to enable interaction with the reorganized topics from one’s own large personal blog data.

8.1. New Materials for Reflection

Natural language processing for re-visiting blogs can break apart the time stream, clustering blogs into reorganized topics with non-temporal associations that often are not clear in temporally ordered blogs, thus offering new materials for multimodal interaction with blogs. In this work, *reorganized topics* are sets of similar, related blog sections from large personal blogs that, in their traditional form, highlight bloggers’ temporally-ordered experiences. Notable experiences are not isolated and involve abstraction and comparison of multiple related experiences. Instead of re-visiting a single experience, NLP can collect experiential items from different times that have never been reflected upon together, combining them as *reorganized topics* to offer novel materials for reflection. From our observation, the minimal interpretability nature of NLP raised an anticipation to examine the NLP-driven results; e.g., A1 and A2 actively interpreted the keywords from the word clouds and A3 deliberated on the numerical information in the visualization. The timeline-breaking ability of NLP produced reorganized topics. Not knowing what will emerge from NLP led to active searching and reflection. Our author sessions imply that, for re-visitation, blog exploration systems should consider the ability of natural language processing to break the order of personal content and recombine related experiences, thus making reflection materials more expressive and triggering curiosity about the reorganized topics [58]. The emerging NLP and machine learning techniques (e.g., NMF and LDA topic modeling) techniques could be utilized to reorganize and model the personal content into similar topics, which could encourage bloggers to identify or generate new perspectives.

8.2. New Paths for Re-Visitation

Visualization of reorganized topics encourages generating new perspectives, providing new paths for blog re-visitation. The NLP- and VIS-augmented content leads to unique paths of

re-visitation in recollecting, reminiscing, retrieving, reflecting, and remembering (“5 Rs”) with blogs [30], expanding the paths to explore new ideas and perspectives [11,58]. In the author session, blog authors started with words describing general feelings or common activities. They reflected on matched or similar experiences by re-visiting NLP-reorganized topics, and selected examples to validate their symbolic memories. In contrast to skimming through the chronologically-ordered blogs, those activities match the mental process of reflection; to reflect on a general idea, people segment experiences and reminisce about particular experience [25]. The visualization of BlogCloud suggests things to search, and reading a particular experience reconstructs general ideas about that type of experience. Visualizing reorganized topics that contradict bloggers’ temporal perception of experiences leads to sensemaking of the similarity-connected experiential items and connecting different experience categories, as alternative paths for reflection. Designers should consider using NLP techniques in future designs to shuffle the order of blogs, increasing the chance to read blogs that are not often re-visited. Re-visiting unrecognized experiential items encourages recovering lost memory and provokes interactions to recover experiences that are dormant in memory. Blog exploration systems should consider visualizations to develop new re-visitation paths such as sensemaking, reminiscing about feelings by re-visiting particular experiences, and encountering less re-visited experiences.

8.3. Interaction with Reorganized Topics

Symbolic words describing symbolic memories should center the identification of reorganized topics, to enable multimodal interaction with memories. Symbolic memories are impressions about experiences, represented by words and descriptions of experiences, that come to bloggers’ mind and express the need for reflection [58]. The reorganization should consider the interactiveness of words in bloggers’ memory: they provide clues for NLP to better retrieve blogs and trigger interactions with the experiences. Symbolic memories are not common and standalone experiential items. They are most meaningful experiences that form general feelings and remembered by bloggers. Three blog authors were focused on particular symbols. Blog authors were interested in retrieving reorganized topics related to symbolic memories when interacting with BlogCloud, leading to a novel re-visitation experience. Technologies supporting reflection of large textual datasets should actively allow users to express their symbolic words. Nevertheless, there are challenges in adopting symbolic words in natural language processing. Symbolic memories rooted in people’s experiences are highly specific and contextual, and people want know how they connect contextually in the blog. Falsely interpreted user symbolic words may confuse bloggers, resulting in dead-end searches. A resolution opportunity could use the user’s ability to find alternate symbolic words. New system designs should consider how to visualize symbolic content to indicate the context, as well as provide hints and clues to encourage users to better describe symbolic memories for blog retrieval. Future design for reflection should consider meaningful ways to capture symbolic experiences in memory, enabling rich interaction with symbolic memories—supporting reflection and avoiding confusion regarding NLP results.

8.4. Towards Better Natural Language Processing and Visualization for Re-Visitation

From our experience of observing bloggers using BlogCloud, we noticed that the reorganization of blogs implemented by NLP and VIS delivered novel multimodal blog re-visitation experiences. There is a trend that intelligent technologies such as machine learning are getting more human-focused and open to non-expert users [59,60]. Advancement in text processing and artificial intelligence will definitely introduce richer and more flexible ways to reorganize large personal content. We see a promising design space to introduce and explore the state-of-the-art information modeling and analysis technologies such as machine-learning and topic modeling to support better extract reorganized topics. Design opportunities raised from the identification of user’s symbolic memories can be applied to other forms of personal content such as pictures and videos, as well as other personal archives such as Twitter and Facebook posts. Future designs can incorporate different manners to collect symbolic words, such as iteratively and interactively asking users to suggest things they know (or do not

know) about their own experiences. From high-level but vague memories about large blog data to the reorganized and against-expectation topics for reflection, NLP and VIS technologies provide new design possibilities to deliver new re-visitation experiences. With advanced data processing methods and alternative interaction techniques, there is an open space to develop rich *reorganization* methods on large personal content to augment re-visitation.

9. Conclusions and Future Directions

In this paper, the term *reorganization* describes a pathway to incorporate natural language processing and visualization in support of re-visitation of large personal blog repositories (see Figure 1). While AI and other automated data-processing approaches are increasingly incorporated to support multi-model interaction, little attention has been drawn to use automation for personal content re-visitation. This work presents our experience observing three long-term bloggers re-visiting personal blogs with computer-organized content, seeking to provide an initial understanding of how automation supports re-visitation. The author sessions investigated how the re-organized blogs lead to new ways to interact with blogs, toward understanding one's own past experiences. From the experience with BlogCloud, we identify that *reorganized topics* offer reflection materials and build new re-visitation paths, and they allow bloggers to interact with symbolic memories. The ever-growing personal content is becoming a valuable repository for people to re-visit and reflect on previous experiences, but following this is the challenge of designing interactive systems for identifying values and creating meanings from large personal datasets [1,9]. Natural language processing as a material for design could be used to bridge the gap between the massive personal content and human's need to re-visit and reflect on meaningful experiences [59]. From the findings of this work, NLP and machine learning techniques can be incorporated in future multi-modality design. Designers should consider intelligent systems to help bloggers identify topics from one's own experience, re-visit experience from anti-time-line paths, and interact with memory symbols with the re-visitation tools and systems.

Certainly this work suggests much-needed follow-up efforts. The small participant number and brief author session was useful in delving deeply into usability issues, but questions remain about generalizability of findings. Similar to other exploratory studies on emerging issues and opportunities in designs for reflection [1,4], we focused on qualitative understanding of how bloggers interact with a multimodal design which presents the NLP reorganized personal blogs to support re-visitation of past experiences. While we have only a small number of participants, we selected three blog authors whom we found have valuable knowledge about large personal blogs as well as usability and engineering. Short but focused sessions with them helped to probe whether the *reorganization* leads to creative re-visiting activities, and what user experience would be expected. This approach can better help capture the opportunities and issues with NLP-based re-visitation than other controlled lab scenarios (e.g., a lab study that recruits participants to read irrelevant blogs). Through the sessions with the dedicated bloggers, we investigated the re-visitation interaction and issues that emerged from the reorganized blog content. With the initial findings from this work, follow-up research needs to consider how to support diverse reading activities with different reader types and reading interests, for example, new readers read other's blogs, bloggers reflect on blogs of different themes, or users read blog collections created by multiple bloggers. In addition, BlogCloud is implemented on large multi-touch displays, which are not universally available to all bloggers. In our study, large displays offer benefits for managing large blog representations and supporting group discussion; however, whether such experience could be generated when interacting with personal devices calls for future investigation. Future work will consider adapt the system to personal devices such as computers or tablets.

Moving forward, we foresee new possibilities for this work in three directions. First, we will seek better interactive designs that encourage people to express their meaningful symbolic words, and collect user's interests implicitly, by using computational approaches to recognize promising reflective topics. Second, alternative visual and interactive designs that support reorganization will be

explored to enable more ways of topic interaction. Re-visitation with other devices such as phones, tablets, or personal computers calls for better design to support interaction with large personal texts. How the new re-visitation paths brought by NLP and VIS could be seamlessly integrated to the existing chronological-ordered blogs requires investigation. It is important to support meaning back-and-forth navigation between high-level topics and blog details. Third, the collaborative understanding of personal blog systems, at collocated events such as conferences and group talks, include possibilities to reflect on and exchange experience topics through experience sharing and collocated re-visitation. New data mining and visualization techniques need to be designed and applied in collaborative tools to support collocated reflection, leading to identification of connections between experience topics of different people.

Author Contributions: Conceptualization, S.N., D.S.M., A.D. and G.D.T.; Data curation, S.N. and T.L.S.; Formal analysis, S.N.; Investigation, S.N.; Methodology, S.N.; Project administration, S.N.; Resources, S.N.; Supervision, D.S.M.; Writing—original draft, S.N.; Writing—review & editing, D.S.M.

Acknowledgments: Thanks to Ellie Harmon for sharing her hiking experience and discussing the system design. Thanks to the Center for Human Computer Interaction, the Department of Computer Science, and the Institute for Creativity, Arts, and Technology for their support of the initiative and the workshop that made this work possible.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Odom, W.; Sellen, A.J.; Banks, R.; Kirk, D.S.; Regan, T.; Selby, M.; Forlizzi, J.L.; Zimmerman, J. Designing for slowness, anticipation and re-visitation. In Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems—CHI '14, Toronto, ON, Canada, 26 April–1 May 2014; ACM: New York, NY, USA, 2014; pp. 1961–1970. doi:10.1145/2556288.2557178.
- Sellen, A.J.; Fogg, A.; Aitken, M.; Hodges, S.; Rother, C.; Wood, K. Do life-logging technologies support memory for the past? In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI '07, San Jose, CA, USA, 28 April–3 May 2007; ACM: New York, NY, USA, 2007; p. 81. doi:10.1145/1240624.1240636.
- Fleck, R.; Fitzpatrick, G. Reflecting on Reflection: Framing a Design Landscape. In Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction OZCHI '10, Brisbane, Australia, 22–26 November 2010; ACM: New York, NY, USA, 2010; pp. 216–223. doi:10.1145/1952222.1952269.
- Thudt, A.; Baur, D.; Huron, S.; Carpendale, S. Visual Mementos: Reflecting Memories with Personal Data. *IEEE Trans. Vis. Comput. Graph.* **2016**, *22*, 369–378. doi:10.1109/TVCG.2015.2467831.
- Jain, A.; Lupfer, N.; Qu, Y.; Linder, R.; Kerne, A.; Smith, S.M. Evaluating TweetBubble with Ideation Metrics of Exploratory Browsing. In Proceedings of the 2015 ACM SIGCHI Conference on C&C '15 Creativity and Cognition, Glasgow, UK, 22–25 June 2015; ACM: New York, NY, USA, 2015; pp. 53–62. doi:10.1145/2757226.2757239.
- Baumer, E.P.; Sueyoshi, M.; Tomlinson, B. Exploring the role of the reader in the activity of blogging. In Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems—CHI '08, Florence, Italy, 5–10 April 2008; ACM: New York, NY, USA, 2008; p. 1111. doi:10.1145/1357054.1357228.
- Vassileva, J.; Gutwin, C. Exploring blog archives with interactive visualization. In Proceedings of the Working Conference on Advanced Visual Interfaces—AVI '08, Napoli, Italy, 28–30 May 2008; ACM: New York, NY, USA, 2008; p. 39. doi:10.1145/1385569.1385578.
- Nardi, B.A.; Schiano, D.J.; Gumbrecht, M.; Swartz, L. Why we blog. *Commun. ACM* **2004**, *47*, 41. doi:10.1145/1035134.1035163.
- Jones, W. Building a Better World with our Information: The Future of Personal Information Management, Part 3. *Synth. Lect. Inf. Concepts Retr. Serv.* **2015**, *7*, 1–203. doi:10.2200/S00653ED1V01Y201506ICR042.
- Nardi, B.A.; Schiano, D.J.; Gumbrecht, M. Blogging as social activity, or, would you let 900 million people read your diary? In Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work—CSCW '04, Chicago, IL, USA, 6–10 November 2004; ACM: New York, NY, USA, 2004; p. 222. doi:10.1145/1031607.1031643.

11. Baugham, W.A.; Mumford, M.D. Process-Analytic Models of creative Capacities: Operations Influencing the Combination-and-Reorganization Process. *Creat. Res. J.* **1995**, *8*, 37–62. doi:10.1207/s15326934crj0801_4.
12. Yogasara, T.; Popovic, V.; Kraal, B.J.; Chamorro-Koc, M. General characteristics of anticipated user experience (AUX) with interactive products. In Proceedings of the 4th World Conference on Design Research: Diversity and Unity (IASDR2011), Delft, The Netherlands, 31 October–4 November 2011; Roozenburg, N., Chen, L.L., Stappers, P.J., Eds.; IASDR Delft University of Technology: Delft, The Netherlands, 2011; pp. 1–11.
13. Korhonen, H.; Arrasvuori, J.; Väänänen-Vainio-Mattila, K. Let Users Tell the Story: Evaluating User Experience with Experience Reports. In Proceedings of the CHI '10 Extended Abstracts on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; ACM: New York, NY, USA, 2010; pp. 4051–4056. doi:10.1145/1753846.1754101.
14. Schau, H.J.; Gilly, M.C. We are what we post? Self- presentation in personal web space. *J. Consum. Res.* **2003**, *30*, 385–404. doi:10.1086/378616.
15. Jolly, J.L.; Matthews, M.S. Why We Blog: Homeschooling Mothers of Gifted Children. *Roeper Rev.* **2017**, *39*, 112–120. doi:10.1080/02783193.2017.1289579.
16. Ab Rahman, N.A.F.W. Why We Blog? The Malaysian Travel Blogger and Tourist Perspective. *J. Tour.* **2017**, *2*, 12–28.
17. Hsu, C.L.; Lin, J.C.C. Acceptance of Blog Usage: The Role of Technology Acceptance, Social Influence and Knowledge Sharing Motivation. *Inf. Manag.* **2008**, *15*, 65–74. doi:10.1016/j.im.2007.11.001.
18. Narayan, S.; Cheshire, C. Not too long to read: The tldr interface for exploring and navigating large-scale discussion spaces. In Proceedings of the 2010 43rd Hawaii International Conference on System Sciences, Honolulu, HI, USA, 5–8 January 2010. doi:10.1109/HICSS.2010.288.
19. Zinman, A.R. Me, Myself, and My Hyperego: Understanding People through the Aggregation of Their Digital Footprints. Ph.D. Thesis, Massachusetts Institute of Technology, School of Architecture and Planning, Massachusetts, MA, USA, 2011.
20. Herring, S.C.; Kouper, I.; Paolillo, J.C.; Scheidt, L.A.; Tyworth, M.; Welsch, P.; Wright, E.; Yu, N. Conversations in the Blogosphere: An Analysis “From the Bottom Up”. In Proceedings of the 38th Annual Hawaii International Conference on System Sciences, Big Island, HI, USA, 6 January 2005; IEEE: Big Island, HI, USA, 2005; p. 107b. doi:10.1109/HICSS.2005.167.
21. Fullwood, C.; Sheehan, N.; Nicholls, W. Blog Function Revisited: A Content Analysis of MySpace Blogs. *CyberPsychol. Behav.* **2009**, *12*, 685–689. doi:10.1089/cpb.2009.0138.
22. Hornecker, E.; Stifter, M. Learning from Interactive Museum Installations About Interaction Design for Public Settings. In Proceedings of the 18th Australia Conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments, Sydney, Australia, 20–24 November 2006; ACM: New York, NY, USA, 2006; pp. 135–142. doi:10.1145/1228175.1228201.
23. Black, L.W. Blog, chat, edit, text, or tweet? Using online tools to advance adult civic engagement. *New Dir. Adult Contin. Educ.* **2012**, *2012*, 71–79. doi:10.1002/ace.20028.
24. Karger, D.R.; Quan, D. What would it mean to blog on the semantic web? In *The Semantic Web—ISWC 2004*; McIlraith, S.A., Plexousakis, D., van Harmelen, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3, pp. 147–157. doi:10.1016/j.websem.2005.06.002.
25. von Glasersfeld, E. Abstraction, Re-Presentation, and Reflection: An Interpretation of Experience and Piagets Approach. In *Epistemological Foundations of Mathematical Experience*; Steffe, L.P., Ed.; Springer: New York, NY, USA, 1991; pp. 45–67. doi:10.1007/978-1-4612-3178-3{_}4.
26. Sengers, P.; Boehner, K.; David, S.; Kaye, J.J. Reflective design. In Proceedings of the 4th Decennial Conference on Critical Computing between Sense and Sensibility—CC '05, Aarhus, Denmark, 20–24 August 2005; ACM: New York, NY, USA, 2005; p. 49. doi:10.1145/1094562.1094569.
27. Kaye, J.J.; Vertesi, J.; Avery, S.; Dafoe, A.; David, S.; Onaga, L.; Rosero, I.; Pinch, T. To have and to hold: Exploring the Personal Archive. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI '06, Montreal, QC, Canada, 22–27 April 2006; ACM: New York, NY, USA, 2006; p. 275. doi:10.1145/1124772.1124814.
28. Petrelli, D.; Whittaker, S.; Brockmeier, J. AutoTopography: What Can Physical Mementos Tell us about Digital Memories? In Proceedings of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems—CHI '08, Florence, Italy, 5–10 April 2008; ACM: New York, NY, USA, 2008; p. 53. doi:10.1145/1357054.1357065.

29. Gulotta, R.; Odom, W.; Faste, H.; Forlizzi, J. Legacy in the age of the internet. In Proceedings of the 2014 Conference on Designing Interactive Systems–DIS '14, Vancouver, BC, Canada, 21–25 June 2014; ACM: New York, NY, USA, 2014; pp. 975–984. doi:10.1145/2598510.2598579.
30. Sellen, A.J.; Whittaker, S. Beyond Total Capture: A Constructive Critique of Lifelogging. *Commun. ACM* **2010**, *53*, 70. doi:10.1145/1735223.1735243.
31. Rule, A. Visualizing Computer Activity to Support the Resumption of Long-term Creative Work. In Proceedings of the 2015 ACM SIGCHI Conference on C&C '15, Creativity and Cognition, Glasgow, UK, 22–25 June 2015; ACM: New York, NY, USA, 2015; pp. 343–344. doi:10.1145/2757226.2764772.
32. Cosley, D.; Sosik, V.S.; Schultz, J.; Peesapati, S.T.; Lee, S. Experiences with designing tools for everyday reminiscing. *Hum.–Comput. Interact.* **2012**, *27*, 175–198. doi:10.1080/07370024.2012.656047.
33. Gaver, W.; Sengers, P.; Kerridge, T.; Kaye, J.; Bowers, J. Enhancing ubiquitous computing with user interpretation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems–CHI '07, San Jose, CA, USA, 28 April–3 May 2007; ACM: New York, NY, USA, 2007; p. 537. doi:10.1145/1240624.1240711.
34. Peesapati, S.T.; Schwanda, V.; Schultz, J.; Lepage, M.; Jeong, S.y.; Cosley, D. Pensieve: Supporting Everyday Reminiscence. In Proceedings of the 28th International Conference on Human Factors in Computing Systems–CHI '10, Atlanta, GA, USA, 10–15 April 2010; ACM: New York, NY, USA, 2010; p. 2027. doi:10.1145/1753326.1753635.
35. Choe, E.K.; Lee, N.B.; Lee, B.; Pratt, W.; Kientz, J.A. Understanding quantified-selfers' practices in collecting and exploring personal data. In Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems–CHI '14, Toronto, ON, Canada, 26 April–1 May 2014; ACM: New York, NY, USA, 2014; CHI '14, pp. 1143–1152. doi:10.1145/2556288.2557372.
36. Plack, M.M. The Reflective Practitioner: Reaching for Excellence in Practice. *Pediatrics* **2005**, *116*, 1546–1552. doi:10.1542/peds.2005-0209.
37. Hoque, E.; Carenini, G. ConVis: A visual text analytic system for exploring blog conversations. *Comput. Graph. Forum* **2014**, *33*, 221–230. doi:10.1111/cgf.12378.
38. Huang, D.; Tory, M.; Adriel Aseniero, B.; Bartram, L.; Bateman, S.; Carpendale, S.; Tang, A.; Woodbury, R. Personal visualization and personal visual analytics. *IEEE Trans. Vis. Comput. Graph.* **2015**, *21*, 420–433. doi:10.1109/TVCG.2014.2359887.
39. Huang, D.; Tory, M.; Bartram, L. A Field Study of On-Calendar Visualizations. In Proceedings of the Graphics Interface 2016, Victoria, BC, Canada, 1–3 June 2016; Canadian Human-Computer Communications Society; School of Computer Science, University of Waterloo: Waterloo, ON, Canada, 2016; pp. 13–20. doi:10.20380/GI2016.03.
40. Cao, N.; Cui, W. Overview of Text Visualization Techniques. In *Introduction to Text Visualization*; Atlantis Press: Paris, France, 2016; pp. 11–40. doi:10.2991/978-94-6239-186-4[_].2.
41. Leskovec, J.; Rajaraman, A.; Ullman, J.D. Finding Similiar Items. In *Mining of Massive Datasets*; Cambridge University Press: Cambridge, UK, 2014; Chapter 3, pp. 73–170.
42. Pousman, Z.; Stasko, J.T.; Mateas, M. Casual information visualization: Depictions of data in everyday life. *IEEE Trans. Vis. Comput. Graph.* **2007**. doi:10.1109/TVCG.2007.70541.
43. Dave, K.; Wattenberg, M.; Muller, M. Flash forums and forumReader. In Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work–CSCW '04, Chicago, IL, USA, 6–10 November 2004; ACM: New York, NY, USA, 2004; p. 232. doi:10.1145/1031607.1031644.
44. Mishne, G. Information access challenges in the blogspace. In Proceedings of the International Workshop on Intelligent Information Access (IIA 2006), Helsinki, Finland, 6–8 July 2006; pp. 6–8.
45. Singh, P.V.; Sahoo, N.; Mukhopadhyay, T. *Seeking Variety: A Dynamic Model of Employee Blog Reading Behavior Seeking Variety: A Dynamic Model of Employee Blog Reading Behavior*; Technical report; Carnegie Mellon University: Pittsburgh, PA, USA, 2010.
46. Miyazoe, T.; Anderson, T. Viewing and participating: Blog visualization and its learning outcomes in blended learning. In Proceedings of the 2011 IEEE International Professional Communication Conference, Cincinnati, OH, USA, 17–19 October 2011. doi:10.1109/IPCC.2011.6087217.

47. Niu, S.; McCrickard, D.S.; Harrison, S. An observational study of simultaneous and sequential interactions in co-located collaboration. In Proceedings of the 16th IFIP TC 13 International Conference on Hum.-Computer Interact. (INTERACT '17), Mumbai, India, 25–29 September 2017; Springer: Mumbai, India, 2017; Volume 10515 LNCS, pp. 163–183. doi:10.1007/978-3-319-67687-6_12.
48. Niu, S.; McCrickard, D.S.; Harrison, S. Investigating Notifications and Awareness for Multi-user Multi-touch Tabletop Displays. In Proceedings of the 16th IFIP TC 13 International Conference on Hum.-Comput. Interact. (INTERACT '17), Mumbai, India, 25–29 September 2017; Springer International Publishing: Cham, Switzerland, 2017; Volume 10515 LNCS, pp. 223–244. doi:10.1007/978-3-319-67687-6_15.
49. Marshall, P.; Morris, R.; Rogers, Y.; Kreitmayer, S.; Davies, M. Rethinking 'multi-user'. In Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems—CHI '11, Vancouver, BC, Canada, 7–12 May 2011; ACM: Vancouver, BC, Canada, 2011; p. 3033. doi:10.1145/1978942.1979392.
50. Wobbrock, J.O.; Morris, M.R.; Wilson, A.D. User-defined gestures for surface computing. In Proceedings of the 27th International Conference on Human Factors in Computing Systems—CHI 09, Boston, MA, USA, 4–9 April 2009; ACM: Boston, MA, USA, 2009; p. 1083. doi:10.1145/1518701.1518866.
51. Andrews, C.; Endert, A.; North, C. Space to Think: Large, High-Resolution Displays for Sensemaking. In Proceedings of the 28th International Conference on Human Factors in Computing Systems—CHI '10, Atlanta, GA, USA, 10–15 April 2010; ACM: Atlanta, GA, USA, 2010; p. 55. doi:10.1145/1753326.1753336.
52. Manning, C.; Surdeanu, M.; Bauer, J.; Finkel, J.; Bethard, S.; McClosky, D. The Stanford CoreNLP Natural Language Processing Toolkit. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, Baltimore, MD, USA, 22–27 June 2014; pp. 55–60. doi:10.3115/v1/P14-5010.
53. Leskovec, J.; Rajaraman, A.; Ullman, J.D. *Mining of Massive Datasets*; Cambridge University Press: Cambridge, UK, 2014.
54. Bryson, B. *A Walk in the Woods*; Random House: New York, NY, USA, 1998.
55. Buck, R. *The Oregon Trail: A New American Journey*; Simon and Schuster: New York, NY, USA, 2015.
56. Baumer, E.P.; Khovanskaya, V.; Matthews, M.; Reynolds, L.; Schwanda Sosik, V.; Gay, G. Reviewing Reflection: On the Use of Reflection in Interactive System Design. In Proceedings of the 2014 Conference on Designing Interactive Systems, Vancouver, BC, Canada, 21–25 June 2014; ACM: New York, NY, USA, 2014; pp. 93–102. doi:10.1145/2598510.2598598.
57. Huang, C.Y.; Shen, Y.Z.; Lin, H.X.; Chang, S.S. Bloggers motivations and behaviours: A model. *J. Advert. Res.* **2007**, *47*, 472–484. doi:10.2501/S0021849907070493.
58. Carroll, E.A.; Latulipe, C.; Fung, R.; Terry, M. Creativity Factor Evaluation: Towards a Standardized Survey Metric for Creativity Support. In Proceedings of the Seventh ACM Conference on Creativity and Cognition C&C '09, Berkeley, CA, USA, 26–30 October 2009; ACM: New York, NY, USA, 2009; pp. 127–136. doi:10.1145/1640233.1640255.
59. Dove, G.; Halskov, K.; Forlizzi, J.; Zimmerman, J. UX Design Innovation: Challenges for Working with Machine Learning as a Design Material. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems—CHI '17, Denver, CO, USA, 6–11 May 2017; ACM: New York, NY, USA, 2017; pp. 278–288. doi:10.1145/3025453.3025739.
60. Lee, T.Y.; Smith, A.; Seppi, K.; Elmqvist, N.; Boyd-Graber, J.; Findlater, L. The Human Touch: How Non-Expert Users Perceive, Interpret, and Fix Topic Models. *Int. J. Hum. Comput. Stud.* **2017**, *105*, 28–42. doi:10.1016/j.ijhcs.2017.03.007.

