

Peereviz: Visualizing Peer Reviews

Kanit Wongsuphasawat, Chih-Chiang Wei, Thiraphat Charoensripongsa

Stanford University

{kanitw, ccwei, tchar}@stanford.edu

ABSTRACT

Peer reviews is one possible solution in scaling assessment of open-ended assignment in massive online open courses. However, the amount of data from peer review process is huge and it becomes difficult to explore and understand. In this paper, we present Peereviz, an exploration tool for large-scale peer review that helps course instructors to understand and gain insights from the peer review activities, the engagement of students, and the quality of the reviews. Peereviz utilized existing text visualization techniques as well as Multiple Coordinated Views and enable the quick navigation through original feedbacks.

Author Keywords

peer review; group collaboration; visualization; MOOC; exploration; text visualization

INTRODUCTION

Massive open online courses(MOOCs) including Khan Academy, Udacity, Coursera and Venture Lab are increasing in numbers [5]. One emerging challenge is how to scale assessment to these high student-teacher ratio environments. For assignment that have clear representative answer such as multiple choices questions, automated grading system can be employed. While human graders are still needed for the assessment of open-ended assignments.

Using peer review system [9], i.e. having each student grade each other, to assess students work is another possible solution. However, in the massive online courses, the number of reviews done in each class is large and it is overly difficult and time-consuming for instructors to get a comprehensive understanding of the review of each task.

To address this problem, we developed Peereviz, a peer review exploration tool on top of Venture Lab platform. It aims for helping instructors to understand the massive amount of peer review results. The design goal of the tool is to create a visualization tool that helps course instructors in three ways:

1. get the overall understanding of peer review activities;
2. be able to dive into specific part of the review results;

Figure 1. the actual evaluation form on the Venture lab platform. The top part shows the quantitative feedbacks which are score ranging from Low (1) to High (10) while the bottom part shows the feedback capture grid which a form where the reviewers give their opinions on four different aspect: notable, constructive, questions, and ideas.

3. quick browse and select higher quality reviews to read.

As a data explorer, it follows “overview first, zoom and filter, then details- on-demand mantra” [8].

The paper organization is as follows. We first summarize previous work in peer review system and visualization. We then describe the peer review data we used. Next, we present our design and visualization techniques, followed by the evaluation and feedbacks from users. Finally, we suggest the possible directions for future work and then summarize our work.

RELATED WORK

A number of studies has confirmed the usefulness of the peer review system to the students in various domains including programming, writing and design classes [11, 12, 14].

However, to the best of our knowledge, work that combines visualization techniques to help instructors understand peer review data is still limited. The only work we discovered is an interactive tool for peer-review exploration proposed by Xiong et al. [18] They primarily work on the improvement of semantic information to help instructor discover interesting patterns and compare different groups of student in the writing class on SWORD system [10]. In contrast to the prior work, we mainly focuses on the visualization and interaction design for exploring multi-dimensional peer review data in the recently emerging large-scale online courses.

DATA SET DESCRIPTION

We use a sample data set from a final project peer review result in the Spring 2012 Entrepreneurship class on the Venture Lab platform. A total of 1,206 peer reviews are collected, which consists of 116 teams being reviewed and 212 individual reviewers. For privacy reasons, the team names are

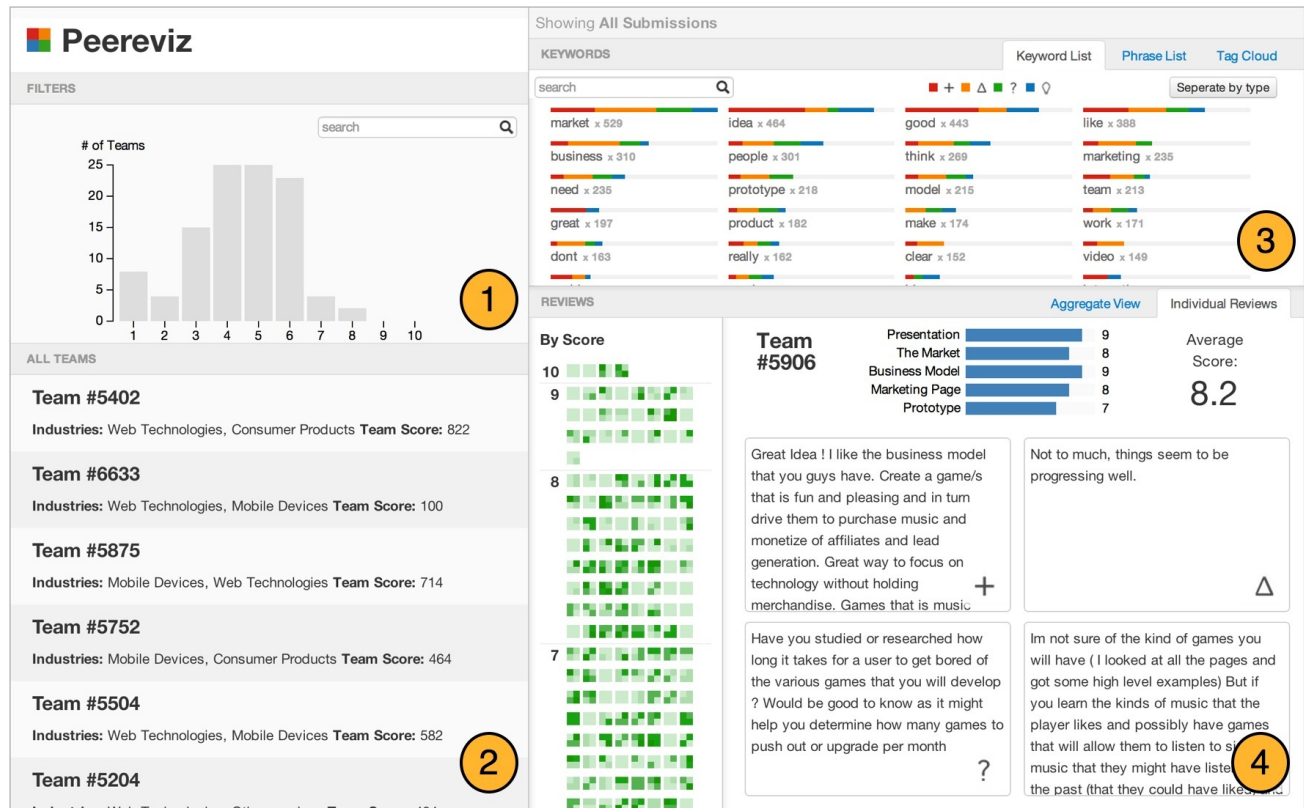


Figure 2. The layout for Peereviz is divided into two parts: the team browser (1,2) and the review browser (3,4). (1) shows the overall score distribution, (2) team description list, (3) keyword list, and (4) Aggregate and individual reviews.

anonymized and we presented team description as much as needed. The actual peer review form has two parts, including multiple quantitative score ranging from Low (1) to High (10) for different aspects of the final project and qualitative text feedback, which utilizes a feedback capture grid [4] where reviewers could express their opinions in four dimensions — what they like (*notable*), constructive criticisms (*constructive*), questions they have about the work (*questions*), and any additional ideas for the project (*ideas*).

VISUALIZATION DESIGN

We use multiple coordinated views techniques to show data in different representations which enable users to interact, explore and understand intricate data [16]. Furthermore, we also use text visualization techniques, including word count list and tag cloud, which are often used to visualize and understand various large text data [13, 17].

Layout

As one of our design goals is to provide fast navigation, we put all views in one page so users do not need to scroll or find hidden menu to achieve their intended tasks. The layout, as shown in Figure 2, consists of two parts: the team browser and the review browser. The first is the team browser on the left side, which contains a team list and a bar chart that visualizes overall score distribution for all team projects. The second is the review browser on the right side which consists

of two sub-parts: the keyword visualization for browsing keyword on the upper part and review display view on the bottom part.

Team Browser

The team browser serves as the main part of the overview where the user can select one or multiple teams using different ways of filtering.

The user can select one team by clicking one of the small block on the stacked bar chart or select multiple teams at the same time by brushing over the bar chart. After the selection, the team descriptions will be listed below as noted by (2) in Figure 2. The information provided in the review browser will be updated accordingly with the data of the selected team(s). In addition, it provides a search tool to select teams by their descriptions as well.

Review Browser

The review browser shows data for the selected team(s) in the team browser. If the user has not selected any team in the team browser, the review browser provides overview by showing aggregated data of all teams.

Keyword List, Phrase List and Tag Clouds

Since one major component of the review is qualitative text feedbacks, we extracted keywords from text feedbacks based on their term frequency in bag-of-words model [2] to provide an overview of what students wrote the most in their reviews.

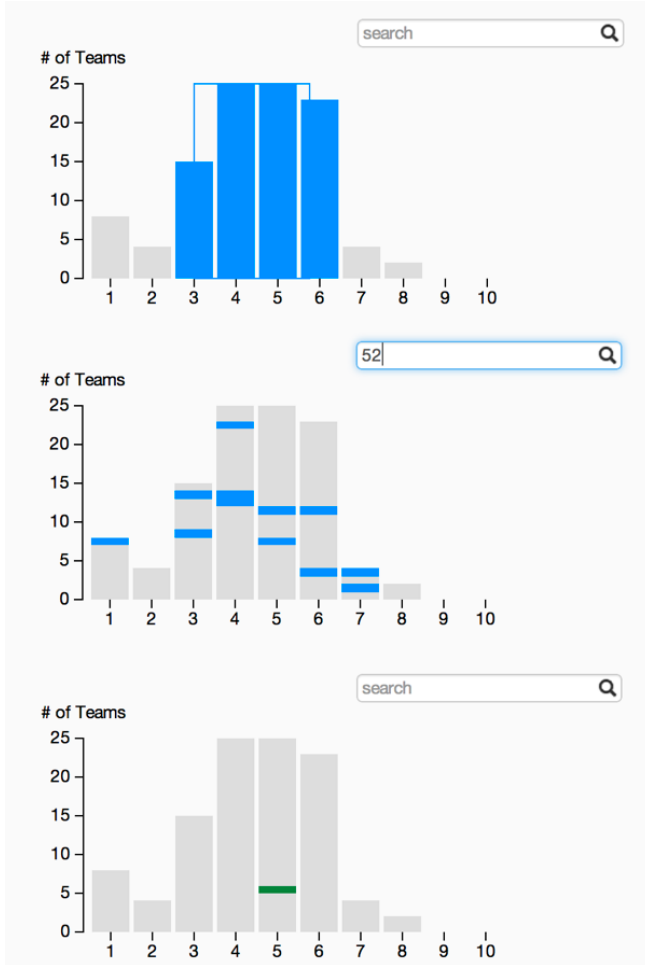


Figure 3. Different selection modes in the team browser. The top one shows brushing mode, which can be used to select teams within a range of scores. The middle one shows search mode, which user can use text by-description search. The bottom one show one team selection, by click on a block on the stacked bar chart.

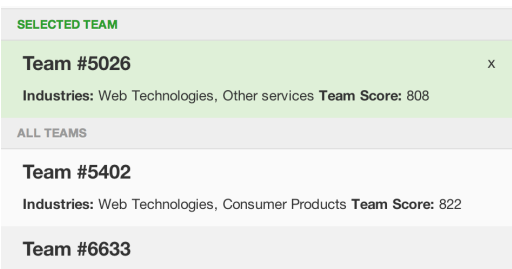


Figure 4. If the user select a team, either from the stacked bar chart or from the list, a selected team page will appear to help user recognize that a team has been selected.

There are three representations: keyword list, phrase list and tag cloud. In any of the three modes, users can click any word to see the actual occurrence of the word to further understand the context of the word.

The keyword list and phrase list are sorted lists by unigram and bigram term frequency respectively. Each word in the frequently list shows a bar representing their frequencies in the histogram on top of the word. There are two representations for the frequency lists: The first view shows keywords in four columns, in which each column represent keywords from different types of feedback in the original feedback capture grid. The second view (Upper in Figure 5) shows the total frequency of the words aggregated from all the feedback types. In this view, we use color to encoding the proportion of the keyword from each type.

Both frequency list views show the most frequent words on the top. User can also use the search box to search for unigram or bigram they are interested in. Meanwhile, the tag clouds view show four tag clouds of keywords group by each type in the feedback capture grid. Our tag clouds display all words horizontally for ease of reading.

Aggregate View and Individual Review View

In aggregate view, the text feedbacks for the selected teams are shown separately categorized by their types from feedback capture grid. This view provides instructors a quick way to read through a number of reviews for each feedback type at a time. As usual, it provides a search tool so users can search through the text feedbacks.

Individual reviews view is used to browse through the original reviews. It consists of two main parts as shown in Figure 8. First part on the left is a list of small multiples, and each small grid corresponds to single original review. We use color encoding in the small grid to represent the length of the review text: the longer the review text, the more intense the color is; so that the instructors can quickly read and select longer reviews which are potentially more valuable to read. The user can click small multiples to navigate through the original review list. Each original review item in the list contains the original scores distribution for the rubrics, the average score, and the feedback capture grid as in the original form.

IMPLEMENTATION NOTES

Peereviz was written in HTML, CSS and Javascript using the d3.js visualization toolkit [7] and a combination of javascript libraries including jQuery, Underscore.js, Backbone.js [1, 3, 6]. The d3s design that utilizes existing web standards including SVG and HTML has greatly facilitated our implementation.

One advantage of developing Peereviz in Javascript and HTML is that it could be easily integrated to any other online platforms since no other server-side setup is needed as opposed to developing the entire system using Ruby on rails and implementing different controllers to handle different pages.

As for the data processing, we performed SQL queries to get all the data we need and imported it directly at once. It also provides the plug-and-play convenience since users don't

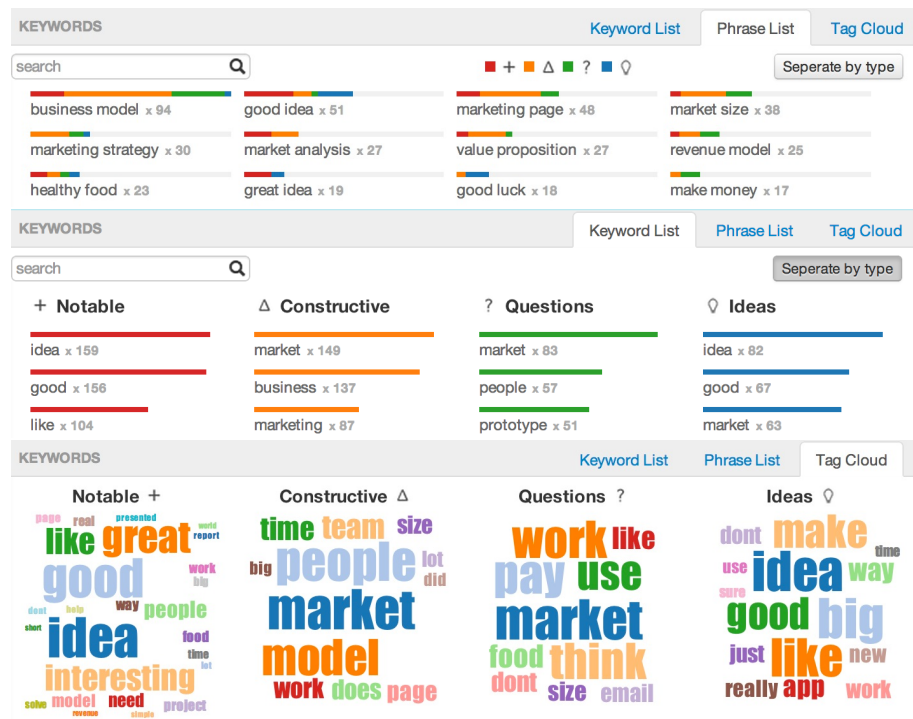


Figure 5. Three views of the keyword browser: (top) the phrase list in integrated mode, (middle) the keyword list in group-by-type mode, (bottom) the tag clouds grouped by type.

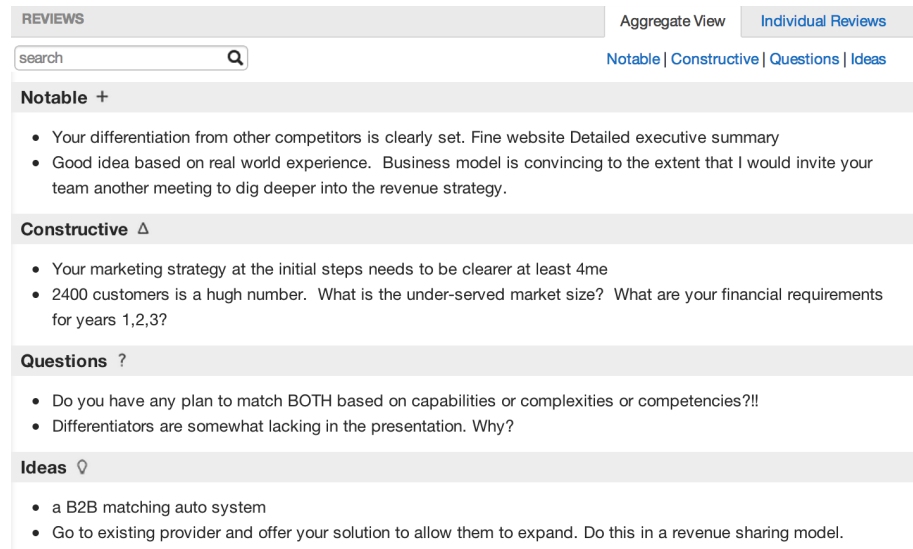


Figure 6. Aggregate View.

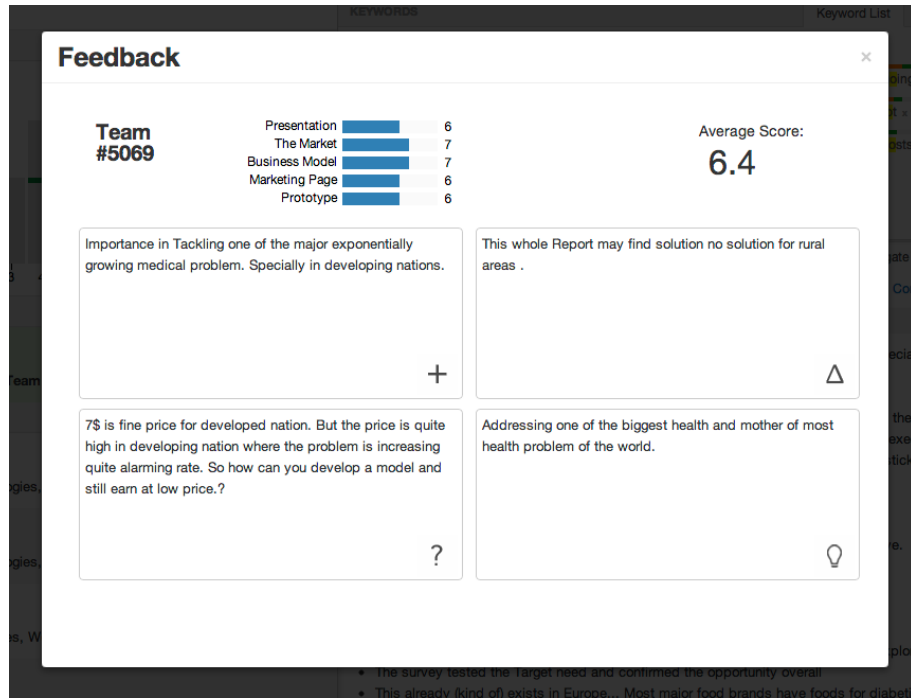


Figure 7. A Modal dialog showing original review rata after a user click on a review data in the Aggregate View.

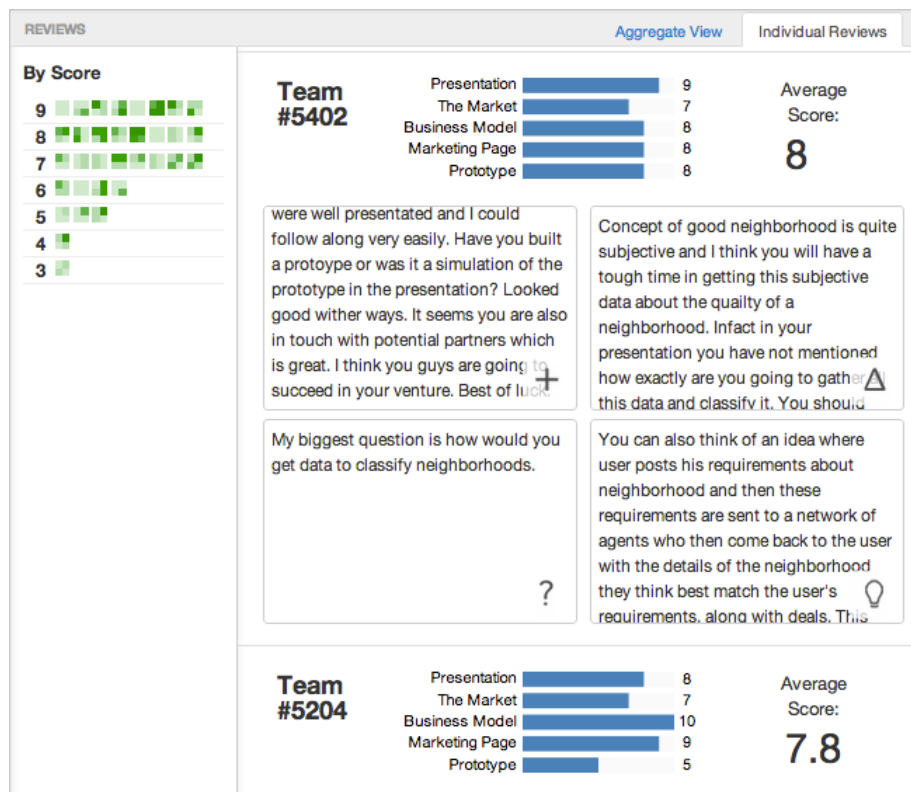


Figure 8. Individual Review View.

have to have a database or remote server to provide the data. But it might not be a good idea when the data size becomes too large. Directly loaded it from database or split the data into separate files and load them as needed might be another alternative but it might slow it down due to the data I/O as well.

DISCUSSION AND FUTURE WORK

This visualization has provided a way for course instructors to quickly browse the peer review data and get senses of the ongoing peer review activities.

We have presented Peereviz to Venture Lab administrator and let her play with it for 30 minutes. She has expressed her positive feedback of this tool and mentioned that she hope to use this to find out what are qualities of team that get good grades and what has gone wrong for teams with a low grade. She believes that the pattern discovered from this tool can provide helpful feedback for both future offering of the courses and the design of the platform. We plan to further evaluate this tool from real instructor usage in online classes after we integrate it into the Venture Lab platform.

In terms of improvement, one area that we can improve is the text visualization techniques. Our current work has only ranked it by the word frequency, while we could easily replace the underlying calculation for ranking the words. For example, we could instead use **TF-IDF model** [15] or other score calculations to find more expressive words.

We have also learned that visualization of the review also largely depends on the peer review form. Pre-structured form facilitates greatly in making the visualization. Our visualization benefits from the feedback capture employed in the platform. Thus, we suggest that providing scaffolding structure in the peer evaluation form does not only improve the review quality but also help the teaching teams to better understand the result as well.

Finally, another possible improvement is to enhance systems ability to compare different group of students by demographic such as GPA, academic major, or nationality.

CONCLUSION

In this paper, we described the visualization design of Peereviz, a system for exploring the peer review data in massive online open courses. We employed techniques from text visualization, simple language modeling and multiple coordinated views to help course instructors explore the result of assessments and learning outcomes as well as the insights from the overall peer review activities. In the near future, we are going to integrate this additional visualization feature to the Venture Lab platform.

ACKNOWLEDGMENTS

The authors of this paper would like to thank Farnaz Ronaghi and Venture Lab team for providing data and suggestions. Jeff Heer and CS448b (Data Visualization) course teaching team's comment and suggestions are also appreciated.

REFERENCES

1. backbone.js. <http://backbonejs.org/>.
2. Bag of Words Model. http://en.wikipedia.org/wiki/Bag-of-words_model.
3. jQuery. <http://jquery.com/>.
4. Stanford d.school's bootcamp bootleg. <http://dschool.stanford.edu/wp-content/uploads/2011/03/BootcampBootleg2010v2SLIM.pdf>.
5. The Year of the MOOC. <http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html?pagewanted=1&r=0>.
6. Underscore.js. <http://underscorejs.org/>.
7. Bostock, M., Ogievetsky, V., and Heer, J. D³ data-driven documents. *Visualization and Computer Graphics, IEEE Transactions on* 17, 12 (2011), 2301–2309.
8. Card, S., Mackinlay, J., and Shneiderman, B. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.
9. Carlson, P., and Berry, F. Calibrated peer review and assessing learning outcomes. In *FRONTIERS IN EDUCATION CONFERENCE*, vol. 2, STIPES (2003).
10. Cho, K., and Schunn, C. Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education* 48, 3 (2007), 409–426.
11. De La Harpe, B., Peterson, J., Frankham, N., Zehner, R., Neale, D., Musgrave, E., and McDermott, R. Assessment focus in studio: What is most prominent in architecture, art and design? *International Journal of Art & Design Education* 28, 1 (2009), 37–51.
12. Hyrynen, V., Hämäläinen, H., Ikonen, J., and Porras, J. Mypeerreview: an online peer-reviewing system for programming courses. In *Proceedings of the 10th Koli Calling International Conference on Computing Education Research*, ACM (2010), 94–99.
13. Kuo, B., Hentrich, T., Good, B., and Wilkinson, M. Tag clouds for summarizing web search results. In *Proceedings of the 16th international conference on World Wide Web*, ACM (2007), 1203–1204.
14. Liu, E., Lin, S., Chiu, C., and Yuan, S. Web-based peer review: the learner as both adapter and reviewer. *Education, IEEE Transactions on* 44, 3 (2001), 246–251.
15. Manning, C., Raghavan, P., and Schütze, H. *Introduction to information retrieval*, vol. 1. Cambridge University Press Cambridge, 2008.
16. Roberts, J. State of the art: Coordinated & multiple views in exploratory visualization. In *Coordinated and Multiple Views in Exploratory Visualization, 2007. CMV'07. Fifth International Conference on*, IEEE (2007), 61–71.
17. Word Count. <http://www.wordcount.org/main.php>.
18. Xiong, W., Litman, D., Wang, J., and Schunn, C. An interactive analytic tool for peer-review exploration.