Vi-ctory Project
The Visualization Factory

Midterm Project Report

Group 8
Evan Smith
Deepak Subramanian
Huanhuan Wang
Alex Chung

Description

The project is aiming to explore the best visual representations to describe and explain complex concepts. Our project is primarily an application project. We are building an application capable of crowd sourcing the production of semantically-rich visuals.

The project is focusing on concepts extracted from class Information 202 - Information Organization and Retrieval.

Social Components

Vi-ctory introduces a system that harnesses the power of the crowds to perform collaborative learning. Collaborative learning is a process that encourages stakeholders to support each other in learning and acquiring new knowledge. While crowdsourcing is beginning to gain the attention of academics, most of today's services focus on displacing the RSS feeds and email listings with an online Q&A platform. The challenge with such platform is the lack of commonality among all participants resulting in silos that only pay attention to their respective threads. Another challenge is the cognitive burden on the recipients of the "wisdom of the crowd" when the mode of communication is limited to text.

Therefore, Vi-ctory focuses on these two issues by introducing a shared platform where the discussions are based on the available lecture slides and contributors are encouraged to submit "smart arts" that captures the essence of the lecture materials.

During a lecture slide discussion, students might have a better explanation of the materials. Yet it is difficult to convey the message if the concept in their mind is based heavily on graphical image. Students would normally have to describe the image in length without much success. However, what if there is a channel for them for post smart arts or upload images along with annotation to get their points easily across. After all, an image speaks a thousand words and our team project is built around that the concept of sharing image/smart art can enrich the learning process while lowering the cognitive burden on the readers.

Yet, the system must also entice the participants with a level of achievement as well as enjoyment to keep them coming back for more. We plan to use a social computing approach to investigate what sort of responses gender more interactions and engagement by both the posters and the readers.

Finally, we believe Vi-ctory is going beyond the current offerings of co-learning platforms and contributing a unique answer to the challenges of collaborative learning online.

Project Stakeholders

For the purposes of this project, users of the Vi-ctory system include current students from School of Information, alumni as well as users interested in information technology, specifically in information retrieval. On one hand, users could create their own version of 'visual notes' for future reference. The 'visual notes' can be used as both reviewing material for the user and a quick way to explain the concept to broader audience. On the other hand, the platform enables users to collaboratively learn from others understanding about certain topics, which potentially could improve the effectiveness of

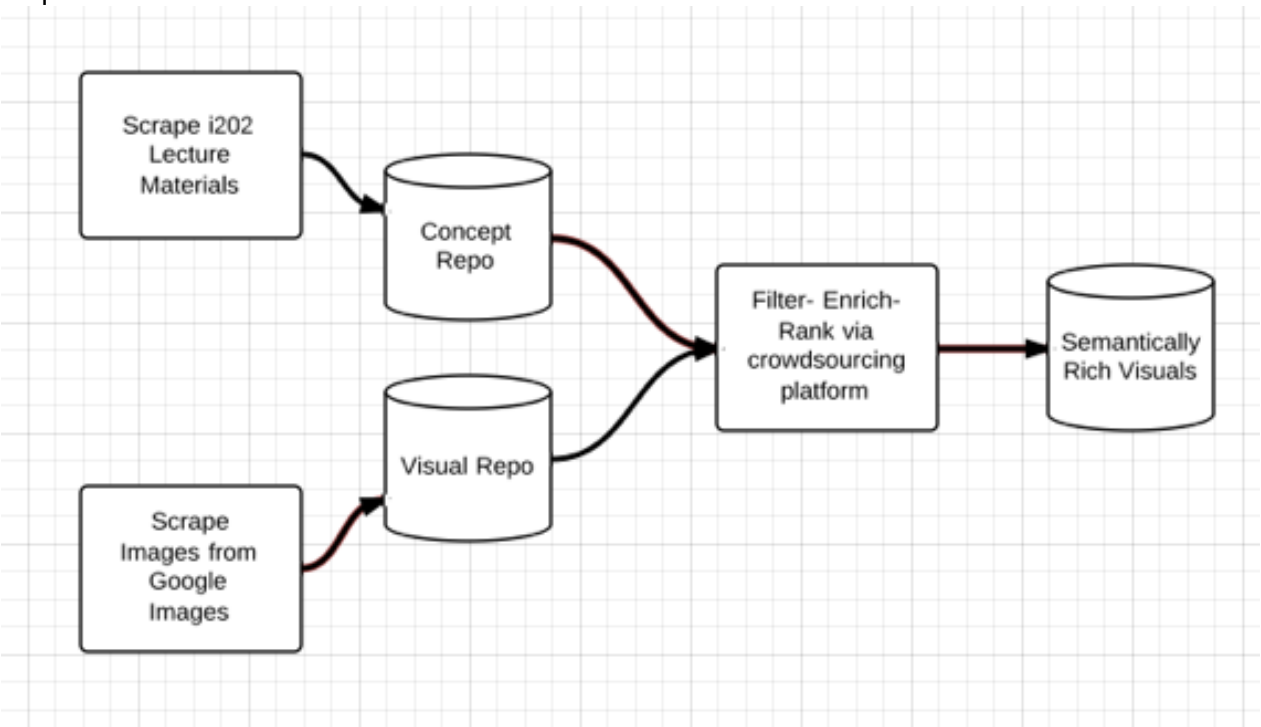
knowledge dissemination.

Another stakeholder to keep in mind is the professor of the i202 course. The professor stands to gain from his material being augmented with visual aides.

An important stakeholder of the system is Making Visible, a larger I School project that contains a repository of visuals, a Saas based editor and a marketplace for services. The visuals created in our crowd-powered visualization factory will be made available to the Making Visible repository.

System Overview

Our system is in essence a crowd-sourcing platform for the creation, organization and enhancement of visual aides. At a high level, the system will work in the following sequence.



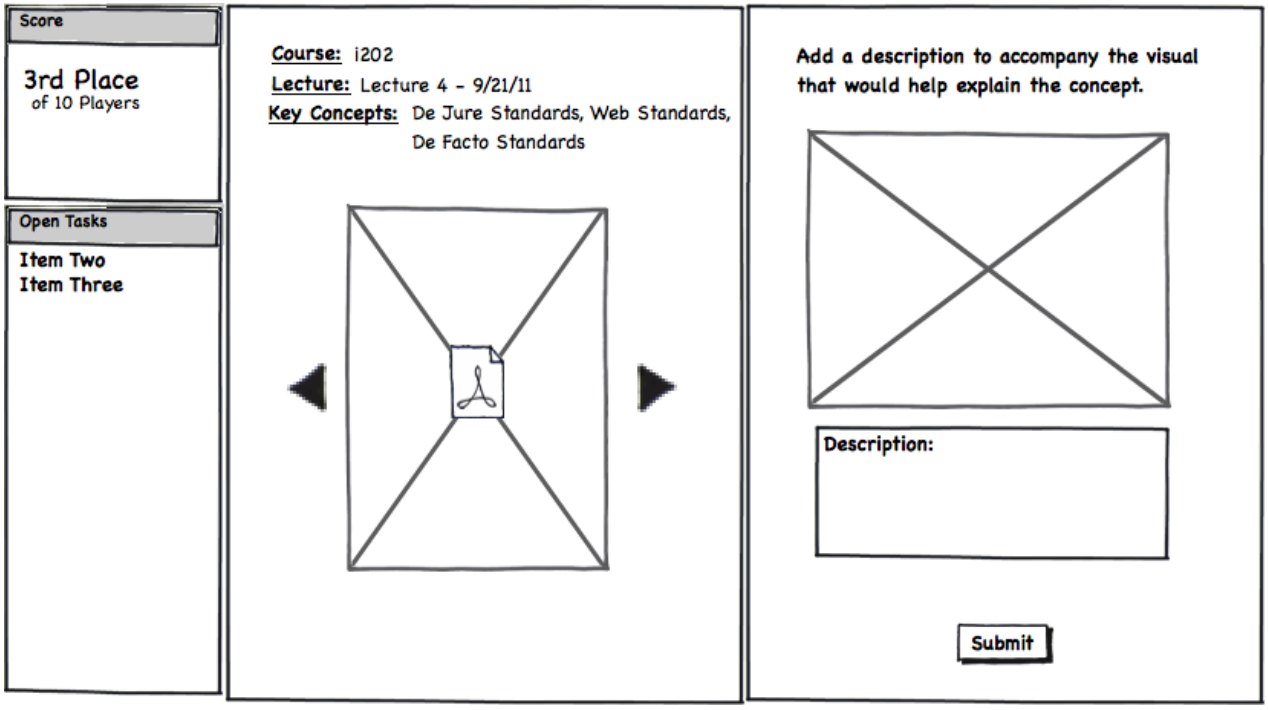
The process for filtering, enriching and ranking the visuals for the purpose of outputting high quality visuals will be facilitated on the crowd-sourcing platform that we create.

Through our brainstorming, we have decomposed the crowd-sourcing platform into a series of functions. Subsets of these functions are as follows:

- Select the best visual to represent an idea
- Rank the effectiveness of a visual and corresponding description at
- Add description to visual
- Draw annotation atop visual
- Enter concept that needs explanation
- Draw a draft of a visual

Our crowd-sourcing platform will coordinate users to perform these tasks in the appropriate series of steps such that the output is of high quality.

Below is a screenshot of one type of task our system will support. This example collects descriptions to accompany visuals. The UIs to support the other tasks will follow this general layout.



The top left “Score” section is a placeholder for the type of reward sub-system we will need to include to ensure users are adequately motivated to continue using the system

Progress Summary

Target Technologies and Paradigms

1. Python Web Development Platform: Django

2. Improving Image Recommendation Algorithm
Inputs:
Human
- Rating
- Descriptions
- Tags

Machine
Corpus of images

3. Algorithm
For each of the concepts, find the highest average rated images and show them to the user. This will keep on improving as more and more users use the system.
Use latent semantic analysis on each description field.
Use a combination of both the above ranks to decide on next image recommendation.

4. Output
A comprehensive rank for each image for a concept using human and machine intelligence

Planned Final Deliverable:
A functional prototype which will use human computation and machine computation to search for and prepare the visuals for a typical college syllabus
1) A web application which has at max 10 pages
1) Two set of inputs
Human metadata input
Image corpus input

Validation

The validation process is to provide evidence and quantitative results to support the usefulness of the system. An underlying assumption is that the hypothesis is independent of platforms.

1. Hypothesis
- Information with visual representation attracts more participation than those without visuals.
- Our system is capable of producing visuals that are able to bring clarity to complex issues

2. Data Analysis Method
Source: Twitter or Facebook
One of the reasons that we choose to use existing platforms to test the hypothesis is to precede the project in a parallel fashion. And because we assume the proposed hypothesis is independent of platforms, the quantitative result should be valid to prove our intention of building a visual learning platform will attract more participation comparing with current learning platform with only texts, with the assumption of all the other components are comparably presented.

3. Measurement:
Percentage of replies or ‘follows’ for comments with/without visuals

4. Method:
We will use statistical model to analyze the data. However, detailed plan is in development.

5. Expected Result:
The percentage of replies for comments with visuals is significantly different from the ones without visuals.

Risks and Challenges

There is a technology risk that the Django platform will prove unsuitable for the type of platform we are trying to build.

There is a risk that there will be insufficient participation by the expected users of the system.

There is a risk of intellectual property issues associated with the visuals.

Reference

INFO290-Social Computing: Social Media in Education
Bebo White, Stanford Linear Accelerator Lab

INFO290-Social Computing: Human Computation/Crowdsourcing
Irwin King, UC Berkeley School of Information

T. Segaran, *Programming Collective Intelligence: Building Smart Web 2.0 Applications*, 1st ed. O'Reilly Media, 2007.

Soylent: A Word Processor with a Crowd Inside. Bernstein, M., Little, G., Miller, R.C., Hartmann, B., Ackerman, M., Karger, D.R., Crowell, D., and Panovich, K. In Proceedings of UIST 2010.

Antin, Judd and Coye Cheshire. 2010. Readers and not Free-Riders: Reading as a Form of Participation in Wikipedia. Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work.

Osterloh, Margit and Bruno S. Frey. 2000. "Motivation, Knowledge Transfer, and Organizational Forms." Organization Science 11:538-550.

Kollock, Peter. 1999. "The Economies of Online Cooperation: Gifts and Public Goods in Cyberspace." Chapter 9 in Communities in Cyberspace, Kollock and Smith, eds.

Rafaeli, Sheizaf , and Daphne R. Raban. 2005. "Information sharing online: a research challenge." International Journal of Knowledge and Learning 1:62-79.

Ubois, Jeff. 2003. "Online Reputation Systems." Pp. 1-33 in Release 1.0: www.edventure.com.