Intelligent Interactive Multimedia Systems and Services
G.A. Tsihrintzis et al. (Eds.)
IOS Press, 2013

doi:10.3233/978-1-61499-262-2-108

Dynamic Word Clouds: Context-based Word Clouds of Presentation Slides for Quick Browsing

Yuanyuan WANG [a,1] and Kazutoshi SUMIYA [b]
[a] *University of Hyogo, Japan, E-mail: ne11u001@stshse.u-hyogo.ac.jp*
[b] *University of Hyogo, Japan, E-mail: sumiya@shse.u-hyogo.ac.jp*

Abstract. With the advent of online education services such as iTunes U and SlideShare, students are able to access various presentation files from many universities for study purposes. However, self-learners retrieving such files still need support to decide which files are worth learning because it can be difficult to understand the context of each file. In this paper, we introduce a quick browsing system to help a user effectively decide whether a presentation content is appropriate for his/her specific learning needs. Our method provides context-based word clouds that summarize slide information to help the user visually understand the context of a presentation content. In our method, words important to the "presentation context," that is, the relevant information on the slides, are first extracted based on components of the presentation content (i.e., intra-slide structure and inter-slide structure). Then the word cloud is generated by weighting the words within the presentation context for each slide. Finally, the word clouds are interactively presented with visual effects that reflect their semantic meanings.

Keywords. presentation slides, dynamic word cloud, quick browsing

Introduction

Presentation slides (e.g., PowerPoint [2], Keynote [3]) are now one of the most frequently used tools for educational purposes. A huge amount of slide-based lecture material, often prepared from teaching material used in actual classes at universities or other educational institutions, is freely shared on Web sites such as iTunes U [1] and SlideShare [2]. Thus, not only students who missed a lecture or presentation, but also anyone interested in the topic can study the presentation on their own. Therefore, techniques are in demand that will efficiently find one or more appropriate slides with content worth learning from the vast numbers of presentations available. Although many techniques for searching and recommending presentation slides have been proposed, some problems remain from the viewpoint of understandability for users browsing search engine results. One problem

[1] Corresponding Author: Yuanyuan Wang, University of Hyogo, Shinzaike-honcho 1–1–12, Himeji, Hyogo 670–0092, Japan; E-mail: ne11u001@stshse.u-hyogo.ac.jp.
[2] http://office.microsoft.com/ja-jp/powerpoint/
[3] http://www.apple.com/iwork/keynote/

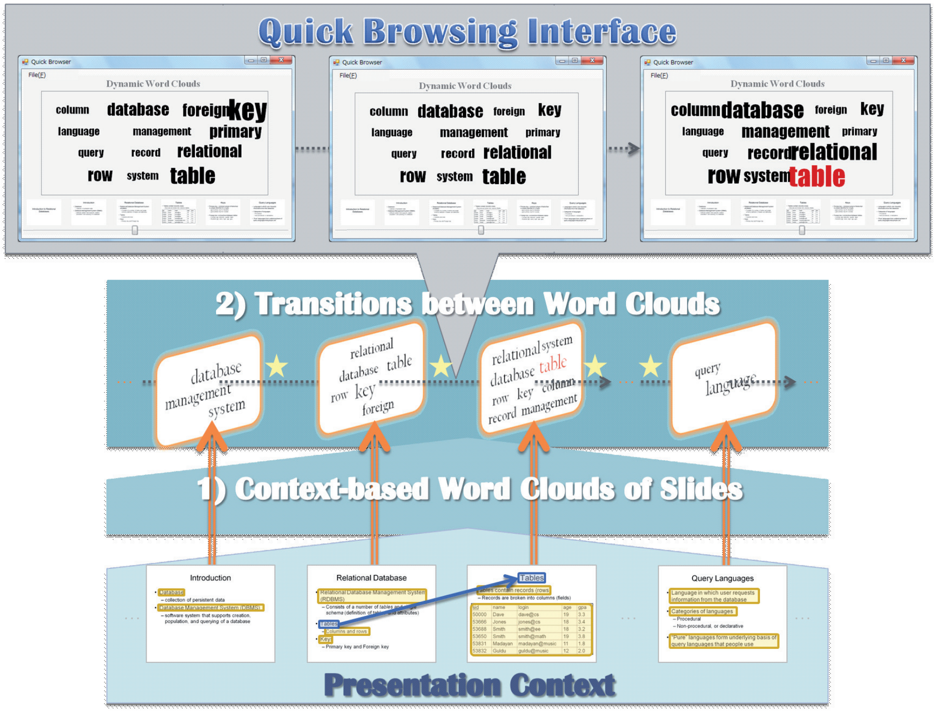

Figure 1. Conceptual diagram of our proposed quick browsing method

is a search engine does not consider context when matching user query words within presentation content, leading to a large number of candidate results. Another problem is the difficulty of general quick browsing, that is, when browsing slide titles only, users cannot grasp specifics of the content. In addition, word clouds of slides based on word frequency can destroy the implicit relevant information between slides and decrease the relevance of words in slides to the overall context. This makes it difficult to understand the context of words in candidate presentations when choosing relevant files.

As depicted in Fig. 1, we present a quick browsing method that 1) generates context-based word clouds for each slide by weighting the words within the context of the presentation (i.e., the intra-slide structure and inter-slide structure) and 2) determines transitions between the word clouds based on relationships between the words in serial slides. We derive the intra-slide structure that slide structure by focusing on the level of indentation in the slide text, and determine the inter-slide structure that relationships between slides by considering words that appear at different indentation levels in the structure of other slides. For example, 'Keys' appears in the body of text in a slide entitled 'Relational model', which related to the slide entitled 'Relational database' and 'Relational model' appears in the body of text in it. 'Keys' is the title of other slide that is related to the slide entitled 'Relational model'.

In this paper, we define presentation context to mean the context for the slide in a presentation, represented by the relevant information on the slide and allowing for relevant information from the rest of the presentation that is not included on the slide. We define two types of presentation context for a slide: link context and structural context, based on the relationships between slides and slide structure, respectively. Using presentation context, we can generate context-based word clouds of slides by weighting words in the presentation. There are two concepts that are particularly helpful when quick browsing presentation content: 1) Presentation flow: link or break [3]. Often, presentations are formed of a chain of slides such that one slide links to the next. Sometimes, however, a slide will move from the point in a previous slide to a completely different point. In this case, there is a break between them. 2) Highlight points: semantics. This occurs when one slide describes a point from a previous slide in detail. In this case, our approach presents words interactively from one slide to another as a streaming word cloud reflecting the flow of points in the slides, helping users to select relevant presentations from search engine results easily and effectively.

This paper is organized as follows. The next section reviews related work, and Section 2 describes how to determine presentation contexts for slides. Section 3 presents the generation of context-based word clouds and the determination of transitions between the word clouds, and Section 4 introduces a prototype application for quick browsing based on our method. Finally, Section 5 concludes this paper with suggestions for further work.

1. Related Work

Our work is directly related to research efforts in two areas: text analytics and information visualization. In the area of text analytics, researchers have developed a number of approaches for text summarization, of which there are two main techniques: sentence-based and keyword-based text summarization. Sentenced-based approaches identify the most salient sentences in a document [4], [5]. For example, Murai and Ushiama [6] proposed a browsing method that presents users with a review-based recommendation of attractive sentences in a novel. However, it may be time consuming for users to read several sentences per document, especially when handling a large number of documents. Alternatively, keyword-based methods summarize documents by topic, each of which is characterized by a set of keywords [7], [8], [9]. Our quick browsing method is built on the latter method, but its focus is on enhancing the summarization results through keyword clouds and visualization. Moreover, we provide users with a novel interface with visual transitions to present the dynamic word clouds.

In the area of information visualization, researchers have developed various visualization approaches to text analysis. These systems can be classified into two categories: metadata-based and content-based text visualization. In email analysis for instance, metadata-based text visualization can use a time-based visualization to explain text summarization results derived by a text analytic engine [10], or a create a relationship-based visualization of email senders and receivers [11]. For content-based text visualization, Viegas et al. [12] used Themail to visualize keywords based on keyword TF-IDF scores in an email collection. Similarly, Strobelt et al. [13] used a mixture of images and TF-IDF-based keywords to create a compact visualization of a document. More recently, Chen et al. [14] and Iwata et al. [15] focused on visualizing document clustering results. In contrast, others have concentrated on representing text content at the word or phrase level, including TextArc (www.textarc.org), WordTree [16], Phrase Net [17], and Fea-

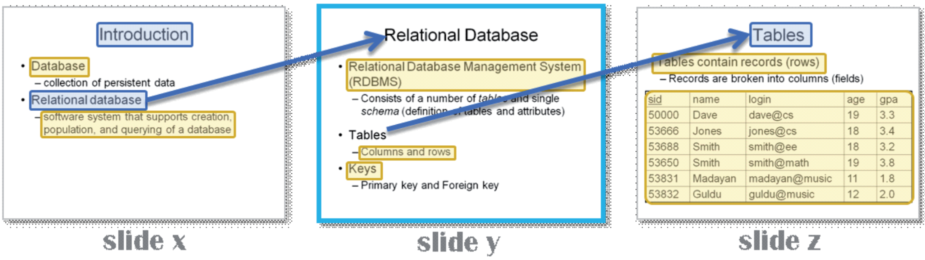

Figure 2. Presentation contexts for slide *y*

tureLens [18]. Our work focuses on visualizing presentation content by generating word clouds of slides, and presenting transitions between the word clouds.

2. Determination of Presentation Contexts for Slides

We determine two types of presentation context for a slide: link context and structural context, based on the relationships between slides and slide structure, respectively. We define the slide whose presentation context is discussed to be the target slide. We construct the slide structure based on the indentations in the slide text. The slide title is the first indentation level, the first item of text within the slide body is the second indentation level, and the depth of the sub-items increases with indentation level. Non-text objects, such as figures or tables, are considered to be at the same indention level as the surrounding text.

2.1. Determination of Link Context for Slides

The link context for a target slide consists of links and anchors (similar to hyperlinks in Web pages) related to the text body and titles of other slides. They refer to words in the title of the target slide and titles of other slides that contain words in the body of the target slide. They also indicate from what type of content the target slide is referred. We extract the link context of the target slide by finding the same words at different levels in the target and other slides.

For a given bag of words M in the title and a given bag of words N in a level in the body of the target slide, words in the titles and levels in the body of other slides are extracted: $\cdots, T_2, B_2, T_1, B_1$. Here, T_i is the title of slide i and B_i are the words in a level of the body of slide i. If B_i corresponds to M, B_i can be considered as a link anchor. Then, B_i links to the target slide such that the words in B_i and its slide title T_i belong to the link context for the target slide, while the words in B_i are similar to that in M. This is calculated using the Simpson similarity coefficient [19], as $Sim(B_i, M) = |B_i \cap M| / min(|B_i|, |M|)$. When $Sim(B_i, M)$ exceeds a predefined threshold, the words in B_i and M are similar. Meanwhile, if N corresponds to T_i, N can be considered as a link anchor. Then, N links to the slide titled T_i in that the words in T_i belong to the link context, while the words in N are similar to that in T_i.

In Fig. 2, the link context for slide y (in blue portions) shows that slide y explains “Relational Database,” which is referred to on slide x as a subheading of ‘Introduction,’ and the subheading ‘Tables’ in slide y is described in slide z.

2.2. Determination of Structural Context for Slides

The structural context for the target slide consists of lower, current, and upper levels of the target levels corresponding to the link context in the target slide, and lower, current, and upper levels of the link context in other slides based on slide structure. When the target slide does not have a link context, we take the title of the target slide as the target level, and then we extract the structural context for the target slide that consists of the levels below the title.

For a given bag of words N at a level in the body of the target slide, words in the lower, current, and upper level of N are extracted: $\cdots$, l_3, l_2, l_1. Here, l_j represents a bag of words at a particular indention level j. When l_j and N are the current level in the target slide, the words at level l_{j+1} are at the lower level of N and l_{j-1} are at the upper level of N. The lower, current, and upper levels of the link context in other slides are extracted in the same way.

In our method, when the target level in the target slide corresponds to the link context in more than one slide, we just extract the link context of the slide nearest to the target slide. The link context and structural context are extracted within a minimal range of surrounding information, containing just enough words to characterize the presentation context. Therefore, the presentation context expresses presentation flow and highlight points well.

Fig. 2 illustrates the structural context for slide y (in yellow portions), where ‘Tables’ is related to ‘RDBMS,’ ‘Keys,’ and ‘Columns and rows’ at the upper, current, and lower levels in slide y and ‘Tables’ includes a link at a lower level, ‘Tables contain records (rows),’ in slide z; ‘Relational Database’ includes a link to a lower level at ‘RDBMS’ in slide y, and ‘Relational Database’ is related to current and lower levels at ‘Database’ and ‘Software system’ in slide x.

3. Presentation of Context-based Word Clouds

3.1. Generation of Word Clouds of Slides

To present a streaming context-based word cloud that reflects the semantics of the words, slide word clouds are generated from words extracted from the presentation context by weighting the words to determine font size. For each type of presentation context, we calculate the degree of the words that 1) appear close to the target slide and 2) appear frequently near the target slide but less frequently around other slides.

Let us consider each word of target slide S as a relevant object, denoted by o. The degree of o for the presentation context $P(S)$ is defined as follow:

$$W(o,P(S)) = \frac{density(o,P(S))}{dist(o,S)} \tag{1}$$

here, $density(o,P(S))$ is the density of o for the presentation context $P(S)$, and $dist(o,S)$ is the distance between o and S. Intuitively, $density(o,P(S))$ means how densely the same word as o appears in $P(S)$. If the same word as o appears frequently in $P(S)$ but less frequently in other presentation contexts, $density(o,P(S))$ becomes large. Suppose that S

is the kth slide among all slides (the target slide). The density of o in $P(S_k)$ is calculated as follows:

$$density(o, P(S_k)) = \frac{N_{A(o,P(S_k)}}{N_{A(o,U)}} \tag{2}$$

where $A(o, P(S_k))$ is a set of relevant objects representing the same word as the object o in the presentation context $P(S_k)$, and $A(o, U)$ is the set of relevant objects in the presentation context of all slides: $U{=}P(S_1) \cup P(S_2) \cup \ldots$. $N_{A(o,P(S_k))}$ and $N_{A(o,U)}$ represent the number of objects in $A(o, P(S_k))$ and $A(o, U)$, respectively. Because it is difficult to identify the set U due to mutual dependencies between the presentation contexts, we approximate the set U as the set of relevant objects of all slides. The distance $dist(o, S)$ indicates the strength of the associations between the relevant object o and the target slide S, and is defined for each type of presentation context. Distance in link context is the number of link relationships from S to o. Distance in structural context is the number of parent, brother, and child nodes to be followed from the target levels in S to o.

We generate context-based word clouds of slides by extracting the words based on the ratio of the degree of each word and the highest degree of the word in each slide. We also sort the font sizes of the words into three groups as follows:

$$STag(c, P(S)) = \{c | \tfrac{W(c,P(S))}{W_{max}(P(S))} \geq \theta_1, \tfrac{W(o,P(S))}{W_{max}(P(S))} < \theta_2\} \tag{3}$$

$$MTag(c, P(S)) = \{c | \tfrac{W(c,P(S))}{W_{max}(P(S))} \geq \theta_2, \tfrac{W(o,P(S))}{W_{max}(P(S))} < \theta_3\} \tag{4}$$

$$LTag(c, P(S)) = \{c | \tfrac{W(c,P(S))}{W_{max}(P(S))} \geq \theta_3\} \tag{5}$$

In Eqs. (3), (4), and (5), $W(c, P(S))$ is the degree of c and $W_{max}(P(S))$ is the highest degree of the word in S using Eqs. (1) and (2). Including too many words in each word cloud does not help users to browse them effectively, so we extract c in S such that the ratio of $W(c, P(S))$ and $W_{max}(P(S))$ is greater than a threshold (i.e., 0.25). $STag(c, P(S))$, $MTag(c, P(S))$, and $LTag(c, P(S))$ are the groups of weighted words to be displayed in small, medium, and large font size such that the ratios satisfy Eq. (3), Eq. (4), and Eq. (5), respectively. In this paper, we empirically set the values of the thresholds to be $\theta_1 = 0.25$, $\theta_2 = 0.50$, and $\theta_3 = 0.75$. Although, in general, the word position is important for word clouds, in this paper, we provide an intuitive interface by changing font sizes only.

3.2. Determination of Transitions between Word Clouds

Based upon the presentation contexts for slides, we present dynamic word clouds with visual effects that reflect the relationships between words interactively. For this purpose, we use relationships between words in the word clouds similar to the relationships between slides defined in our previous work [20], which fall into four types based on the presentation contexts for slides:

- *Detailed*: titles of other slides belonging to the link context for the target slide in the word clouds. The word clouds of the other slides have more information about the link context than the word cloud of the target slide.
- *Generalized*: bodies of other slides belonging to the link context for the target slide in the word clouds. The word cloud of the target slide contains the words about the link context in the outline given in the word cloud of the other slides.

- *Parallel*: titles of other slides in the word clouds belonging to the link context that link to the current levels in the target slide, these word clouds of other slides are parallel with each other.
- *Independent*: slides do not have a link context for each other in their word clouds.

To present dynamic word clouds, the transitions discussed here explain the kinds of visual effects added to the relationship types, reflecting presentation flow or highlights. Presentation flow consists of many chains of serial slides such that each chain and each transfer switch between chains must be presented. For one chain of serial slides, *detailed*, *generalized* or *parallel* relationships exist between them. For a transfer switch between different chains of slides, the *independent* relationship exists between them. Highlights are the words belonging to the link context of one chain in detail. The effects for three types of transitions between the words in the generated word clouds are as follows:

- *Font size changes*: a *shrinking* or *expanding* effect is set between serial slides when presenting one chain in the presentation flow. When the font sizes of the words in the current word cloud are smaller than those in the previous one, a *shrinking* effect is implemented. When the font sizes of the words in the current word cloud are larger than those in the previous one, an *expanding* effect is applied after a *shrinking* effect. Users can easily understand that they are following a chain of slides where the words are mentioned.
- *Color changes*: a *coloring* effect is set between serial slides to highlight detailed points in one chain of the presentation flow. For a current word cloud, the words belonging to the next word cloud, which are described in detail on the next slide, are drawn in red. Users easily see that these words are highlighted in one chain. When the highlighted words in the current word cloud are not detailed in the next word cloud, these words are drawn the default color (black).
- *Switching*: a *dissolve* effect is applied to a transfer switch between different chains in the presentation flow. The current word cloud disappears and the next word cloud appears gradually in its place. Users easily grasp that a transfer switch has occurred.

4. Application

Based on the method described above, we built a novel quick browsing interface to support users to quickly gain a broad understanding of presentation contents (see Fig. 3). For example, when a user wants a presentation about ‘relational database overview: tables and keys,’ he/she can specify any presentation content for quick browsing from the results, and the browser presents all words from all word clouds of slides with an initial font size (i.e., 20pt) in a dynamic word cloud display. He/she can use a seekbar to turn over slides, and the weighted words belonging the word clouds of slides with their font sizes (i.e., small: 30pt, medium: 40pt, large: 50pt) are dynamically presented with visual transitions in the word cloud display.

An example of browsing the dynamic word clouds of two presentations from our dataset, P_A and P_B, is shown in Fig. 4. The user browses the flow of slide 3 to slide 5 in P_A [4] named “Introduction to Relational Databases,” and the flow of slide 4 to slide

[4] http://web.cecs.pdx.edu/˜howe/cs410/lectures/Relational_Intro_1.ppt

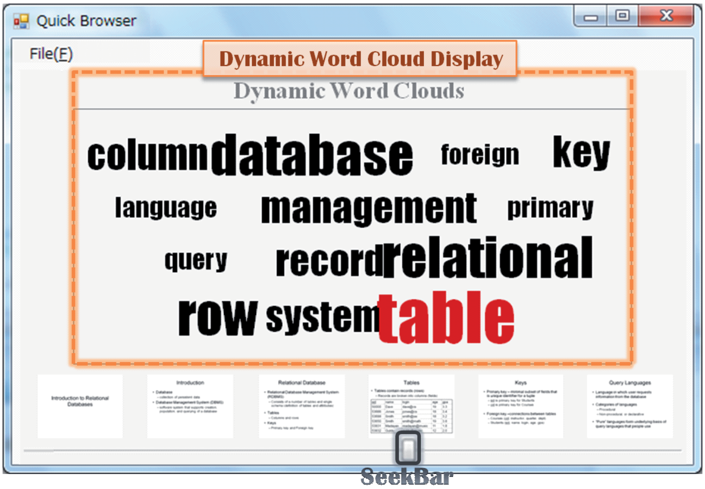

Figure 3. Screenshot of quick browsing interface

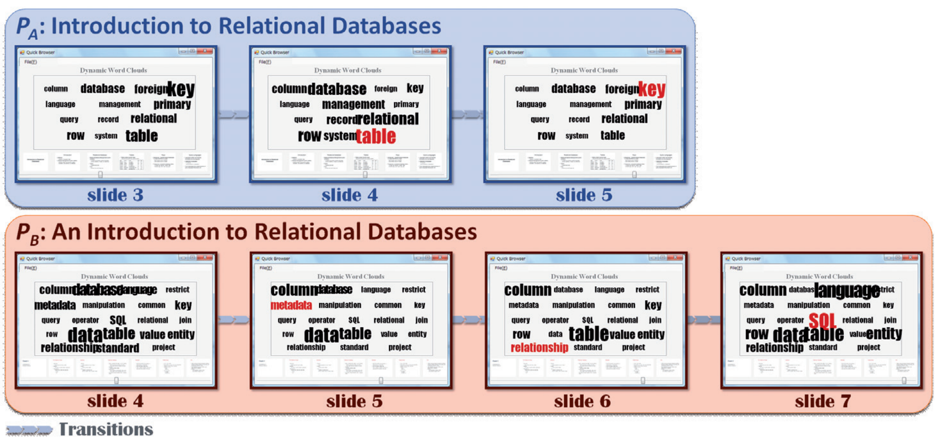

Figure 4. Examples of dynamic word clouds in P_A and P_B

7 in P_B [5] named "An Introduction to Relational Databases." When the user moves the seekbar to go from slide 3 to 5 in P_A, the font sizes of 'table,' 'database,' and 'relational' are increased in slide 4, and the font sizes of 'key,' 'foreign,' and 'primary' are increased in slide 5. In particular, 'table' and 'key' are drawn in red in slides 4 and 5, respectively. Meanwhile, when the user moves the seekbar to go from slide 4 to 7 in P_B, the font sizes of 'column' is increased in slide 5, the font sizes of 'table' and 'relationship' are increased in slide 6, and the font sizes of the words such 'SQL,' 'language,' and 'data' are increased in slide 7. In particular, 'metadata,' 'relationship,' and 'SQL' are drawn in red in slides 5, 6, and 7, respectively.

[5] http://www.atilim.edu.tr/~mrehan/Chapter%203.ppt

In the case of P_A, we find that 'table' and 'key' are core points in the flow of slides 3 to 5, the presentation explains 'tables and keys in relational databases,' detailed relationships exist between 'table' in slides 3 and 4, and 'key' in slides 3 and 5. There is also a parallel relationship between 'table' and 'key' in slides 4 and 5. For P_B, 'metadata,' 'relationship,' and 'SQL' are core points in the flow of slides 4 to 7, the presentation explains 'characteristics of data in relational database,' and detailed relationships exist between 'metadata' in slides 4 and 5, 'relationship' in slides 4 and 6, and 'SQL' in slides 4 and 7. A parallel relationship exists among 'metadata,' 'relationship,' and 'SQL' in slides 5, 6, and 7. Therefore, P_A is worth learning in that it better meets the user's needs. Although we confirmed that our proposed dynamic word clouds enables a user to effectively and easily select presentations with contents that meet his/her needs, we encountered difficulties when presentation contents (i.e., P_A and P_B) had a similar title. We need to consider how best to present the differences in similar presentation content (e.g., different topics or same topics with different context information, etc.). Additionally, we must provide a quick browsing interface that compares similar presentation contents simultaneously with the differences clearly marked.

5. Concluding Remarks

In this paper, we proposed a quick browsing method for presentation content that uses dynamic word clouds to present words interactively with visual effects to help a user visually understand the context of content within a presentation. We described how presentation context can be determined from slide structure and the relationships between slides. In order to generate context-based word clouds of slides, we extracted weighted words from presentation context, and then presented dynamic word clouds with transitions that highlighted the relationships between slides.

In the future, we plan to develop various other visual effects for presenting transitions in the word clouds to help users intuitively understand presentation content. We have to confirm that our quick browsing method enables users to select presentation contents meet their needs effectively and easily. We also plan to develop a quick browsing interface that presents the contents of multiple presentations for comparison.

References

[1] iTunes U, http://www.apple.com/jp/education/itunes-u/.

[2] SlideShare, http://www.slideshare.net/.

[3] Presentation Flow: Link Or Break, http://philpresents.wordpress.com/2011/11/23/presentation-flow-link-or-break/.

[4] G. Carenini, R. Ng, and X. Zhou, "Summarizing Email Conversations with Clue Words," in *Proc. of the 16th International World Wide Web Conference (WWW 2007)*, May 2007, pp. 91–100.

[5] D. Wang, T. Li, S. Zhu, and C. Ding, "Multi-document Summarization via Sentence-level Semantic Analysis and Symmetric Matrix Factorization," in *Proc. of the 31st Annual International ACM SIGIR Conference (SIGIR 2008)*, July 2008, pp. 307–314.

[6] S. Murai and T. Ushiama, "Review-Based Recommendation of Attractive Sentences in a Novel for Effective Browsing," in *International Journal of Knowledge and Web Intelligence*, vol. 3, no. 1, 2012, pp. 58–69.

[7] M. Dredze, H. Wallach, D. Puller, and F. Pereira, “Generating Summary Keywords for Emails using Topics,” in *Proc. of International Conference on Intelligent User Interfaces (IUI 2008)*, January 2008, pp. 199–206.
[8] A. McCallum, X. Wang, and A. Corrada-Emmanuel, “Topic and Role Discovery in Social Networks with Experiments on Enron and Academic Email,” in *Journal of Artificial Intelligence Research*, vol. 30, 2007, pp. 249–272.
[9] J. Kaye, A. Lillie, D. Jagdish, J. Walkup, R. Parada, and K. Mori, “Nokia Internet Pulse: A Long Term Deployment and Iteration of a Twitter Visualization,” in *Proc. of International Conference on Human-Computer Interaction (CHI 2012)*, May 2012, pp. 829–844.
[10] S. Liu, M. X. Zhou, S. Pan, Y. Song, W. Qian, W. Cai, and X. Lian, “TIARA: Interactive, Topic-Based Visual Text Summarization and Analysis,” in *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 3, no. 25, 2012, pp. 1–28.
[11] A. Perer and M. Smith, “Contrasting Portraits of Email Practices: Visual Approaches to Reflection and Analysis,” in *Proc. of Advanced Visual Interfaces (AVI 2006)*, May 2006, pp. 389–395.
[12] F. Viegas, S. Golder, and J. Donath, “Visualizing Email Content: Portraying Relationships from Conversational Histories,” in *Proc. of the 2006 Conference on Human Factors in Computing Systems (CHI 2006)*, April 2006, pp. 979–988.
[13] H. Strobelt, D. Oelke, C. Rohrdantz, A. Stoffel, A. D. Keim, and O. Deussen, “Document Cards: A Top Trumps Visualization for Documents,” in *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, 2009, pp. 1145–1152.
[14] Y. Chen, L. Wang, M. Dong, and J. Hua, “Exemplar-Based Visualization of Large Document Corpus,” in *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, 2009, pp. 1161–1168.
[15] T. Iwata, T. Yamada, and N. Ueda, “Probabilistic Latent Semantic Visualization: Topic Model for Visualizing Documents,” in *Proc. of the 14th ACM SIGKDD International Conference and Knowledge Discovery and Data Mining (KDD 2008)*, August 2008, pp. 263–271.
[16] M. Wattenberg and F. Viegas, “The Word Tree, an Interactive Visual Concordance,” in *Proc. of Information Visualization Conference (InfoVis 2008)*, October 2008, pp. 1221–1228.
[17] F. V. Ham, M. Wattenberg, and F. B. Viegas, “Mapping Text with Phrase Nets.” in *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, 2009, pp. 1169–1176.
[18] A. Don, E. Zheleva, M. Gregory, S. Tarkan, L. Auvil, T. Clement, B. Shneiderman, and C. Plaisant, “Discovering Interesting Usage Patterns in Text Collections: Integrating Text Mining with Visualization,” in *Proc. of the 16th ACM Conference on Information and Knowledge Management (CIKM 2007)*, November 2007, pp. 213–222.
[19] E. H. Simpson, “Measurement of Diversity,” in *Nature*, vol. 163, April 1949, p. 688.
[20] D. Kitayama, A. Otani, and K. Sumiya, “An Extracting Method of Semantic Relations between Scenes for Presentation Contents [in Japanese],” in *IPSJ Transaction on Database (TOD)*, vol. 2, no. 2, June 2009, pp. 71–85.