Algorithms, Models and Systems for Eigentaste-Based Collaborative Filtering and Visualization

Tavi Nathanson

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2009-85
http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-85.html

May 26, 2009

Copyright 2009, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

Acknowledgement

These projects have been supported in part by the Berkeley Center for New Media and the NSF Graduate Research Fellowship Program.

Algorithms, Models and Systems for Eigentaste-Based Collaborative Filtering and Visualization

by Tavi Nathanson

Research Project

Submitted to the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, in partial satisfaction of the requirements for the degree of **Master of Science, Plan II**.

Approval for the Report and Comprehensive Examination:

Committee:

Professor Ken Goldberg

Research Advisor

(Date)

* * * * * * *

Professor Kimiko Ryokai

Second Reader

(Date)

Dedication

This report is dedicated to my grandfather, Illes Jaeger, who has given me unconditional love and support throughout my life. In many ways my academic accomplishments reflect his intellect and interest in engineering and technology. Despite the fact that he has never owned a computer, the amount he knows about computers never ceases to surprise me. Perhaps one day he will decide to use a computer, at which point I would be thrilled to recommend him a few recommender systems!

Abstract

We present algorithms, models and systems based on Eigentaste 2.0, a patented constant-time collaborative filtering algorithm developed by Goldberg et. al. [16]. Jester 4.0 is an online joke recommender system that uses Eigentaste to recommend jokes to users: we describe the design and implementation of the system and analyze the data collected. Donation Dashboard 1.0 is a new system that recommends non-profit organizations to users in the form of portfolios of donation amounts: we describe this new system and again analyze the data collected. We also present an extension to Eigentaste 2.0 called Eigentaste 5.0, which uses item clustering to increase the adaptability of Eigentaste while maintaing its constant-time nature. We introduce a new framework for recommending weighted portfolios of items using *relative* ratings as opposed to *absolute* ratings. Our Eigentaste Security Framework adapts a formal security framework for collaborative filtering, developed by Mobasher et. al. [33], to Eigentaste. Finally, we present Opinion Space 1.0, an experimental new system for visualizing opinions and exchanging ideas. Using key elements of Eigentaste, Opinion Space allows users to express their opinions and visualize where they stand relative to a diversity of other viewpoints. We describe the design and implementation of Opinion Space 1.0 and analyze the data collected. Our experience using mathematical tools to utilize and support the wisdom of crowds has highlighted the importance of incorporating these tools into fun and engaging systems. This allows for the collection of a great deal of data that can then be used to improve or enhance the systems and tools. The systems described are all online and have been widely publicized; as of May 2009 we have collected data from over 70,000 users. This master's report concludes with a summary of future work for the algorithms, models and systems presented.

Contents

1 Introduction

In this age of the Internet, information is abundant. In fact, it is so abundant that "information overload" is increasingly common. Recommender systems aim to reduce this problem by predicting the information that is likely to be of interest to a *particular* user and recommending that information. A collaborative filtering system is a recommender system that uses the likes and dislikes of other users in order to make its predictions and recommendations, as opposed to content-based filtering that relies on the content of the items.

Eigentaste 2.0 is a patented constant-time collaborative filtering algorithm developed by Goldberg et. al. [16]. In this report we present a number of algorithms, models and systems based on Eigentaste: Jester 4.0, Donation Dashboard 1.0, Eigentaste 5.0, a framework for recommending weighted portfolios of items using relative ratings, the Eigentaste Security Framework and Opinion Space 1.0. We will describe each algorithm, model and system, as well as describe our analysis of data collected from our systems.

Jester is an online joke recommender system that uses Eigentaste to recommend jokes to users, developed alongside Eigentaste by Goldberg et. al. Jester 4.0 is our new version of Jester that we architected from the ground up and released in November 2006; as of May 2009 it has collected over 1.7 million ratings from over 63,000 users. Jester 4.0 has been featured on Slashdot, the Chronicle of Higher Education, the Berkeleyan and other publications (see Appendix F for the articles).

Donation Dashboard is a second application of Eigentaste to the recommendation of non-profit organizations, where recommendations are in the form of portfolios of donation amounts. We released Donation Dashboard 1.0 in April 2008, and it has since collected over 59,000 ratings from over 3,800 users. It has been featured on ABC News, MarketWatch, Boing Boing and notable philanthropy news sources such as the Chronicle of Philanthropy and Philanthropy News Digest (see Appendix F).

Eigentaste 5.0 is an an extension to the Eigentaste 2.0 algorithm that uses item clustering to increase its adaptability while maintaining its constant-time nature. Our new framework for recommending weighed portfolios of items using *relative* ratings avoids the biases inherent in absolute rating systems. The Eigentaste Security Framework is an adaptation to Eigentaste of a formal framework for modeling attacks on collaborative filtering systems developed by Mobasher et. al. [33].

Finally, Opinion Space is a new system that builds upon our work with collaborative filtering systems by using key elements of the Eigentaste algorithm for the visualization of opinions and exchange of ideas. It allows users to express their opinions and visualize where they stand relative to a diversity of other viewpoints. Opinion Space 1.0 was released on April 22, 2009 and has collected over 18,000 opinions from over 3,700 users. It has been featured by publications including Wired and the San Francisco Chronicle (see Appendix F).

2 Background

As described in Section 1, recommender systems predict information that is likely to be of interest to a particular user and recommend that information. A collaborative filtering system is a kind of recommender system that makes its predictions and recommendations using the preferences of other users.

Eigentaste 2.0, also referred to simply as Eigentaste, is a constant-time collaborative filtering algorithm that collects real-valued ratings of a “gauge set” of items that is rated by all users [16]. This gauge set consists of the highest variance items in the system, as this allows for the quick identification of other users with similar preferences.

By requiring all users to rate the gauge set of items, Eigentaste 2.0 handles the cold start problem for new users. It is a common problem in collaborative filtering systems where a lack of information about new users results in an inability to make good recommendations, but the gauge set ratings provide that information. This requirement that all users rate the gauge set also ensures that the gauge set ratings matrix is dense, which allows for a straightforward application of principal component analysis (described below).

Collecting real-valued, continuous ratings is advantageous for the following reasons: discretization effects in matrix computations are avoided, taste is captured more precisely, and users find it easier to provide continuous ratings [16].

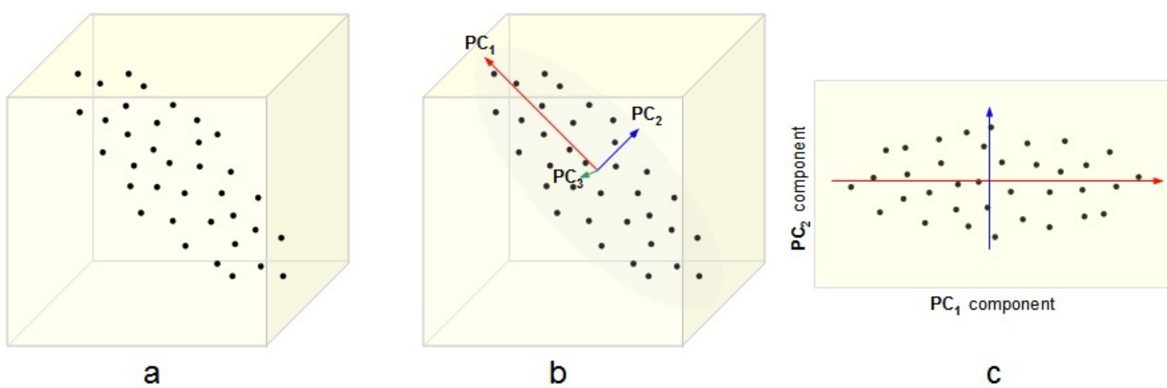


Figure 1: Illustration of principal component analysis.

Principal component analysis (PCA) is a mathematical procedure introduced in 1901 by Pearson [39] and generalized to random variables by Hotelling [23], and it is a key component of the Eigentaste algorithm. It works by transforming data to a new coordinate system of orthogonal “principal components,” where the first principal component represents the direction of highest variance in the data, the second principal component represents the direction of second highest variance in the data, and so on [24]. It can be used for dimensionality reduction by projection high-dimensional data points onto a space comprised of only the most primary principal components, such as the plane consisting of the first and second principal components. One method of performing PCA involves finding the eigenvectors of the the covariance matrix of the data.

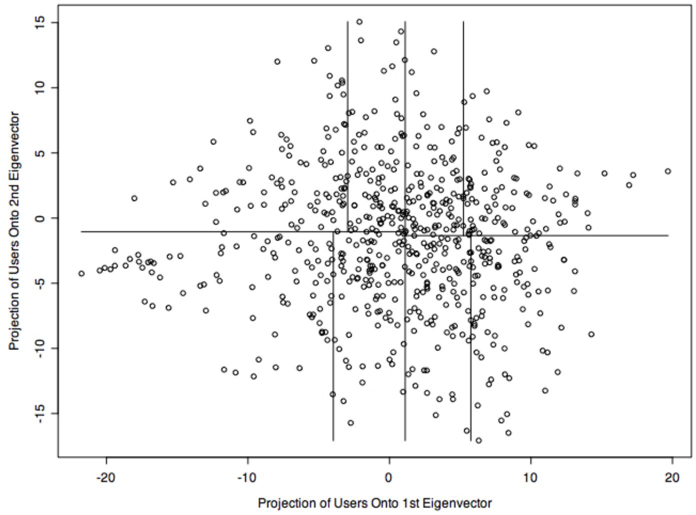


Figure 2: Recursive rectangular clustering resulting in 8 clusters.

Eigentaste 2.0 is divided into an offline phase and an online phase. Offline, it first applies PCA to the ratings matrix of the gauge set, and each user is projected onto the resultant eigenplane comprised of the first and second principal components. Due to a high concentration of users around the origin, a median-based algorithm referred to as recursive rectangular clustering is used on the lower-dimensional space to divide users into clusters (see Figure 2 for an illustration). This method ensures that cluster cell size decreases near the origin, resulting in evenly populated clusters.

Online, a new user's position on the eigenplane is determined in constant time via a dot product with the principal components that were generated in the offline phase. The user falls into one of the clusters defined offline, and the predicted rating for user u of item i is the average rating of item i by user u's cluster.

3 Related Work

Collaborative Filtering D. Goldberg et. al. [15] coined the term "collaborative filtering" in the context of filtering documents by their collective "annotations." Resnick et. al. [46] describe GroupLens, a news recommender and one of the earliest collaborative filtering systems developed. Sarwar et. al. [48] analyze different item-based collaborative filtering algorithms and argue that they provide better quality and performance than user-based algorithms. Item-based algorithms generate predictions by measuring similarity between items while user-based algorithms use similarity between users. Herlocker et al. [21] give an extensive analysis of methods for evaluating collaborative filtering systems, and Thornton maintains an excellent online survey of the literature on collaborative filtering: [51]. A few companies that employ collaborative filtering commercially are Netflix, Amazon, and TiVo. Many new algorithms have been developed as a result of the Netflix Prize, an ongoing competition for the best collaborative filtering algorithm using the Netflix dataset.

Eigentaste 2.0 Eigentaste 2.0 specifically scales well because in constant online time, it matches new users with user clusters that are generated offline. Linden et al. [29] at Amazon use an *item-based* collaborative filtering algorithm that scales independent of the number of users and the number of items. Rashid et al. [45] propose an algorithm, ClustKnn, that combines clustering (*model-based*) with a nearest neighbor approach (*memory-based*) to provide scalability as well as accuracy. Earlier, Pennock et al. [41] evaluated the method of *personality diagnosis*, another technique that combines *model-based* and *memory-based* approaches by aiming to determine a user's personality type. Deshpande and Karypis [11] evaluate *item-based* collaborative filtering algorithms and show that they are up to two orders of magnitude faster than *user-based* algorithms.

We deal with rating sparseness in Eigentaste 2.0 by ensuring that all users rate the common set of items, but there are many alternative solutions to this problem. Wilson et al. [58] approach sparseness by using data mining techniques to reveal implicit knowledge about item similarities. Xue et al. [60], on the other hand, fill in missing values by using user clusters as smoothing mechanisms. Wang et al. [56] fuse the ratings of a specific item by many users, the ratings of many items by a certain user, and data from similar users (to that user) rating similar items (to that item) in order to predict the rating of that item by the given user. This approach both deals with sparsity and combines *user-based* and *item-based* collaborative filtering. Herlocker et al. [21] evaluate several different approaches to dealing with sparseness.

As described in Section 2, Eigentaste 2.0 is well-suited to handle the cold-start problem for new users. Park et al. [38] use *filterbots*, bots that algorithmically rate items based on item or user attributes, to handle cold-start situations. Schein et al. [50] discuss methods for dealing with the cold-start problem for new items by using existing item attributes, which is symmetric to the cold-start problem for new users when user attributes are available. Rashid et al. [44] survey six techniques for dealing with this situation. Cosley et al. [10] investigate the rating scales used with recommender interfaces.

Other improvements to Eigentaste 2.0 include Kim and Yum's [25] iterative PCA approach that eliminates the need for a common set of items. Lemire's [28] scale and translation invariant version of the Eigentaste 2.0 algorithm improves NMAE performance by 17%.

Jester Many websites provide large databases of jokes, such as as Comedy Central's `Jokes.com` and JibJab's JokeBox; however, none of them provide user-specific recommendations. Ziv [63] categorize taste in humor using predefined categories, while we avoid such categories with Eigentaste and rely solely on user ratings. The International Journal of Humor Research [36] investigates the understanding of humor.

We use item similarity metrics in order to analyze data collected from Jester. Vozalis and Margaritis [54], as well as Ding et al. [12], compare the prediction errors resulting from using adjusted cosine similarity and Pearson correlation as item similarity metrics. Sarwar et al. [48], referenced earlier, also compare the cosine similarity in addition to the first two. Herlocker et al. [20] compare prediction errors resulting from using mean squared difference, Spearman rank correlation, and Pearson correlation as user similarity metrics.

Donation Dashboard There are also many websites and organizations that provide information about non-profit organizations, such as GuideStar, Charity Navigator, the BBB Wise Giving Alliance and the American Institute of Philanthropy. These websites provide a wealth of information about any specific non-profit organization and also provide rankings to make it easier for people to find the best non-profit organizations to donate to. Charity Navigator, for example, features top ten lists of non-profit organizations; however, user-specific recommendations are not available.

There are several existing examples of recommender systems that form an implicit portfolio of recommendations that may include items already rated by the user. In music recommenders, for example, the collection of songs can be implicitly viewed as a portfolio of recommendations, weighted by the frequency at which each song will be played. Anglade et. al. present a graph-based approach to create personalized music channels that broadcast the music shared by community members in a peer-to-peer fashion [2].

Ali and van Stam describe the TiVo recommender system in [1] and propose the idea of portfolio-based recommendations, arguing that the system should strive to recommend the best *set* of items as opposed to maximizing the probability that individual items will each be rated highly. Recommendation lists generated by the latter strategy tend to exhibit low levels of diversification; that is, items of the user's favorite and/or most frequently rated genre are recommended more frequently, commonly referred to as the *portfolio effect*. Ziegler et al. study ways of mitigating the portfolio effect by improving topic diversification in recommendation lists [62]. They model and analyze properties of recommendation lists, including a metric for intra-list similarity that quantifies the list's diversity of topics. Zhang and Hurley [61] model the goals of maximizing diversity while maintaing similarity as a binary optimization problem and seek to find the best subset of items over all possible subsets.

While most recommender systems are "single-shot," conversational recommender systems utilize dialog between the system and the user, allowing users to continually refine their queries. Like Donation Dashboard, conversational recommenders will often present items that have been previously recommended. Recent papers on conversational recommender systems include: [18, 59, 6, 52]. McGinty and Smyth consider various forms of user feedback in [31].

We also label the rating scale of Donation Dashboard in a specific way so as to encourage higher variance in ratings. More generally, Cosley et al. [10] investigate the rating scales used with recommender interfaces and the effects of their designs on user actions.

Eigentaste 5.0 Related to our usage of item clustering in Eigentaste 5.0, Quan et al. [43] propose a collaborative filtering algorithm that adds item clusters to a *user-based* approach: when predicting an item's rating for a user, it looks at similar users with respect to ratings within that item's cluster. George and Merugu [14] use Bregman co-clustering of users and items in order to produce a scalable algorithm. Vozalis and Margaritis [55] compare other algorithms that combine *user-based* and *item-based* collaborative filtering.

Ziegler et al. [62], as cited above, propose a new metric for quantifying the diversity of items in a recommendation list that is used to address the portfolio effect. Konstan et al. [26] extend this work to apply recommender systems to information-seeking tasks. While these works improve user satisfaction by increasing topic diversity in recommendation lists, we do with Eigentaste 5.0 by dynamically reordering recommendations in response to new ratings.

Recommending Weighted Item Portfolios Using Relative Ratings In Section 7 we introduce a new graphical model for collaborative filtering motivated by social choice theory or group decision theory. Hochbaum and Levin derive an axiomatic and algorithmic approach to optimally combine a set of user preferences [22]. This is also known as group ranking, and is a key problem within group decision theory. The first connection between collaborative filtering and group decision theory was made by Pennock et. al. [40]. Satzger, B. et. al. present user similarity metrics that compare various forms of item preference lists [49].

We use relative ratings in our model because rating individual items on an absolute scale has significant dependencies on the order in which the items are presented; as mentioned earlier, Herlocker et al. analyze methods for evaluating collaborative filtering systems [21], and one specific issue discussed is the bias resulting from the order in which items are presented.

Eigentaste Security Framework Lam and Riedl [27] wrote one of the earlier works describing attacks on collaborative filtering and their effects. More recently, Bryan et. al. [7] introduced metrics originating in bioinformatics for the unsupervised detection of anomalous user profiles in collaborative filtering systems. Resnick and Sami [47] analyzed the amount of (potentially useful) information that must be discarded in order to create a manipulation-resistant collaborative filtering system. Massa and Avesani [30] described how they use explicit trust information to affect recommendations, which results in a good defense against ad hoc user profiles generated by attackers.

Mobasher et. al. present an in-depth analysis of security in recommender systems in [35]. They have contributed greatly to the area of security in collaborative filtering with the following papers, as well: [3, 34, 8, 9, 32].

Opinion Space Freeman's article [13] provides background on visualization in social network analysis, from hand-drawn to computer-generated. Viégas and Donath [53] explore two visualizations based on email patterns, a standard graph-based visualization and a visualization that depicts temporal rhythms, and find that the latter is much more intuitive. Pang and Lee [37] review several techniques for gathering and understanding political and consumer opinions. At Berkeley, Heer and boyd [19] developed the interactive "Vizster" visualization system. Vizster displays social network users as nodes in a graph; it is based on binary connectivity models and does not represent gradations of opinion.

The Exploratorium's "my evidence" project attempts to create collaborative belief maps based on the various types of evidence that might lead to any belief. "Poligraph" by the Washington Post maps users onto a two-dimensional plane based on their answers to two political questions, and allows them to compare their positions with public figures. The Citizen Participation Map by Studio Mind represents individuals, their opinions and other characteristics. Opposing Views is a system that lists various issues and arguments both for and again them in an easy to digest format. KQED's "You Decide" project is an online "devil's advocate" that responds to a user's beliefs by presenting facts that might sway him or her in the other direction; upon changing beliefs, facts are presented that might cause the user to change back, and so on.

4 Jester (System)

Jester is a online joke recommender: it uses Eigentaste to recommend jokes to users based on their ratings of the initial gauge set. Goldberg et. al. used jokes as an item class to explore whether humor could be successfully broken up into different clusters of taste, and whether an automated system could successfully recommend a funny joke [16]. Further, jokes are particularly suitable for Eigentaste because all users, without any prior knowledge, are able to evaluate a joke after reading the text of the joke. Thus, all users are able to rate the "gauge set."

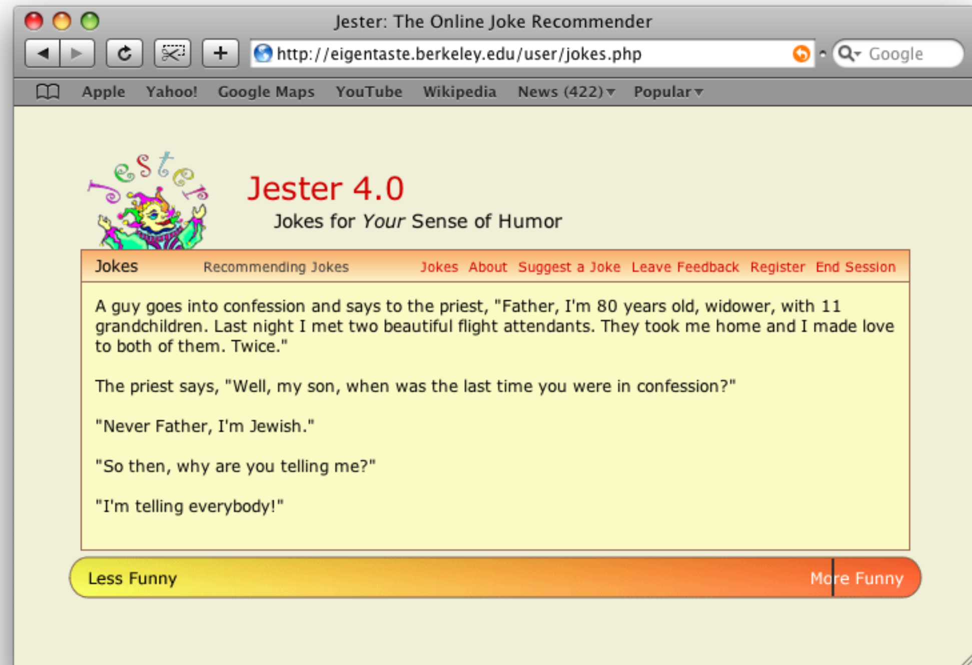


Figure 3: Screenshot of Jester 4.0, accessible at `http://eigentaste.berkeley.edu`.

Jester works as follows: users are presented with a set of initial jokes. This set is the "gauge set" of items in the Eigentaste algorithm 2, where items are now jokes. After they rate this initial set, Eigentaste uses those ratings to match them with a cluster of other users who have similar ratings, and recommends the highest rated jokes from that cluster.

4.1 Jester 1.0 through 3.0 Descriptions

Jester 1.0 [17] went online in October 1998, using a nearest neighbor algorithm. It was mentioned in Wired News on December 2, 1998, and was then picked up by various online news sites. This resulted in 7,000 registered users by December 25, 1998, and 41,350 page requests, overwhelming it with traffic.

Goldberg and his students developed Eigentaste in response. Eigentaste is a constant-time collaborative filtering algorithm as described in Section 2 and therefore highly scalable. Jester 2.0 went online on March 1, 1999, using Eigentaste to handle scalability. Jester 3.0 was the next version that added various improvements to the user interface and architecture. Between 1999 and 2003, Jester 3.0 was mentioned in publications such as USA Today, and collected 4.1 million ratings from over 73,000 users; the first dataset is available at:

`http://eigentaste.berkeley.edu/dataset`

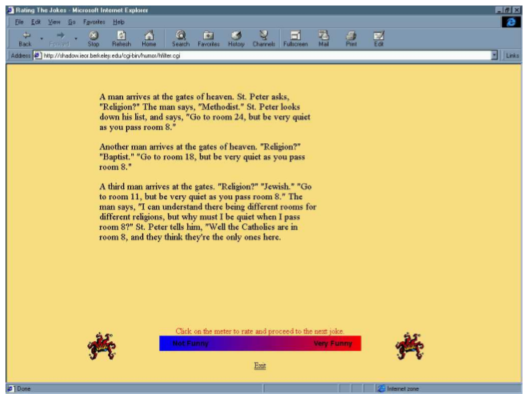


Figure 4: Screenshot of Jester 3.0.

A screenshot of Jester 3.0 is shown in Figure 4.

4.2 Jester 4.0 Description

We developed Jester 4.0 in the Summer of 2006, architected from the ground up based on the functionality of Jester 3.0. As was mentioned in Section 1, Jester 4.0 was released in November 2006, and as of May 2009 it has collected over 1.7 million ratings from over 63,000 users. The new version has been featured in various publications, such as Slashdot, the Chronicle of Higher Education and the Berkeleyan. Appendix A shows selected jokes from the system.

Jester 4.0 was architected using PHP and a MySQL database, along with JavaScript, HTML and CSS. In addition to the user-facing application, data processing scripts were also written using PHP. Jester 3.0, on the other hand, was written using C, CGI and HTML. The latest version of Jester is more automatic than its predecessor: running principal component analysis on Jester 3.0 involved running a few different C scripts, and the process of changing the principal components in the system involved manually editing a settings file. This whole process was automated in Jester 4.0 and reduced to a single command. Jester 4.0 also introduces an administration area for the approval of suggested jokes.

Jester 4.0 and the latest dataset are online at:

`http://eigentaste.berkeley.edu`
`http://eigentaste.berkeley.edu/dataset`

Below we describe Jester 4.0 in detail.

4.2.1 System Usage Summary

When a user enters Jester 4.0, he or she is presented with a short description of the system and can click on "Show Me Jokes!" to begin viewing jokes. The user is then presented with jokes one at a time; upon rating one joke the next joke is shown. The first set shown is the "gauge set" of jokes, as decribed earlier in Section 2.

Upon rating every joke in the gauge set, the user is shown the "seed set" of jokes, which is the set of jokes that has collected the fewest number of ratings in the system. This seeding process ensures that all jokes get rated enough; without it, the highly rated jokes might tend to be rated more and more, while other jokes might stagnate. Seeding is crucial upon the introduction of new jokes into the system; otherwise they might never collect positive ratings, and thus might never be recommended. This is a common strategy for ranking and recommendation systems; a similar system is the "Upcoming" section of `Digg.com` that presents new articles rather than highly rated ones. Jester 3.0 had a fixed set of jokes that was selected as the "seed set," and this set needed to be manually changed by an administrator. Jester 4.0 introduces dynamic seeding, where the seed jokes shown are those that have the fewest ratings at the moment the user is viewing jokes. Note that seed jokes are not currently being shown in the online version of Jester, which now uses both Eigentaste 2.0 and Eigentaste 5.0 as described in Section 6.

After rating the "seed set," the user is shown the "recommendation set": the system uses Eigentaste to predict the user's rating for every joke, and then sorts the jokes by their predicted ratings. Thus, every joke in the system that has not yet been displayed is a part of the recommendation set, where the jokes with highest predicted ratings are shown first. Jester 4.0 also interleaves seed jokes in the recommendation set in order to increase the seeding effect: for every 5 jokes recommended, 2 jokes are shown from the seed set.

Registration is optional in Jester 4.0, as compared with a required email and password in Jester 3.0. We made this decision in the hopes of collecting more data as users are often turned away by a registration page, and this has worked as planned.

4.2.2 Design Details

A joke in Jester 4.0, like in Jester 3.0, is presented as text and a rating bar that the user can click on to indicate his or her rating of the joke, as shown in Figure 3. We do not include any additional information such as the joke's title or author: this keeps the system quick and easy to use and avoids unnecessary biases.

Jester has always used continuous ratings in order to capture the gut feeling of the user, as opposed to many other systems that collect discrete ratings. Jester 4.0 retains continuous ratings, and, like Jester 3.0, displays the rating bar beneath the text of the current joke. In Jester 4.0 we introduce a JavaScript slider component to the rating bar: a vertical line is displayed above the bar, following the user's mouse to indicate that the position clicked affects the rating of the joke. Jester 3.0 used a simple HTML image map as the rating bar, but we found that people did not understand that the position that they clicked on indicating their rating of the joke. Jester 4.0 also introduces an "Instructions" display above the rating bar that is shown to users until they rate a certain amount of jokes, to further alleviate confusion regarding where to click.

Jester 4.0 also introduces a navigation menu, as shown in Figure 3. Jester 3.0 had a completely linear user interaction scheme: upon entering the system, the user was shown one joke at a time until there were none left, at which point he or she was asked for feedback. The navigation menu allows users to submit jokes, leave feedback or access the "About" page at any time. It also informs users of their state in the system, with messages such as "Displaying Initial Jokes (2/8)" and "Recommending Jokes."

A major change to Jester 4.0 is two-dimensional ratings, although this feature has never appeared online. We wanted users to be able to rate jokes on the *confidence* dimension, in addition to the "Less Funny" to "More Funny" dimension, in order to account for the difference between quick, "sloppy ratings" and ratings that were carefully thought out.

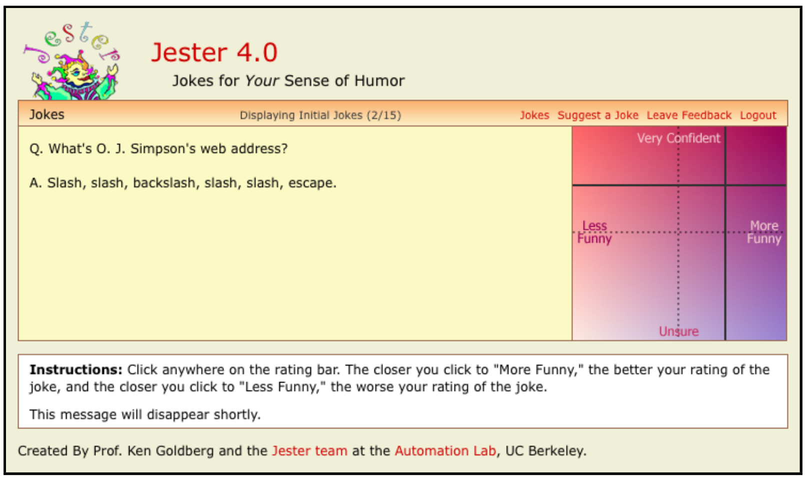


Figure 5: Screenshot of a two-dimensional rating bar in Jester 4.0.

We came up with several ideas for a two-dimensional rating bar, including both a triangular and square design. We implemented the square design, as shown in Figure 5. Initial user testing indicated that the addition of this second dimension made the process of rating each joke extremely time consuming, as a lot of thought was required in deciding on the rating. Users simply did not know where to click, which also resulted in a great deal of confusion. As a result, we decided not to add two-dimensional rating to the online system.

Figure 6: Mockup image of a blue vertical line indicating predicted rating in Jester 4.0.

Another addition that we considered was an extra vertical line, displayed on top of the rating bar, to indicate the system's predicted rating. We decided that displaying such a predicted rating *before* the user rated the joke would significantly bias the user's rating, but we gave serious thought to displaying it *after* the user rated the joke. Figure 6 shows the design of what would be displayed to the user upon rating a joke, where the blue vertical line represents the predicted rating. As it turned out, this was still problematic as it allowed users to see patterns in the predictions. For example, Eigentaste 2.0 always recommends jokes in descending order of their predicted ratings; thus, users would see the prediction bar shift more and more to the left as they progressed through the system, and would be able to guess the next prediction. This would have introduced bias into the ratings, so we did not implement this feature.

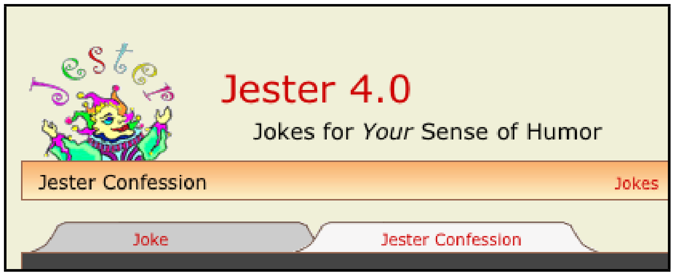


Figure 7: Screenshot of tabs in Jester 4.0.

In addition to displaying information on top of the rating bar, we also experimented with displaying data about each joke in a separate tab called "Jester Confession," as shown in Figure 7. Possible data might include statistics about the number of ratings, the average rating, etc. We implemented this feature, but once again did not introduce it into the online system due to concern about biasing the ratings.

4.2.3 Populating the System

Jester 4.0 uses many of the jokes from Jester 3.0, in addition to 30 new ones. Certain jokes from Jester 3.0 were not included in Jester 4.0, such as those that were outdated. In searching for the new jokes, we tried to select a very diverse set of jokes that would appeal to specific audiences. We did not want jokes that would appeal to all audiences, as this would not exploit the user-specific nature of a recommender system.

Some of the new jokes appealed to specific countries, others appealed to specific age groups, and so on. Three of the five highest variance jokes in the system are from this new set, and Figure 8 shows the highest variance joke in the system:

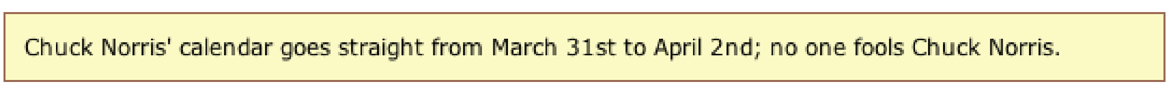


Figure 8: Joke 140 from Jester 4.0.

Other high variance jokes are shown in Appendix A.

4.3 Jester 4.0 Data Collected

We hypothesized that because jokes are generally considered in a positive light, the average rating in Jester would be positive. At the time of this writing the average rating is 1.62 (in the interval [-10.00, +10.00]), where the histogram of all ratings is shown in Figure 9. As is shown in the histogram, users did indeed rate positively far more often than they rated negatively. The

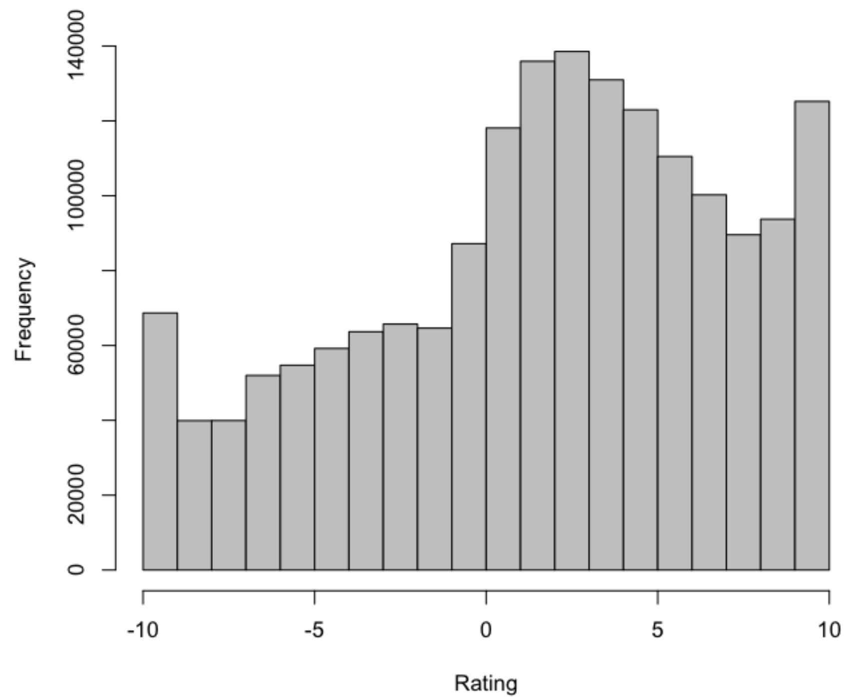


Figure 9: Histogram of Jester 4.0 ratings (in the interval [-10.00, +10.00]).

histogram also shows a general tendency for users to rate slightly to the right of the center, or to rate extremely positively.

Appendix A shows the top five highest rated jokes, the top five lowest rated jokes and the top five highest variance jokes.

4.4 Jester 4.0 Item Similarity

After building Jester and collecting data, we became curious whether items with similar ratings have similar content as well, where items are jokes in Jester. In order to investigate this we consider five standard similarity metrics: Pearson correlation, cosine similarity, mean squared difference, Spearman rank correlation, and adjusted cosine similarity. For each metric, we report the five most similar joke pairs based on ratings similarity, referenced by joke ID. Each pair is labeled with the other metrics that included this pair among their top five.

4.4.1 Notation

U	the set of all users
I	the set of all items
$U(i, j)$	the set of all users who have rated item i and item j
$\vec{i}$	the vector containing item i's ratings
$r_{u,i}$	user u's rating of item i
$\overline{r_i}$	the mean rating of item i
$\overline{r_u}$	user u's mean item rating
$k_{u,i}$	user u's rank of item i
$\overline{k_i}$	the mean rank of item i
$\overline{k_u}$	user u's mean item rank

Items are ranked such that the user's highest rated item corresponds to rank 1. Tied items are given the average of their ranks [20].

4.4.2 Metrics

Pearson Correlation The Pearson correlation of item i and item j is defined as follows:

$$P(i, j) = \frac{\sum_{u \in U(i,j)} (r_{u,i} - \overline{r_i})(r_{u,j} - \overline{r_j})}{\sqrt{\sum_{u \in U(i,j)} (r_{u,i} - \overline{r_i})^2} \sqrt{\sum_{u \in U(i,j)} (r_{u,j} - \overline{r_j})^2}} \tag{1}$$

Below are the top 5 pairwise clusters using this metric:

$P(86, 94)$	$= 0.650904$	[Pearson correlation, mean squared difference, adjusted cosine similarity]
$P(123, 140)$	$= 0.627227$	[Pearson correlation]
$P(138, 139)$	$= 0.605676$	[Pearson correlation]
$P(30, 85)$	$= 0.601248$	[Pearson correlation]
$P(68, 148)$	$= 0.587297$	[Pearson correlation]

Cosine Similarity The Cosine similarity of item i and item j is defined as follows:

$$C(i, j) = \frac{\vec{i} \cdot \vec{j}}{|\, \vec{i} \,||\, \vec{j} \,|} \tag{2}$$

Below are the top 5 pairwise clusters using this metric:

$C(114, 117)$	$= 0.720056$	[Cosine similarity, mean squared difference]
$C(53, 127)$	$= 0.703861$	[Cosine similarity, adjusted cosine similarity]
$C(105, 117)$	$= 0.696183$	[Cosine similarity, mean squared difference]
$C(53, 69)$	$= 0.689465$	[Cosine similarity]
$C(53, 117)$	$= 0.682699$	[Cosine similarity]

Mean Squared Difference The Mean squared difference (reverse of similarity) of item i and item j is defined as follows:

$$M(i,j) = \frac{\sum_{u \in U(i,j)} (r_{u,i} - r_{u,j})^2}{\mid U(i,j) \mid} \tag{3}$$

Below are the top 5 pairwise clusters using this metric:

$M(114,117)$	$= 20.5012$	[Cosine similarity, mean squared difference]
$M(86,94)$	$= 20.8009$	[Pearson correlation, mean squared difference, adjusted cosine similarity]
$M(105,117)$	$= 21.0734$	[Cosine similarity, mean squared difference]
$M(72,134)$	$= 21.1944$	[Mean squared difference]
$M(108,117)$	$= 21.3194$	[Mean squared difference]

Spearman Rank Correlation The Spearman rank correlation of item i and item j is defined very similarly to the Pearson correlation (described below). The only difference is that the Spearman rank correlation uses item ranks instead of item ratings. It is defined as follows:

$$S(i,j) = \frac{\sum_{u \in U(i,j)} (k_{u,i} - \overline{k_i})(k_{u,j} - \overline{k_j})}{\sqrt{\sum_{u \in U(i,j)} (k_{u,i} - \overline{k_i})^2} \sqrt{\sum_{u \in U(i,j)} (k_{u,j} - \overline{k_j})^2}} \tag{4}$$

Below are the top 5 pairwise clusters using this metric:

$S(7,13)$	$= 0.809066$	[Spearman rank correlation]
$S(15,16)$	$= 0.798144$	[Spearman rank correlation]
$S(7,15)$	$= 0.796643$	[Spearman rank correlation]
$S(16,18)$	$= 0.79648$	[Spearman rank correlation]
$S(7,8)$	$= 0.792346$	[Spearman rank correlation]

Adjusted Cosine Similarity The Adjusted cosine similarity of item i and item j is defined as follows:

$$A(i,j) = \frac{\sum_{u \in U(i,j)} (r_{u,i} - \overline{r_u})(r_{u,j} - \overline{r_u})}{\sqrt{\sum_{u \in U(i,j)} (r_{u,i} - \overline{r_u})^2} \sqrt{\sum_{u \in U(i,j)} (r_{u,j} - \overline{r_u})^2}} \tag{5}$$

Below are the top 5 pairwise clusters using this metric:

$A(124,141)$	$= 0.525723$	[Adjusted cosine similarity]
$A(53,127)$	$= 0.513244$	[Cosine similarity, adjusted cosine similarity]
$A(86,94)$	$= 0.492048$	[Pearson correlation, mean squared difference, adjusted cosine similarity]
$A(58,124)$	$= 0.491805$	[Adjusted cosine similarity]
$A(57,58)$	$= 0.481541$	[Adjusted cosine similarity]

4.4.3 Analysis

In order to provide intuition, we list pairs of jokes with high similarity; for example, Jokes 86 and 94 had the highest similarity using the Pearson correlation measure. Joke pairs that are not displayed can be viewed at:

http://eigentaste.berkeley.edu/viewjokes

Warning: some of these jokes have ethnic, sexist, or political content (some have all three). Although we tried our best, it is hard to eliminate all such factors and retain a sense of humor.

Joke 86

A neutron walks into a bar and orders a drink. "How much do I owe you?" the neutron asks.

The bartender replies, "For you, no charge."

Joke 94

Two atoms are walking down the street when one atom says to the other, "Oh, my! I've lost an electron!"

The second atom says, "Are you sure?"

The first replies, "I'm positive!"

Joke 123

When most people claim to be "killing time", it's only an expression. When Chuck Norris kills time, the minutes actually cease to exist.

Joke 140

Chuck Norris' calendar goes straight from March 31st to April 2nd; no one fools Chuck Norris.

Joke 138

WASHINGTON (Reuters) - A tragic fire on Monday destroyed the personal library of President George W. Bush. Both of his books have been lost.

Presidential spokesman Ari Fleischer said the president was devastated, as he had not finished coloring the second one.

Joke 139

In a Veteran's Day speech, President Bush vowed, "We will finish the mission. Period." Afterwards, he was advised that he doesn't have to read the punctuation marks.

Joke 30
Q: What's the difference between a lawyer and a plumber? A: A plumber works to unclog the system.
Joke 85
Q: How many presidents does it take to screw in a light bulb? A: It depends upon your definition of screwing a light bulb.

Joke 68
A man piloting a hot air balloon discovers he has wandered off course and is hopelessly lost. He descends to a lower altitude and locates a man down on the ground. He lowers the balloon further and shouts, "Excuse me, can you tell me where I am?" The man below says, "Yes, you're in a hot air balloon, about 30 feet above this field." "You must work in Information Technology," says the balloonist. "Yes I do," replies the man. "And how did you know that?" "Well," says the balloonist, "what you told me is technically correct, but of no use to anyone." The man below says, "You must work in management." "I do," replies the balloonist, "how did you know?" "Well," says the man, "you don't know where you are, or where you're going, but you expect my immediate help. You're in the same position you were before we met, but now it's my fault!"
Joke 148
Recently a teacher, a garbage collector, and a lawyer wound up together at the Pearly Gates. St. Peter informed them that in order to get into Heaven, they would each have to answer one question. St. Peter addressed the teacher and asked, "What was the name of the ship that crashed into the iceberg? They just made a movie about it." The teacher answered quickly, "That would be the Titanic." St. Peter let him through the gate. St. Peter turned to the garbage man and, figuring Heaven didn't really need all the odors that this guy would bring with him, decided to make the question a little harder: "How many people died on the ship?" Fortunately for him, the trash man had just seen the movie. "1,228," he answered. "That's right! You may enter." St. Peter turned to the lawyer: "Name them."

Joke 114
Sherlock Holmes and Dr. Watson go on a camping trip, set up their tent, and fall asleep. Some hours later, Holmes wakes his faithful friend. "Watson, look up at the sky and tell me what you see." Watson replies, "I see millions of stars." "What does that tell you?" Watson ponders for a minute. "Astronomically speaking, it tells me that there are millions of galaxies and potentially billions of planets. Astrologically, it tells me that Saturn is in Leo. Timewise, it appears to be approximately a quarter past three. Theologically, it's evident the Lord is all-powerful and we are small and insignificant. Meteorologically, it seems we will have a beautiful day tomorrow. What does it tell you?" Holmes is silent for a moment, then speaks. "Watson, you idiot, someone has stolen our tent."
Joke 117
A man joins a big corporate empire as a trainee. On his very first day of work, he dials the pantry and shouts into the phone: "Get me a coffee, quickly!" The voice from the other side responds, "You fool, you've dialed the wrong extension! Do you know who you're talking to, dumbo?" "No," replied the trainee. "It's the CEO of the company, you fool!" The trainee shouts back, "And do YOU know who YOU are talking to, you fool?!" "No." replied the CEO indignantly. "Good!" replied the trainee, and puts down the phone.

Joke 7
How many feminists does it take to screw in a light bulb? That's not funny.
Joke 13
They asked the Japanese visitor if they have elections in his country. "Every morning," he answers.

Joke 124
Person 1: Hey, wanna hear a great knock-knock joke? Person 2: Sure, What is it? Person 1: Okay, you start. Person 2: Knock-knock. Person 1: Who's there? Person 2: ... Person 1: Hah!
Joke 141
Jack Bauer can get McDonald's breakfast after 10:30.

We were surprised by how differently the metrics perform. For example, the top four pairwise clusters produced by Pearson correlation are pairs $(86, 94)$, $(123, 140)$, $(138, 139)$, and $(30, 85)$. The first pair are both about elementary particles, the second pair are both about Chuck Norris, the third pair are both about George Bush, and the fourth pair are both about the US legal/political system. This reflects that users' ratings are consistent for jokes with similar content. The converse does not necessarily follow: a pair of jokes with very different content and structure may get consistent ratings. The 5th most similar pair resulting from the Pearson correlation, joke pair (68, 148), is not that similar in terms of content, although one might argue that both jokes have cynical views of bureaucrats: managers (68) and lawyers (148), respectively.

None of the other similarity metrics found pairs with as much similarity in content as the Pearson correlation. The Adjusted cosine similarity metric takes users' different scales into account by subtracting a user's average item rating from specific ratings, and has been found to perform even better than Pearson correlation within item-based collaborative filtering [48]. The two metrics that had the most top-five pairs in common were Cosine similarity and Mean squared difference. Cosine similarity ignores the differences in rating scales between users [48], as does Mean squared difference.

5 Donation Dashboard (System)

Donation Dashboard is our second application of Eigentaste that recommends non-profit organizations to users in the form of a weighted portfolio of donation amounts. It is motivated by the fact that there are over 1 million registered non-profit institutions in the United States, but effectively allocating funds among these causes can be a daunting task. Many non-profit organizations do not have the resources to effectively advertise their causes; as a result, most people have only heard of a select few non-profit organizations.

Figure 10: Screenshot of Donation Dashboard 1.0, accessible at `http://dd.berkeley.edu`.

5.1 Donation Dashboard 1.0 Description

We began development of Donation Dashboard in the Fall of 2007, and Donation Dashboard 1.0 launched on April 21, 2008. As of May 2009 it has collected over 59,000 ratings of 70 non-profit organizations from over 3,800 users. As mentioned in Section 1, it has been covered by ABC News, MarketWatch, Boing Boing and other notable news sources. Appendix B lists selected organizations from the system. Donation Dashboard 1.0 was built upon the Jester 4.0 code base, using PHP and a MySQL database along with JavaScript, HTML and CSS.

Donation Dashboard 1.0 and the latest dataset are online at:

`http://dd.berkeley.edu`
`http://dd.berkeley.edu/dataset`

Below we describe Donation Dashboard 1.0 in detail.

5.1.1 System Usage Summary

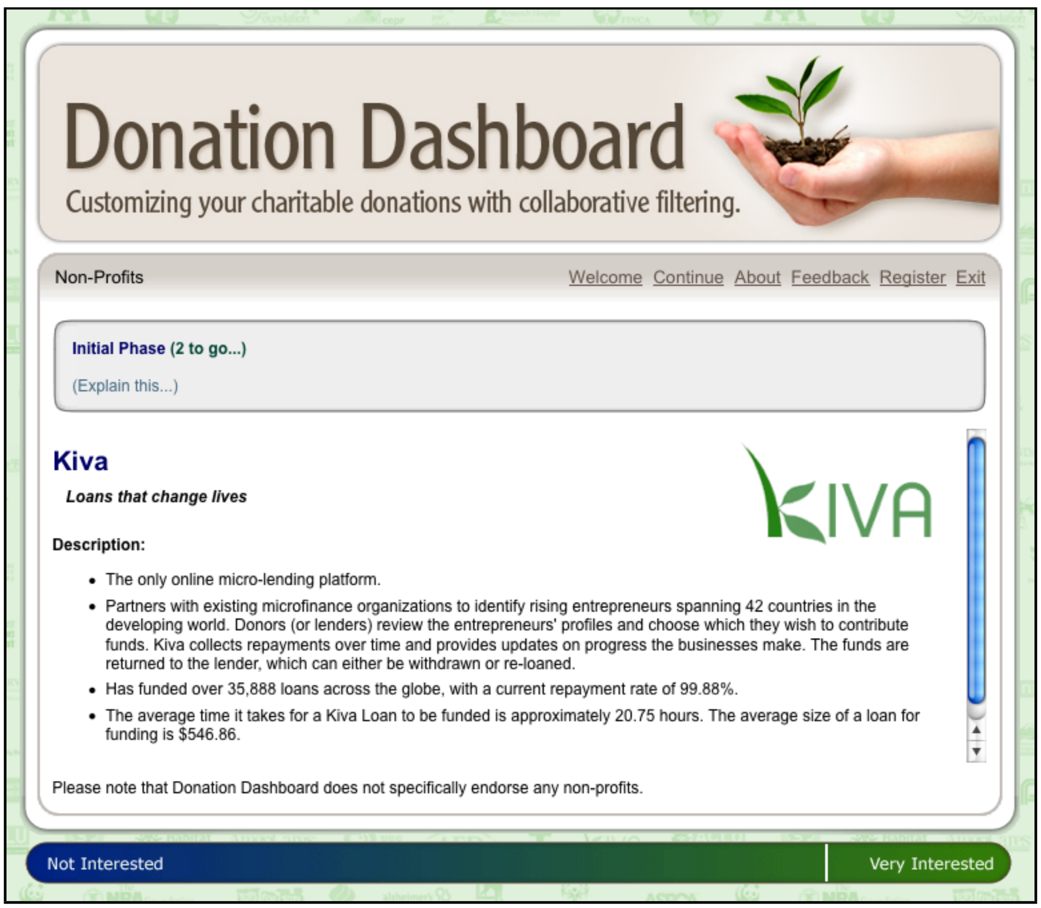


Figure 11: Screenshot of a single non-profit in Donation Dashboard 1.0, with the continuous rating bar at the bottom.

From the perspective of a new user, Donation Dashboard 1.0 works as follows: first, the user is presented with an initial set of 15 non-profit organizations, displayed one at a time. Every time he or she rates an organization the system displays the next one, until all 15 are rated. At this point the system presents the user with a portfolio of donation amounts as described below in Section 5.1.3. The user can opt to continue rating items to improve his or her portfolio.

Behind the scenes, the first group of non-profit organizations presented is a "gauge set" of items that every user is asked to rate, as described in Section 2. The next group is the "seed set," or organizations with the least amount of ratings in the system. We elaborate on the reasons for using a seed set in Section 4.2.1. The last group of organizations are items in the "recommendation set": items with the highest predicted ratings via Eigentaste. If the user opts to continue rating items to improve his or her portfolio, he is presented with items in descending order of their predicted ratings.

5.1.2 Design Details

Non-profit organizations are displayed one at a time, as a name, logo, motto, website URL and short description. Figure 11 shows a screenshot of that display.

We chose the short descriptions by manually picking out statements from non-profit information sources that we felt best described the activities of each organization. While we aimed to be unbiased, the statements we chose most likely affected the ratings collected. Nevertheless, they allowed users to easily digest the nature of each organization and quickly give an educated rating.

Below this information we display a real-valued slider in the hopes of eliciting more of a gut reaction than is possible with discrete ratings. We carefully chose the wording of the end-points of the slider to be "Not Interested" on the left and "Very Interested" on the right, instead of other possible word choices such as "Like" and "Dislike" or "Love" and "Hate." Non-profits are all generally attempting to do good in the world, and therefore it is (generally) hard to say that one "dislikes" a non-profit. It is usually a lot easier to say that one is "Not Interested" in donating to a particular non-profit.

Up until January 23, 2009 we also included an "Efficiency" percentage for each non-profit, defined as the percent of its funds spent on programs as opposed to overhead. We included this metric to give people a quick quantitative sense of each organization, but we received a lot of feedback from users who felt that it was not fair to include this one metric without including others. One user wrote, "There are so many ratings other than 'efficiency' that should be baked into a contributors decision." Instead of including an entire page of quantitative information that would have most likely resulted in information overload, we simply removed the "Efficiency" percentage from the system.

5.1.3 Generating Portfolios

We chose to present each user with a weighted item portfolio for a few reasons. First, we wanted to make the giving process more fun and interesting; both the concept of an automatically generated portfolio and the fact that a weighted portfolio can be visualized as a pie chart seemed

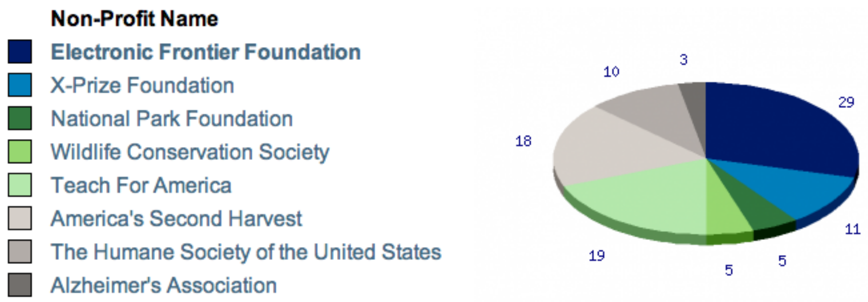


Figure 12: Screenshot of a donation portfolio in Donation Dashboard 1.0.

to accomplish that. Second, as described earlier, it is hard to say that one dislikes a non-profit; with the portfolio, we make it clear that this process is about the amount of money donated as opposed to like versus dislike. The portfolio is also very easy to share with peers; in fact, each user is given a permanent link to his personal portfolio.

The concept of a portfolio is certainly applicable to other item classes as well. Most obvious would be stock portfolios, although there are plenty of other applications: for example, a movie recommender system might present users with a portfolio of movies for their given budgets.

For Donation Dashboard, it would not be interesting to the user if we populated the portfolio solely with organizations that the he rated highly; on the other hand, populating the portfolio only with items having high predicted ratings would not take into account the actual ratings of the user. Thus, we decided to split the portfolio into two sections, section A and section B. Section A consists of items not yet rated by the user and makes up ρ_A percent of the portfolio. It is populated with items that have the highest predicted ratings, as generated by the CF algorithm (currently Eigentaste 2.0). Section B contains items already rated by the user and comprises ρ_B percent of the portfolio, where $\rho_B = 1 - \rho_A$. It is populated with items that are rated most highly by the user. Sections A and B each contain a fixed number of items for now, and items in the portfolio are weighted as follows:

We denote the rating of item i by user u with r_i^u and we denote the mean rating of item i by users in cluster c with r_i^c. Consider user u who belongs to cluster c, and let item i be the item with highest average rating in c such that $i \notin A$. Then, for each item $a \in A$, we allocate $(r_a^c - r_i^c)\rho_A$ percent of u's total donation amount to a. Similarly, let item j be user u's highest rated item such that $j \notin B$. Then, for each item $b \in B$, we allocate $(r_b^u - r_j^u)\rho_B$ percent of u's total donation amount to b. Figure 12 displays a screenshot of a sample portfolio from Donation Dashboard 1.0.

5.1.4 Populating the System

Ephrat Bitton chose the non-profits to include in version 1.0 of the system, and she looked at a few criteria. First, to make things simple, she only selected non-profits based in the United States. Further, she only included national non-profits, as we wanted to ensure that anyone using the system would be able to donate to the recommended non-profits. Note that in the future we are considering building localized instances of Donation Dashboard that recommend local non-profits. Local non-profits are often less known than larger, national non-profits, and therefore would benefit even more from Donation Dashboard.

Ephrat aimed for a sample of non-profits that would span a large range of interests. Our current set of 70 non-profits contains a number of large, well-known organizations as well as a number of smaller ones. The non-profits include the American Cancer Society, Doctors Without Borders, Kiva and Ashoka. The full list is available upon downloading our dataset, linked to in Section 5.1.

5.2 Donation Dashboard 1.0 Data Collected

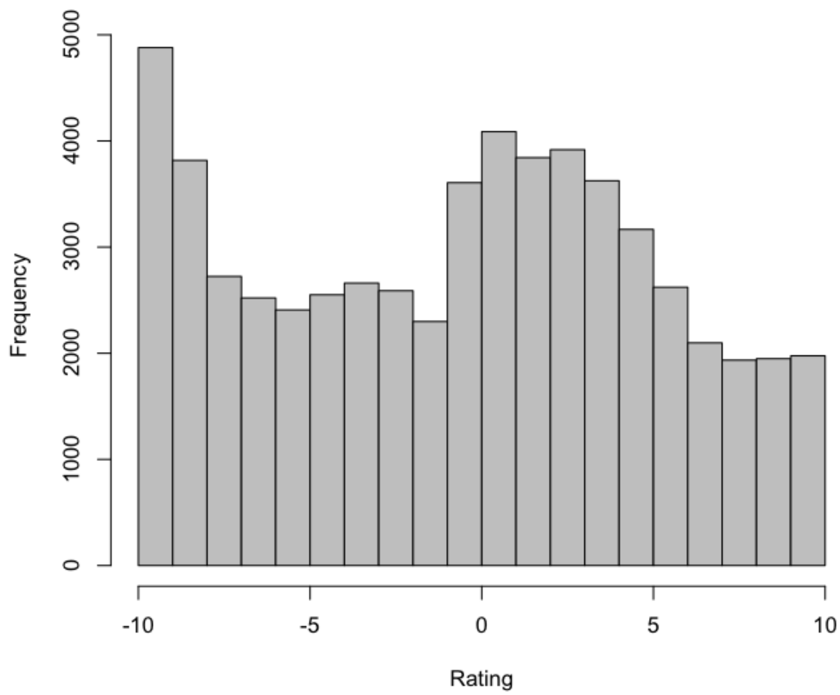


Figure 13: Histogram of Donation Dashboard 1.0 ratings (in the interval [-10.00, +10.00]).

As with Jester, we hypothesized that because non-profit organizations are generally considered in a positive light the average rating in Donation Dashboard would be positive. As it turns out, at the time of this writing the average rating is -0.65 (in the interval [-10.00, +10.00]), where the histogram of all ratings is shown in Figure 13. The histogram shows that users did indeed tend to rate organizations slightly to the right of the center, but there was also a tendency for very extreme negative ratings.

The most highly rated organizations are Doctors Without Borders, Kiva, the Public Broadcasting Service, Planned Parenthood Federation of America and Engineers Without Borders, with average ratings of 2.68, 2.29, 2.28, 1.60 and 1.46, respectively. The organizations with the lowest ratings are the NRA Foundation, the Heritage Foundation, PETA (People for the Ethical Treatment of Animals), Boy Scouts of America and Prison Fellowship, with average ratings of -6.37, -5.10, -4.90, -3.74 and -3.65, respectively. Those with the largest variance in ratings are

National Public Radio, the Humane Society of the United States, the Wikimedia Foundation, the American Society for the Prevention of Cruelty to Animals and St. Jude Children's Research Hospital. Appendix B shows the information displayed by the system for these organizations.

We use Mean Absolute Error (MAE) as described in [5] to evaluate Eigentaste 2.0 on the Donation Dashboard dataset, and we also list the Normalized Mean Absolute Error (NMAE) that normalizes MAE to the interval [0, 1]. We compare to the global mean algorithm and the results are as follows in Table 1:

Algorithm	**MAE**	**NMAE**
Global Mean	4.295	0.215
Eigentaste 2.0 (2 Clusters)	4.196	0.210
Eigentaste 2.0 (4 Clusters)	4.120	0.206
Eigentaste 2.0 (8 Clusters)	4.043	0.202
Eigentaste 2.0 (16 Clusters)	4.059	0.203
Eigentaste 2.0 (32 Clusters)	4.083	0.204
Eigentaste 2.0 (64 Clusters)	4.244	0.212

Table 1: MAE and NMAE for the global mean algorithm and Eigentaste 2.0.

The error initially decreases as the cluster count increases; however, the error increases once the cluster count reaches 16. As we collect more data and more users, the MAE for higher cluster counts should decrease. Note that users are currently divided into 8 clusters on the live system, and this data confirms that such a division is the correct choice for now.

In contrast to these NMAE values, Goldberg et. al. [16] show that the expected NMAE for randomly guessing predictions is 0.333 if the actual and predicted ratings are uniformly distributed and 0.282 if the actual and predicted ratings are normally distributed.

5.3 Donation Dashboard 1.0 Item Clusters

After we began collecting data from Donation Dashboard 1.0, we wanted to find out whether clustering items based on that ratings data would result in clusters of non-profit organizations with similar purposes. This is similar to what we measure in Section 4.4 after collecting data from Jester 4.0.

We apply standard k-means clustering to the non-profit organizations, where each organization is represented as a vector of its user ratings. We only consider users who rated *all* non-profits in order to avoid sparsity in these vectors. We run k-means with $k = 5$ and $k = 10$, and the clusters of organizations in Figure 14 are notable in terms of their relatively consistent purposes. As mentioned earlier, Appendix B shows the information displayed by the system for some of these organizations.

PETA (People for the Ethical Treatment of Animals) **American Society for the Prevention of Cruelty to Animals** **The Humane Society of the United States** **NAACP Legal Defense and Educational Fund**
One Laptop Per Child **X-Prize Foundation** **The Wikimedia Foundation** **Electronic Frontier Foundation**
Wildlife Conservation Society **Alzheimer's Association** **World Wildlife Fund** **Puppies Behind Bars** **National Park Foundation**

Figure 14: Clusters of Donation Dashboard 1.0 organizations with consistent purposes.

There are, however, many clusters that do not have consistent purposes. Figure 15 is an example:

Fisher House Foundation **Downtown Women's Center** **The Leukemia and Lymphoma Society** **United Way of America** **St. Jude Children's Research Hospital** **Center for Economic Policy Research** **William J. Clinton Foundation** **Guide Dog Foundation for the Blind** **Marine Toys for Tots Foundation** **Save the Children** **American Jewish World Service** **AVID Center** **Asha for Education** **Donors Choose American Foundation for Suicide Prevention**

Figure 15: Cluster of Donation Dashboard 1.0 organizations with inconsistent purposes.

We we would like to run this clustering again in the future after more data is collected; perhaps then there will be more clusters with consistent purposes.

To further understand this data, we also use PCA (as described in section 2) to project the non-profit organizations onto a two-dimensional plane, using the same vectors as described above. This is the inverse of the way Eigentaste 2.0 uses PCA, which is to project *users* onto a two-dimensional plane. Figure 16 illustrates the resultant projection, as well as the names of a few related organizations.

Appendix C shows a legend for Donation Dashboard 1.0 organizations, mapping ID to name.

Figure 16: PCA applied to the Donation Dashboard 1.0 organizations.

6 Eigentaste 5.0 (Algorithm)

Although Eigentaste provides a fast cold-start mechanism for new users, each user is permanently assigned to a user cluster and thus a fixed item presentation order that is determined by predicted ratings for that cluster. A disadvantage is that the system does not respond to the user's subsequent ratings. Another disadvantage is the potential for presenting similar items sequentially, referred to as the *portfolio effect* [62, 26]. For example, a Jester user who rates a Chuck Norris joke highly might then be recommended several more Chuck Norris jokes. In humor, as in many other contexts like movies or books, the marginal utility of very similar items decreases rapidly.

To address these problems we present Eigentaste 5.0, an adaptive algorithm that, in constant online time, dynamically reorders its recommendations for a user based on the user's most recent ratings. Eigentaste 5.0 also addresses the problem of integrating new items into the system. Although there are no versions 3.0 or 4.0, we adopted this version numbering to maintain consistency with our recommender system; that is, Eigentaste 5.0 was developed after Jester 4.0.

6.1 Notation

U	the set of all users
I	the set of all items
$r_{u,i}$	user u's rating of item i
$\overline{r_i}$	the mean rating of item i
$\overline{r_u}$	user u's mean item rating
$\vec{a}_u$	moving average of u's ratings per item cluster

6.2 Dynamic Recommendations

In order to maintain the constant online running time of the algorithm, we exploit the dual nature between users and items in a recommendation system. By partitioning the item set into groups of similar items, we can make recommendations based on user preferences for certain classes of items, instead of moving users to different user clusters as their preferences change. While clustering users into groups with similar taste and aggregating their ratings is helpful in predicting how a new user would rate the items, we can tailor recommendations in real-time as we learn more about the user's preferences.

We use the k-means algorithm to cluster the item space offline and across all user ratings. The Pearson correlation function is used as a distance metric between items, where the correlation between item i and item j is defined by Equation 1.

Pearson correlation provides a real-valued measure on the scale of $[-1, +1]$, where greater values correspond to higher correlation. To use this as a distance metric, we compute 1 minus the Pearson correlation, where larger values reflect greater distances. Note that it is standard procedure to only consider users who have rated both items when computing Pearson correlation, but this is problematic for k-means clustering when dealing with sparse data.

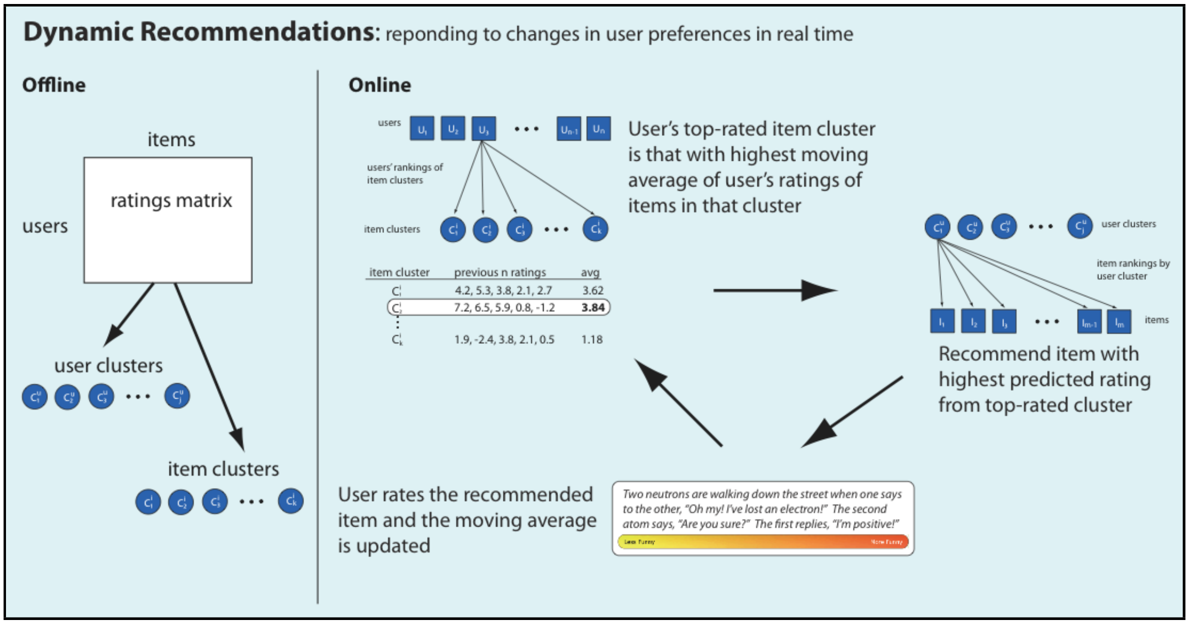


Figure 17: Illustration of dynamic recommendations in Eigentaste 5.0.

For each user $u \in U$ we maintain a vector $\vec{a}_u$ of a moving average of u's ratings of the items in each item cluster; that is, $\vec{a}_u[c]$ corresponds to user u's average rating of the last n items in item cluster c, where n is some constant. We initialize a new user's moving average slots with his or her ratings of items from the common set. For new users who do not have n ratings for each item cluster, we seed the remaining slots of their moving average with the average ratings for the corresponding item cluster across users from the same user cluster.

In Eigentaste 2.0, a user is always recommended items in decreasing order of their predicted ratings, where predictions are determined by the average rating for that item by users in the same user cluster. In essence, a user's ratings of additional items have no influence on which item is recommended next.

We give a constant online time solution as follows: user u is recommended the top-predicted item (not yet rated by u) from the item cluster corresponding to the highest value in $\vec{a}_u$. As u tires of the items from that cluster, the moving average of his or her ratings for the cluster will begin to decrease and eventually fall below that of another item cluster.

For clarity, we provide a numerical example that walks through the process of recommending the next item to some user u. Suppose u maintains the following table, $\vec{a}_u$ (Table 2), of his or her last 5 ratings of items in each item cluster.

Item Cluster	**User u's last 5 ratings**	**Average**
1	4.2, 5.3, 3.8, 2.1, 2.7	3.62
2	7.2, 6.5, 5.9, 0.8, -1.2	**3.84**
3	1.9, -2.4, 3.8, 2.1, 0.5	1.18

Table 2: Sample ratings table.

At present, user u's moving average of items in cluster 2 is the highest, and so Eigentaste 5.0 presents u with the item from cluster 2 that has the highest predicted rating and has not yet been evaluated by u. Suppose u rates this item with -2.3. The user's new moving average of item ratings for cluster 2 becomes 1.94, which is lower than the moving average for item cluster 1. Subsequently, in the next iteration, user u will be recommended the item from cluster 1 that has the highest predicted rating, and the process is repeated.

6.3 Cold Starting New Items

The cold start problem for new users, as described in Section 2, is the problem where a lack of information about new users results in an inability to make good recommendations. Similarly, good recommendations cannot be made when there is a lack of information about new *items*: that is the cold start problem for new items.

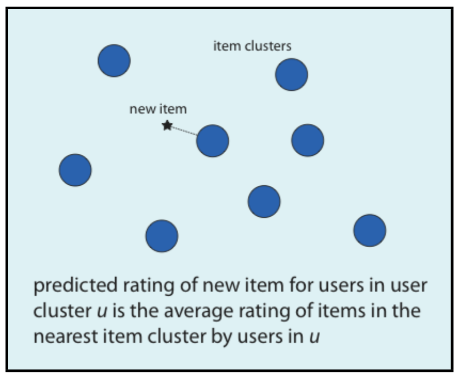


Figure 18: Illustration of cold starting in Eigentaste 5.0.

Difficulty in introducing new items to the system stems from the fact that they have so few ratings overall, and even fewer ratings within individual user clusters. Hence, the predictions generated by Eigentaste 2.0 are subject to more variability due to the small sample sizes.

Eigentaste 5.0 uses sparse ratings for a new item i to find the closest item cluster based on the Pearson distance metric described in section 6.2. User u's predicted rating of i is determined by the average rating of all items within i's nearest item cluster across users in u's user cluster. We use confidence intervals to determine the appropriate time to switch from this estimate to the estimate based only on actual ratings of item i.

6.4 Results

It is impossible to truly compare how a user would react to different item recommendation orders, as the first test would greatly bias the second. We first compare the algorithm with its predecessor by backtesting on old data collected from Jester, and then we evaluate the algorithms more conventionally by randomly assigning new online users to either Eigentaste 2.0 or 5.0.

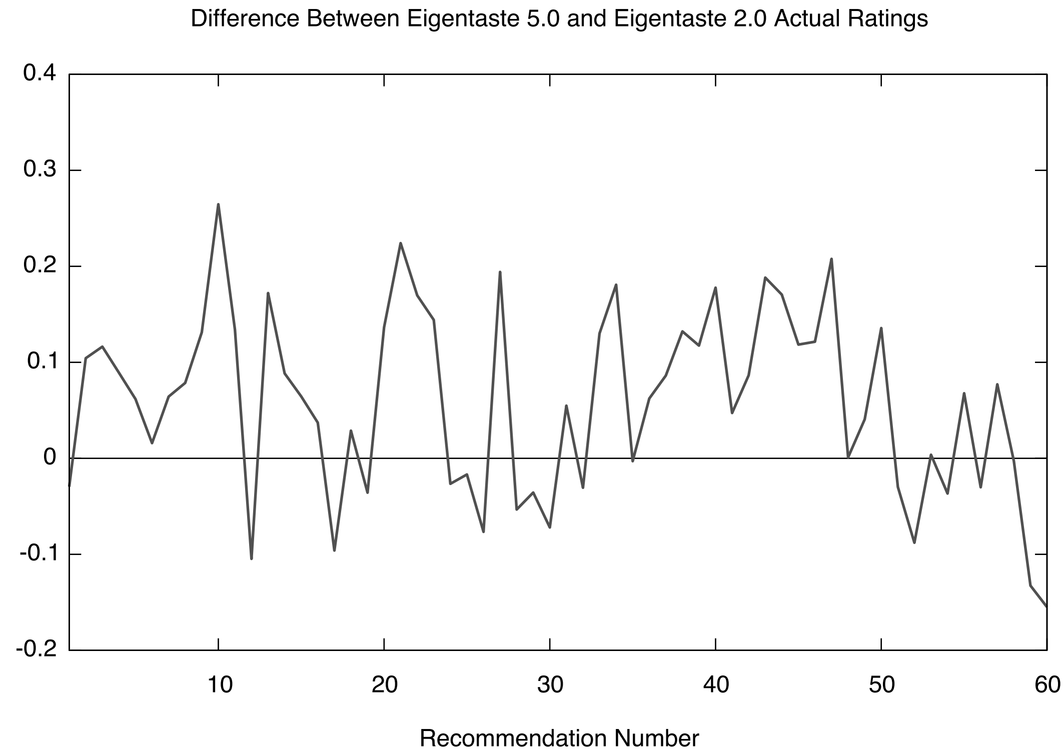


Figure 19: Average difference (across 7,000 users) between actual ratings for Eigentaste 5.0 and 2.0 for the i^{th} recommended item.

6.4.1 Backtesting Results

We simulated Eigentaste 2.0 and the dynamic recommendations of Eigentaste 5.0 with the Jester data collected between 1999 and 2003. The users are randomly partitioned into equal sized sets of "existing" and "new" users, and "existing" users are clustered using principal component analysis. The item space is clustered using a k-value of 15, and we use a 5 item moving average for each item cluster to track user preferences. We iteratively introduce the "new" users into the system and determine the order that the respective algorithms would recommend items. The two sequences of ratings (and corresponding predictions) for each user are recorded in this order. We average the ratings for the i^{th} recommended item across all "new" users for both actual and predicted rating sequences.

Figure 19 shows the difference between the average actual ratings for Eigentaste 5.0 and 2.0. In accordance with our objective, we find that Eigentaste 5.0 provides a distinct advantage over Eigentaste 2.0 for earlier recommendations, particularly within the range of the first 50 items.

The differences between the average actual ratings and the average predicted ratings for each algorithm are shown in Figure 20, which illustrates that the error in predictions is significantly greater for Eigentaste 5.0 than for Eigentaste 2.0. This is because the user clusters used to generate predictions only consider ratings for the common set of items, while with Eigentaste 5.0,

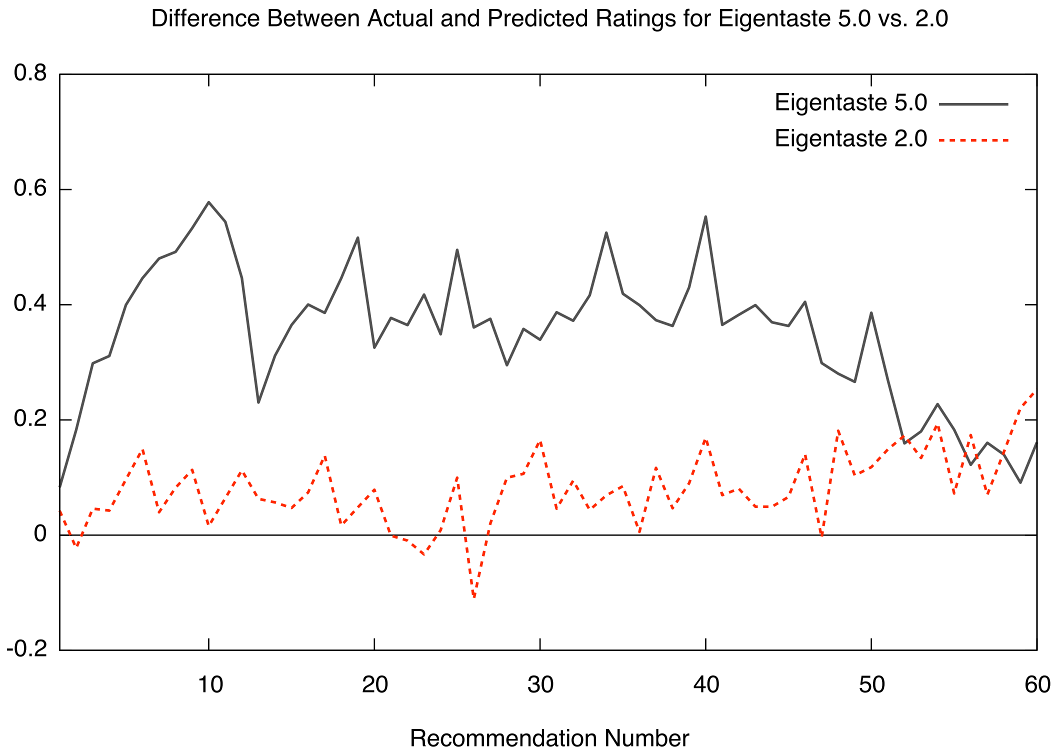


Figure 20: Average differences (across 7,000 users) between actual and predicted ratings for Eigentaste 5.0 and 2.0 for the i^{th} recommended item.

we can recommend items better suited to a user's interests by using item clusters to take into account the latest user-specific information.

6.4.2 Recent Results: Jester with Eigentaste 5.0

We incorporated the dynamic recommendations of Eigentaste 5.0 into Jester 4.0 by randomly partitioning new users into two groups: one group was recommended jokes using the Eigentaste 2.0 algorithm, while the other group was recommended jokes using the Eigentaste 5.0 algorithm. Note that we disabled seeding in Jester 4.0 at that point, as it interfered with our ability to compare Eigentaste 5.0 and 2.0 fairly.

Between September 22 and October 14, 2007 we collected 22,000 ratings from 912 new users, some who used Eigentaste 2.0 and some Eigentaste 5.0. The results of comparing the two groups of users are as follows in Figure 21:

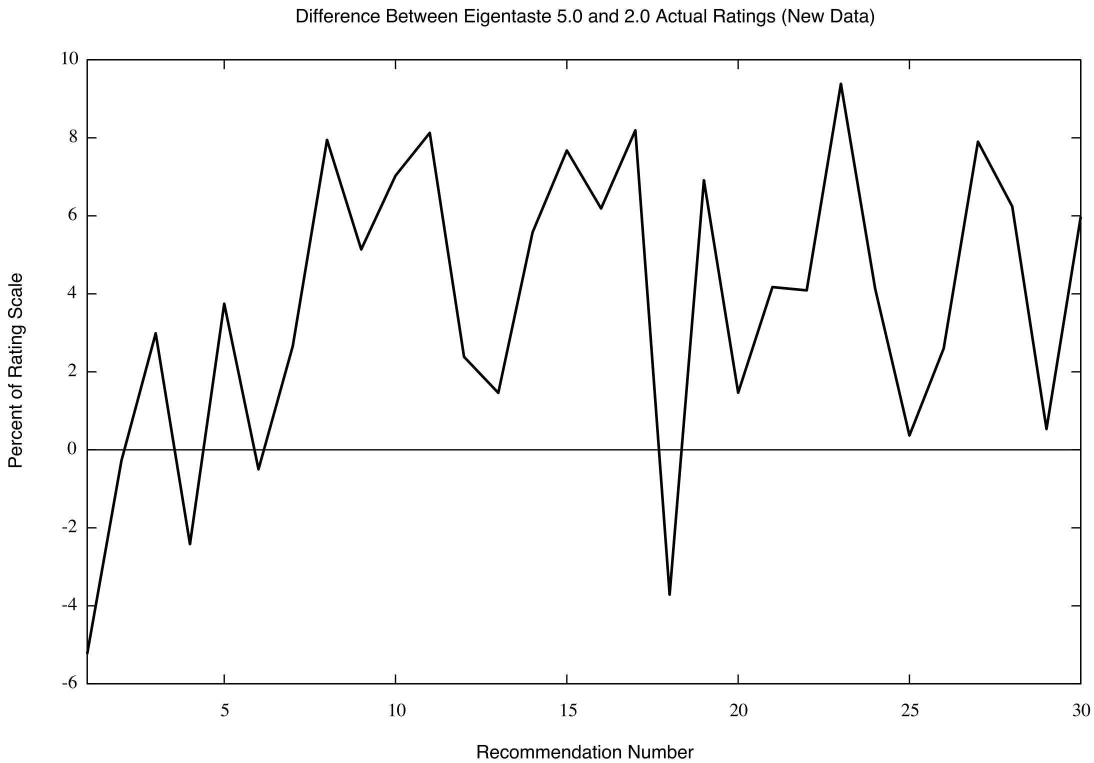


Figure 21: Average difference (across 912 users) between actual ratings for Eigentaste 5.0 and 2.0 for the i^{th} recommended item.

The differences in the rating scale for these results are far greater than those found with backtesting, as can be seen by comparing Figure 21 with Figure 19. The histogram shows that as the item recommendation number increases beyond 5, Eigentaste 5.0 generally results in significantly higher ratings for the next 25 items than Eigentaste 2.0. This makes sense, as Eigentaste 5.0 does not have much information to work with prior to the 6th item presented.

7 Recommending Weighted Item Portfolios Using Relative Ratings (Model)

Recommender systems that collect explicit ratings, including the systems discussed in the previous sections, typically prompt users to rate items on an absolute scale rather than relative to each other. When providing absolute ratings, users are implicitly comparing the item they are rating with every other item in the universal set. As a result, they cannot provide accurate ratings if they are not aware of all the items in the set. Furthermore, it is often difficult for users to remember which items they were presented and how they rated them, which can lead to unintentional inconsistencies.

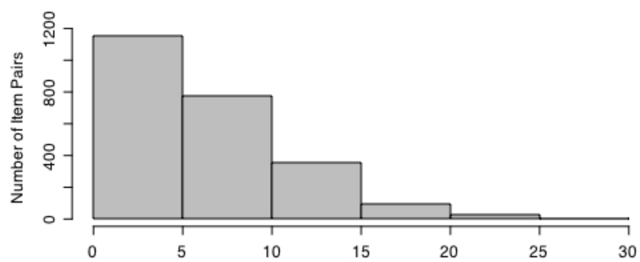


Figure 22: Difference between the average rating of an item pair when the presentation order is reversed, as a percent of the rating scale.

Based on data collected with Donation Dashboard 1.0, we found that the order in which items are presented can significantly bias the user's ratings. Let $\bar{\delta}_{ij}$ be the average difference of user ratings for item i and item j when i is shown before j. We measured $\Delta_{ij} = |\bar{\delta}_{ij} - \bar{\delta}_{ij}|$ for every item pair (i, j), and plot the number of item pairs that fall into different ranges of Δ in Figure 22. Ratings are normalized to a $[0, 1]$ scale. We observe that the Δ values of more than half of all item pairs exceed 5% on the rating scale, and more than 20% of the pairs exceed 10%.

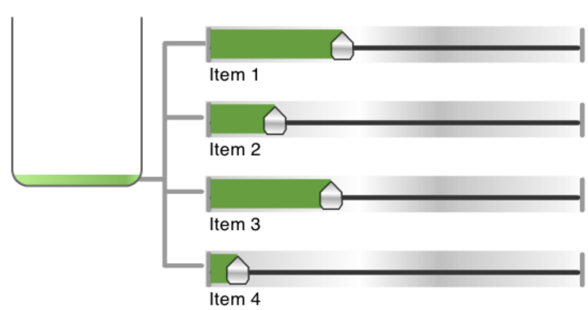


Figure 23: A potential user interface for relative ratings.

We hypothesize that these findings result from the implicit comparison with the universal set inherent in absolute ratings. Motivated by the observed bias, we present a graph-based model and algorithm for collaborative filtering that ask users for *relative* ratings as opposed to absolute ratings. We expect that using relative ratings will reduce the bias demonstrated by our data, because they eliminate the need for users to compare each item against the universal set. A potential user interface for relative ratings is shown in Figure 23.

In this new framework, we maintain the structure of Eigentaste 2.0, but the user provides relative ratings for portfolios of items instead of absolute ratings for individual items. The user is first asked to rate a gauge portfolio, which consists of the items whose relative ratings have the highest combined variance. He or she is then placed in a cluster generated by PCA and recursive rectangular clustering, and is asked to rate a series of item portfolios about which his cluster has little or no information. The final portfolio of recommendations is based on the user's relative ratings and those of the users in his or her cluster.

Our objectives are to determine the intermediate portfolios that maximize our confidence in the user's preferences and to determine a portfolio of recommendations that consistently aggregates the user's ratings with those of his or her cluster.

7.1 Notation

I	item set
U	set of users
C	set of user clusters
n	the fixed number of items in a portfolio
$r^u_{i,j}$	user u's preference for item i over item j
$r^c_{i,j}$	cluster c's aggregate preference for item i over j
$p^u_{i,j}$	prediction of user u's preference for item i over j
$\epsilon^u_{i,j}$	prediction error of u's preference for item i over j

7.2 Relative Ratings

As an alternative to absolute ratings, we propose asking users to specify their relative preference for one item with respect to a set of other items. To do so, users provide ratings in the form of a multi-way comparison, as defined below.

Definition 1 *A* ***multi-way*** *comparison of a set of items is specified as a distribution of preference among those items. More practically, the user must indicate the degree to which each item is preferred to the other items, using a continuous scale and normalized to a percentage. If user u allocates r^u_i percent to item i and r^u_j percent to item j, then we define the intensity of u's preference for i over j as $r^u_{i,j} = r^u_i - r^u_j$. Hence, we say that user u prefers i to j if $r^u_{i,j} > 0$. A set of multi-way comparison ratings is* ***consistent*** *if for each i, j, k we have that $r^u_{ik} = r^u_{ij} + r^u_{jk}$.*

We consider the graphical representation (Figure 24) of the active user's relative ratings by the directed graph $G(V, E_u)$, where the set of vertices V is equal to the set of items I. Add a directed edge (i, j) with weight $r^u_{i,j}$ to E_u if u prefers item i to item j.

Hochbaum and Levin prove in [22] that a necessary condition for consistency is that the graph be acyclic. If a user's ratings are consistent, then a topological sort of the vertices will yield a consistent preference ranking of the items. They also prove that if the ratings are consistent, then all paths between any two nodes will be of equal length. Hence, given an incomplete but

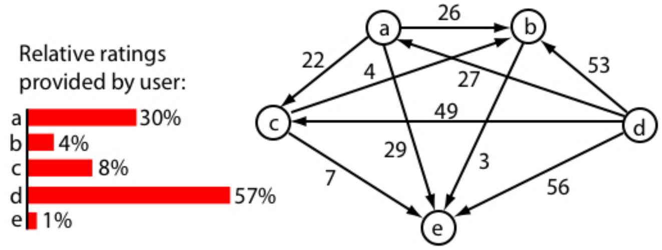


Figure 24: Normalized histogram of a sample multi-way rating and its corresponding graphical representation.

consistent matrix, we can construct the *consistent closure* of the user's ratings by setting r^u_{ij} to be the length of a path from i to j, should one exist. Otherwise, i and j are said to be *incomparable*. We assume for now that users always give consistent ratings and discuss later how to handle inconsistencies.

It is important to note that users are clustered based on their multi-way or *relative* ratings of a set of items and not based on how much they like each *individual* item. Hence, a user who likes every item in the gauge portfolio an equal amount will be placed in the same cluster as a user who dislikes every item equally.

7.3 Prediction Error Model

According to our model, if a user u always gives consistent preference ratings, then $r^u_{i,k} = r^u_{i,j} + r^u_{j,k}$ for all $(i, j, k) \in I$. Hence, if we can predict u's preferences for i over j and for j over k, then we can predict u's preference for i over k by invoking the same transitive principle. However, there will naturally be a degree of error with every prediction; furthermore, when the path that we are predicting over is longer, we should expect a greater level of error.

If we make the assumption that the relative ratings $r^u_{i,j}$ of users in the same cluster can be modeled as independent and identically distributed random variables, then according to the law of large numbers (LLN) the average of the ratings will approach their expected value [42].

Let $\epsilon^u_{i,j} = p^u_{i,j} - r^u_{i,j}$ be a random variable representing the prediction error of u's preference for item i over j. If $p^u_{i,j}$ is equal to the average preference for i over j in u's user cluster c, then $\epsilon^u_{i,j}$ will have an expected value of 0 and some corresponding variance. When a prediction is made by finding the consistent closure of a pair of connected edges, the variance of the error is dependent on the covariance of the two error terms.

7.4 Determining the Next Portfolio

We treat the variance of a cluster's prediction error for a pair of nodes as the degree of confidence we possess about users in that cluster. That is, when variance is low our prediction error should be much lower than when variance is high. To model our confidence of the active user u, we can construct an undirected graph with the weight of each edge corresponding to the variance of our prediction error for that edge. For every relative rating provided by u, the variance and thus the weight on the corresponding edge is equal to 0, since the value is already known. For simplicity, we assume that users are consistent in their ratings, and so we assign a weight of 0 for any edge in the consistent closure of u's ratings. Any edge that is not in the consistent closure is assigned a weight corresponding to the variance of prediction error for that edge across users in u's cluster. If there exists a path of lower total variance between the same two nodes, we replace the original weight with the lower one. We may wish to model user inconsistency in future work.

Now that we have a model reflecting our knowledge of a user's preferences, *our goal when collecting ratings is to reduce the variance of our prediction error.* Mathematically, we seek to determine the set of n items that when rated will minimize the total variance of prediction error for the remaining item pairs. For now, we use a simple heuristic and select the set of n nodes whose interconnected edges have the greatest combined variance.

7.5 Final Portfolio Selection

To determine user u's final recommended portfolio, we wish to rank the items in I according to u's preferences and those of the users in his cluster c. We require any final ranking of the items to hold two properties: 1) u's preferences must be preserved, and 2) the aggregated preferences must be consistent. We first describe how to optimally and consistently aggregate a set of relative ratings and then show how we can preserve the ratings of a single user when doing so.

Under our model, the individual preference ratings of users in cluster c can be combined by superimposing their graphical representations, allowing the possibility for there to be more than one edge between pairs of nodes. In most practical applications, the aggregate graph $G\left(V, E_c = \cup_{u \in c}\{E_u\}\right)$ of user ratings will yield inconsistencies. The problem is then to find a consistent relative rating of the items in the universal set that is close to the individual ratings of the users.

Let the decision variables $r^c_{i,j}$ be the preference for i over j aggregated across all users in cluster c, and let r^c_i be the weighting given to item i such that $r^c_{i,j} = r^c_i - r^c_j$. Given a set of preference ratings, a unique set of weights can be found by normalizing r^c_1 to 0. Then as shown in [22], this problem can be formulated by the following unconstrained convex optimization problem:

$$\begin{aligned} \min \quad & \sum_{v \in U} \sum_{i>j} \left(r^c_i - r^c_j - r^v_{i,j}\right)^2 \\ \text{(subject to} \quad & r^c_1 = 0) \end{aligned} \tag{6}$$

Using either Newton's method or the conjugate gradient method, this problem can be solved in $O\left(n^3\right)$, where n is the number of items in I. [22] To preserve the ratings provided by user u, we reconstruct the optimization problem by replacing $r^c_{i,j}$ with $r^u_{i,j}$ for all (i, j) item pairs u has rated. Once we have a consistent aggregate ratings graph that preserves u's preferences, we can perform a topological sort on the graph to order the items according to highest preference. This information can be used to select items for the final portfolio and allocate resources appropriately, similar to the method described in Section 5.1.3.

8 Eigentaste Security Framework (Model)

8.1 Attacks on Collaborative Filtering

Malicious users can alter collaborative filtering recommendations by rating items in a certain manner. Such attacks can have devastating consequences; for example, an attack might prevent an item from being recommended on `Amazon.com`, which would drastically affect its sales. Similarly, an attack might cause Donation Dashboard to assign more funds than it should to a particular non-profit.

Mobasher et. al. [33] present a formal framework for modeling attacks on general collaborative filtering systems. It includes various attack models and two attack objectives: making an item more likely to be recommended ("push") and making an item less likely to be recommended ("nuke"). We adapt this framework to the Eigentaste algorithm, as described in the sections below.

8.2 Attack Scenario

We consider an attacker attempting to influence recommendations given by Eigentaste to new users. Since the Jester dataset is available online for public access, we assume the attacker has prior knowledge of each and every rating, as well as the ability to calculate the individual item ratings distributions and the global ratings distribution. The algorithm is published, so the attacker also has the ability to recreate the system and calculate the existing user clusters.

The scenario we propose is as follows: an attacker creates malicious profiles ("profile injection") to target an item for either "push" or "nuke," causing Eigentaste to rank that item higher or lower in the recommendation set than it otherwise would. For example, Chuck Norris may wish to see his jokes recommended ahead of blonde jokes. He has two options available to him: push Chuck Norris jokes to increase their ranks or nuke blonde jokes to decrease their ranks. Similarly, in the case of Donation Dashboard, a non-profit organization might want to increase its own rank or decrease the ranks of its competitors.

We do not consider attacks on target items that appear in the gauge set; in the case of Jester, gauge set items are never recommended, so "push" or "nuke" attacks on such items would be undefined.

As previously mentioned, Mobasher et. al. introduce an attack framework for collaborative filtering in general. Their user profiles comprise a target item t (rated either r_{max} or r_{min}), a set of unrated items, a set of randomly rated "filler" items and a "selected" set items rated in a particular way. In the case of user-based collaborative filtering, attack profiles need to be "close" to other profiles to affect their predictions: "filler" items increase the number of rated items so that the attack profile is close to other profiles in general, while "selected" items position the attack profile close to particular profiles. There are similar reasons for the existence of these sets in the case of item-based collaborative filtering.

In the case of a general collaborative filtering system, any item can be rated at any time, and there is no clean split between any of the sets described above. However, Eigentaste requires that all users rate the gauge set in order to be matched with a cluster; furthermore, items are recommended one at a time after a user is matched with a cluster. Thus, a target item cannot be rated until it appears as a recommendation. Taking all of this into account, the general framework maps onto Eigentaste as follows: Eigentaste's gauge set serves the purpose of the "selected" set of items by positioning the user close to a specific group of others users (the cluster). Furthermore, the ratings of recommended items prior to reaching the target item are similar to "filler" items. Finally, items that would be recommended after the target item are unrated.

Note that Eigentaste does introduce a clean split between the gauge set items used for determining closeness with other users and the other items used for getting to the target item to push or nuke. As a result, we propose that a user profile consist of a Cluster Entry Vector (CEV) and a Recommendation Vector (RV), where the CEV is the user's gauge set ratings and the RV is the user's recommendation set ratings (ratings of recommended items, including the target item).

We describe each attack documented by Mobasher et. al. as well as define its Eigentaste variants (both CEV and RV) if applicable.

8.2.1 Segment Attack

A Segment attack allows an attacker to push or nuke items to a particular group (segment) of users. When pushing, the attacker creates a malicious profile that rates a set of well liked items in the targeted segment with r_{max}, some filler items are rated r_{min}, other items are unrated, and the target item is rated r_{max}. The nuke attack can be performed similarly by rating the target item with r_{min}.

Eigentaste Variants:

The differentiating element of the Segment attack compared to other attacks is that it focuses on a particular group of users; thus, we focus on the CEV attack. The r_{max} rating of the target item and the r_{min} ratings of the "filler" items are not unique to the Segment attack, so we introduce a corresponding RV attack later on when such ratings are brought up again (see Section 8.2.4).

There are two possible cases for a Segment attack in Eigentaste: a segment can be within (or equal to) a user cluster or it can span multiple clusters. In the former case, any rating of the gauge set items is a Segment attack; in the latter case, such an attack is not possible as there is no closeness metric in Eigentaste other than membership in a cluster.

We choose to focus our efforts on the Random and Average attacks because their CEV variants degenerate into specific Segment attacks.

8.2.2 Random Attack

In a Random attack, the attacker creates a profile with some "filler" ratings drawn from a normal distribution $N(\mu_{global}, \sigma^2_{global})$ with mean and standard deviation calculated from all ratings in the system, where we re-draw from the normal distribution if the rating we draw is greater than r_{max} or less than r_{min}. A target item to be pushed is rated r_{max} and a target item to be nuked is rated r_{min}. Finally, some items remain unrated.

Eigentaste Variants:

A Random CEV attack draws the ratings of all gauge set items from N(μ_{global}, σ^2_{global}). A Random RV attack draws the ratings of recommendation set items from N(μ_{global}, σ^2_{global}) until the target item is recommended, at which point it is rated r_{max} or r_{min} (push or nuke). No other items are rated.

8.2.3 Average Attack

An Average attack proceeds similarly to the Random attack except that, for each item assigned a rating, it samples from that item's rating distribution N(μ_{item}, σ^2_{item}) rather than the global ratings distribution.

Eigentaste Variants:

The Eigentaste variants for both the CEV and RV are equivalent to those for the Random attack, with the per-item ratings distribution used in place of the global ratings distribution.

8.2.4 Love/Hate Attack

The Love/Hate attack is an intuitive attack on collaborative filtering systems. For push, some "filler" items are rated with r_{min}, the target item is rated with r_{max} and some items are left unrated. For nuke, r_{min} and r_{max} are flipped. The target item is either promoted or demoted and the rest of the items experience a prediction shift in the opposite direction of the target item.

Eigentaste Variants:

The corresponding CEV attack would degenerate into a Segment attack targeting the cluster that hates everything or the cluster that loves everything, but such an attack does not seem particularly useful or noteworthy. More importantly, the Love/Hate attack aims to create as large a differential as possible between the target item and other items that might be recommended, which maps perfectly onto an RV attack: for push, all items are given a rating of r_{min} until the the target item is recommended, at which point it is rated r_{max}. No other items are rated, and r_{min} and r_{max} are flipped for nuke.

8.2.5 Bandwagon Attack

The Bandwagon attack aims to create a profile that is "close" to as many other profiles as possible, and it achieves this by rating popular items highly: i.e. items that have been rated highly by a large number of users. This allows an attacker to influence a large number of other profiles upon rating a target item. More specifically, a "selected" set of popular items is rated r_{max}, some other "filler" items are given random ratings, other items are unrated and the target item is rated r_{max}.

The nuke variant of this attack is called the Reverse Bandwagon attack, which works similarly to the Bandwagon attack. The attacker crafts a malicious profile with low ratings of a set of unpopular items, items which have been rated poorly by a large number of users. The target item is given a low rating to associate it with the unpopular items. Specifically, a set of unpopular items is rated r_{min}, the remaining items are rated as before (random and unrated) and the target item is rated r_{min}.

Eigentaste Variants:

The RV component of the Bandwagon attack does not need to be discussed, as it would consist of random ratings and a target item rating of r_{min} or r_{max}, which is equivalent to the Random RV attack (see Section 8.2.2).

The CEV component would be more noteworthy; however, Eigentaste's median-based clustering ensures that clusters have an (almost) even number of users, so entering one cluster would not affect more users than entering another. Furthermore, the gauge set items are chosen to have high variance; as a result, none of them are particularly popular or unpopular. Thus, the Bandwagon attack is not applicable to Eigentaste at all.

8.2.6 Eigentaste Attack Summary

The following tables summarize the attacks described above.

Summary of attacks on cluster entry vectors:

Attack Type	Gauge Set Ratings
Random	r $\sim$ N(μ_{global}, σ^2_{global})
Average	r $\sim$ N(μ_{item}, σ^2_{item})

Summary of attacks on recommendation vectors:

Attack Type	Attack Objective	Recommendation Set Ratings	Target Item Rating
Love/Hate	Push	r_{min}	r_{max}
Random	Push	r $\sim$ N(μ_{global}, σ^2_{global})	r_{max}
Average	Push	r $\sim$ N(μ_{item}, σ^2_{item})	r_{max}
Love/Hate	Nuke	r_{max}	r_{min}
Random	Nuke	r $\sim$ N(μ_{global}, σ^2_{global})	r_{min}
Average	Nuke	r $\sim$ N(μ_{item}, σ^2_{item})	r_{min}

9 Opinion Space (System)

Opinion Space is an experimental new system for visualizing opinions and exchanging ideas. Users can express their opinions and visualize where they stand relative to a diversity of other viewpoints. Designs for early versions of Opinion Space began in Fall 2008, and development began in December 2008.

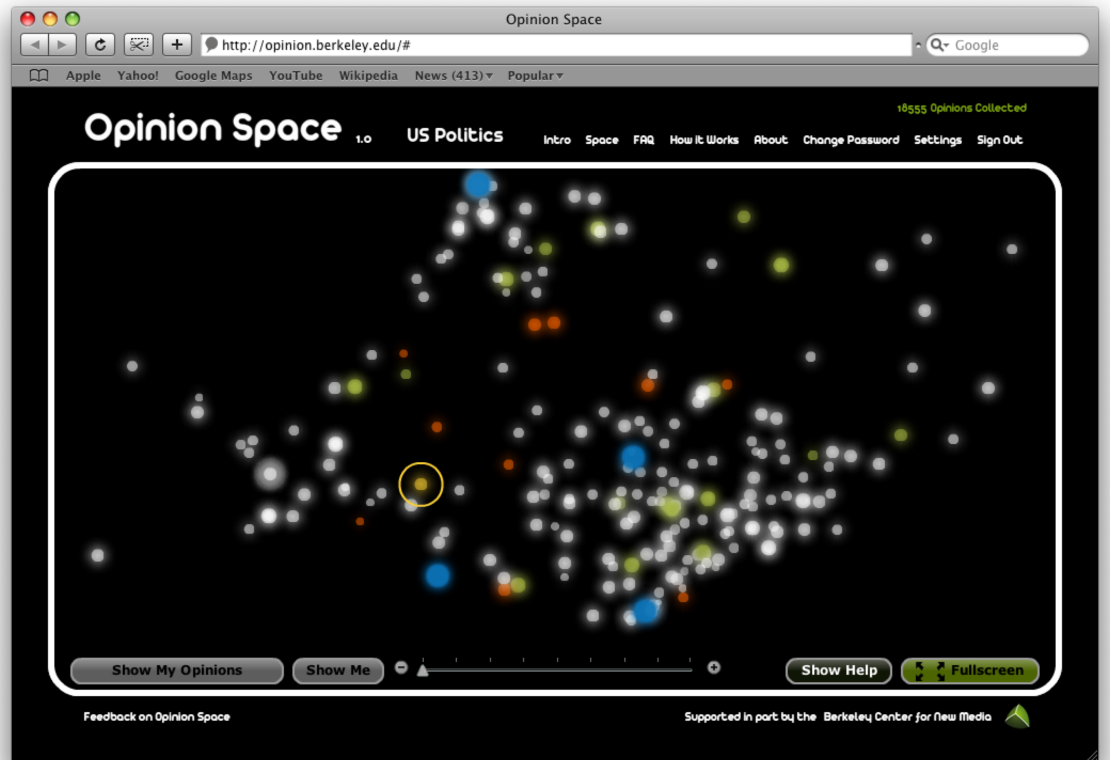


Figure 25: Screenshot of Opinion Space 1.0, accessible at `http://opinion.berkeley.edu`**.**

9.1 Early Designs

Opinion Space went through several iterations in Fall 2008 before we arrived at the visualization system that it is today. These early versions are described below:

9.1.1 News Filter

In the first design of Opinion Space the focus was on filtering news. The intuition was that if reading a news article resulted in a change of opinion for one user, that news article would likely change the views of similar users.

It would have worked as follows: a user would input his or her opinions on specific issues, at which point the system would present relevant news articles to the user. The user would then be able to change his or her original responses after the presentation of every news article. Figure 37 in Appendix E shows a mockup of this design.

9.1.2 "Devil's Advocate" System

The next design was a "devil's advocate" system, where the key question for users was "How much you can influence others' opinions?" It was a system where users could submit comments that they felt would affect others' opinions on various statements, and that would recommend comments to users based on what comments successfully changed similar users' opinions. In other words, it was a comment recommender system, and in some senses an automated version of systems like KQED's "You Decide" project.

Specifically, users would be able to submit statements that could be either agreed upon or disagreed upon, and would also be able to submit comments that they felt would affect others' views on any of the user submitted statements. A user would click on a statement and be able to view the comments that were most successful in changing the opinions of similar users.

The system would highlight unbiased comments by rewarding those that were able to affect the opinions of other users, and would therefore be a resource in finding unbiased opinions. For example, consider the following question: "Is torture justifiable if it prevents a terrorist attack?" Typical responses to this question would tend to be biased either pro-torture or anti-torture, which makes them untrustworthy. However, when comments are rewarded for *changing* opinions, a very successful pro-torture response would be one that changed the opinion of someone who was anti-torture, or vice versa. Such unbiased comments would be helpful in exploring the issues.

Figure 38 in Appendix E shows a mockup of this design.

9.1.3 Eigentaste-Based Visualization

The final design of Opinion Space that became Opinion Space 1.0 builds on Eigentaste by using a key component: PCA (see Section 2). Users rate their opinions on five "propositions" or statements, which places them as points in a five-dimensional space. The propositions are essentially the Eigentaste "gauge set."

PCA is applied to the five-dimensional ratings data of all the users, who are then projected onto the resultant two-dimensional space. We then allow users to *visualize* this space, which is the core of Opinion Space 1.0.

9.2 Opinion Space 1.0 Description

Opinion Space 1.0 launched to the public on April 22, 2009. As mentioned in Section 1, it has collected over 18,000 opinions from over 3,700 users as of May 2009. It has been featured by Wired, the San Francisco Chronicle, and other news sources. Figure 25 shows a screenshot of

Opinion Space 1.0, and Figure 39 in Appendix E shows screenshots of Opinion Space during its development. Appendix D lists selected responses to the discussion question described below.

The system uses Flex 3 as its frontend and Django 1.0 as its backend, as well as additional libraries such as Flare and Degrafa for visualization. Like Jester 4.0 and Donation Dashboard 1.0, it accesses a MySQL database.

Opinion Space 1.0 is online at `http://opinion.berkeley.edu`. Below we describe it in detail.

9.2.1 System Usage Summary

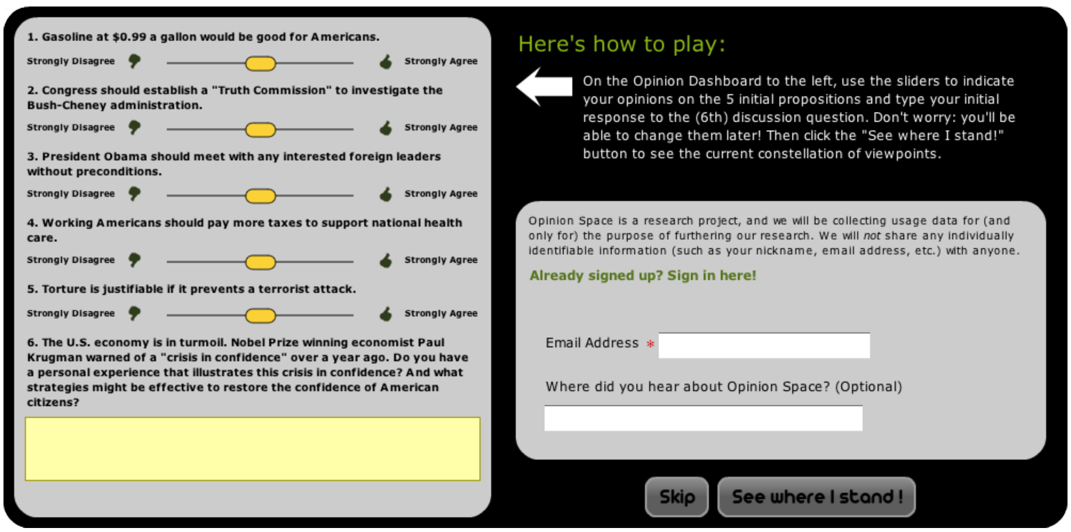


Figure 26: Screenshot of the initial propositions page in Opinion Space 1.0.

Upon entering Opinion Space, the active user is shown five "propositions" (statements), each with a corresponding slider bar that ranges from "Strongly Disagree" to"Strongly Agree." The user is also shown a discussion question below the five propositions with a corresponding text field, as shown in Figure 26. After adjusting the five sliders, optionally responding to the discussion question and entering an email address, the user can click "See where I stand!"

At this point a two-dimensional map or "space" is displayed, as shown in Figure 25. Both the active user and other users are represented as points in this space based on their ratings of the five propositions, where the transformation from proposition ratings to a point in the space is via the application of PCA (as described in Section 2). The space also contains "landmark" users (public figures), whose ratings (and therefore, positions in the space) correspond to educated guesses about how the corresponding public figure would have rated the propositions. Closer points in the space indicate higher levels of agreement on the five propositions; thus, the active user is able to visualize his or her opinions on the five propositions with respect to others.

There is a button on the map called "Show My Opinions" that, when clicked, pops up "My Opinion Dashboard," as shown in Figure 27. This dashboard allows the active user to adjust his or her ratings of the five propositions and corresponding position in space at any point, along

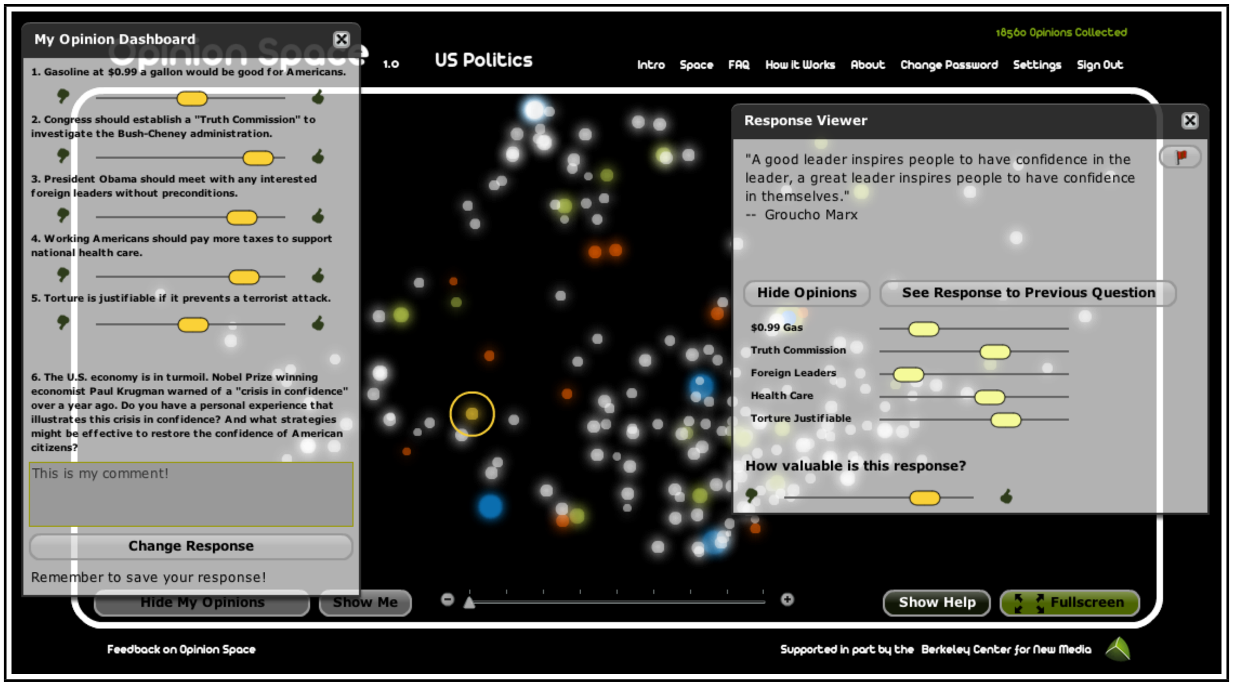
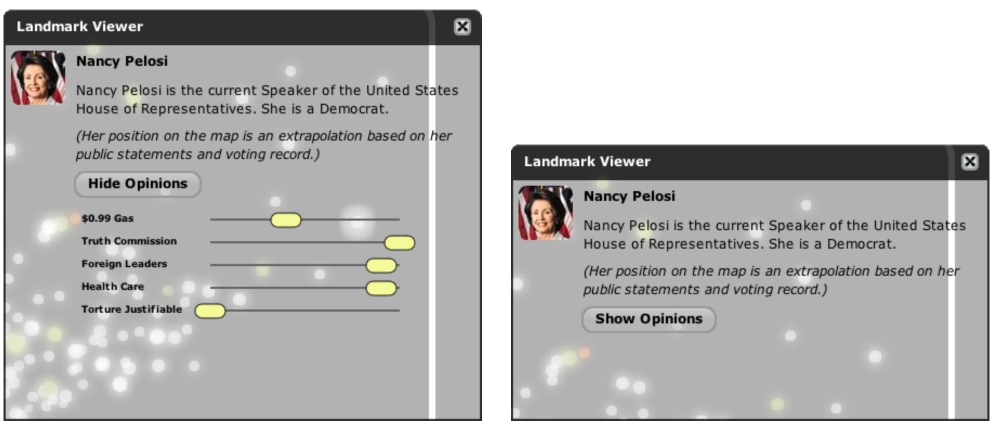


Figure 27: Screenshot of "My Opinion Dashboard" and the "Response Viewer" in Opinion Space 1.0.



(a) Opinions shown (b) Opinions hidden

Figure 28: Screenshot of the "Landmark Viewer" in Opinion Space 1.0.

with his or her response to the discussion question. The user can also click on any of the points in the space: clicking on a normal user's point pops up a "Response Viewer" window (as shown in Figure 27) that allows him to view both that user's response to the discussion question and ratings of the five propositions. Similarly, clicking on a landmark user's point pops up a "Landmark Viewer" window (Figure 28) that identifies the landmark and lists the landmark's ratings. Ratings are hidden by default in both the "Response Viewer" and the "Landmark Viewer"; users must click "Show Opinions" to make them visible.

After viewing a user's response to the discussion question in the "Response Viewer," the active user can rate that response with another slider. The rating will affect the rated user's "score," and therefore the size of the rated user's point, as explained in Section 9.2.3. The rating will also affect the *color* of the rated user's point as displayed to the active user, which we explain in Section 9.2.2. The active user can also click the red flag at the corner of the "Response Viewer" in order to report any response as inappropriate, and such reports are currently monitored manually.

9.2.2 Design Details

We used a metaphor of space in designing the layout and colors of Opinion Space 1.0. The background of the space is black to signify an open expanse, and the active user is represented in the space as a yellow point with a yellow halo as a reference to the sun. Other points in the space that have not been rated by the active user are colored white to look like stars and signify neutrality, while "landmark" users are colored blue to be differentiated; all of these points have a star-like pulsing effect to engage users. Points rated positively by the active user are colored green and points rated negatively are colored orange, where the color applies only to the active user's view of the space; these colors make it easier for the active user to remember which points he or she has previously rated, a common request from early users. As shown in Appendix E, we iterated through a number of color choices before arriving at these.

Figure 29: Screenshot of the zoom bar in Opinion Space 1.0.

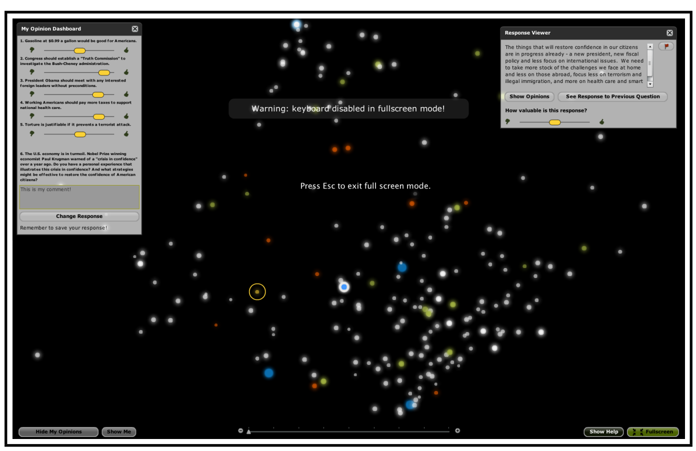


Figure 30: Screenshot of fullscreen mode in Opinion Space 1.0.

"Information Overload" A key factor in our design was the avoidance of "information overload," and the first feature to deal with that problem was zooming: if there are too many points in the space, the user can zoom in to focus on a particular area. This functionality is accessible via the zoom bar at the bottom of the screen, as shown in Figure 29. The user can pan the map in any direction when zoomed in, and the space will follow his or her point if it moves outside the border. Another way that we actively avoid "information overload" is by only showing a small set of possible points at any given time. Specifically, we show k users who have discussion question responses that the active user has rated, limited to 100. We then show l users that are the most recent in the system and who have responded to the discussion question, limited to $200 - k$. Finally, if $k + l < 200$ we show the most recent users who did not respond to the discussion question. Section 10 discusses some alternatives to zooming and limiting the number of points.

As shown in the screenshots of earlier versions of Opinion Space (see Appendix E), the dashboard and the viewers (e.g. the "Response Viewer") used to have permanent locations at the left hand side and top or bottom of the space, respectively. However, this left little room to show the space itself, which is why the dashboard and the viewers can now be dragged around by the user, as well as be closed and opened at any point. They are overlaid on top of the space and transparent to allow the points to show through as much as possible. Another important and related feature is fullscreen mode: upon clicking "Fullscreen," the space takes up the user's entire screen and spreads the points out accordingly, as shown in Figure 30; the large size of the space makes it easier to process the information.

Our scoring system makes it easier to process the information in the space by highlighting the most compelling points: it is described further in Section 9.2.3.

Understanding the Space The most common area of confusion since we began showing Opinion Space to users has been that users don't understand what their point in the space actually *means*. Users wanted the axes of the space to have semantic meaning, such as "Liberal" to "Conservative"; however, because we use PCA to generate the eigenvectors of the space, they have no meaning that is easy to describe. One simple feature that addresses this problem is "live dragging": when users move their proposition sliders, their point in space moves at exactly the same rate. This allows users to begin to understand the layout of the space. Another feature that has already been discussed is "landmark" users; in fact, this area of confusion is what prompted us to add these "landmark" users into the system. The landmarks begin to give the space semantic meaning, as certain directions in the space correspond to certain public figures. For example, if a user is closer to Nancy Pelosi than another user, some meaning can be inferred from that.

Another approach to giving the space meaning is a slideshow tutorial that we created and embedded into the website, as shown in Figure 31. We went through several iterations of text for the different slides so that it would make sense to people unfamiliar with advanced mathematics.

Some users did not understand where their point was in the space, or what to do with the space after arriving. To address these problems we introduced a "Show Me" button that centers the space on the user when clicked, a "Me" label that appears next to the user's dot, and the yellow halo that surrounds the user's dot to clearly differentiate it from all the users. We also added a "Show Help" button that displays helpful text when clicked, and a large "Suggestions" label that directs the user to click that button when he or she first arrives at the space. Figure 32 shows screenshots of the "Me" and "Suggestions" labels, as well as the "Show Help" button and screen.

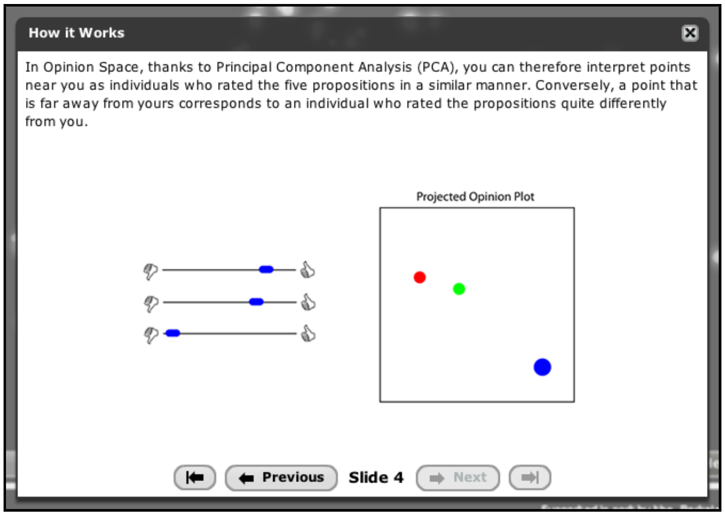
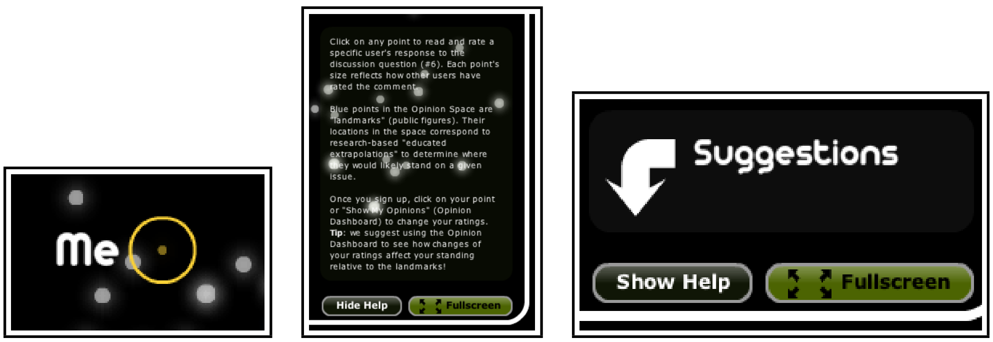


Figure 31: Screenshot of the "How it Works" tutorial in Opinion Space 1.0.



(a) Large "Me" label (b) Help screen (c) Large "Suggestions" label

Figure 32: A few guidance elements in Opinion Space 1.0.

Changing the Discussion Question The optional discussion question changes every so often to keep the dialogue fresh and exciting, which resets the sizes and colors of all points in the space because their ratings and scores are no longer applicable; that is, the ratings were for the responses to the *previous* discussion question. However, the active user can click on "See Response to Previous Question" in the "Response Viewer" to view any user's response to a previous question, a feature we added so that users would not be completely overwhelmed by a discussion question change.

Saving When the user adjusts the proposition sliders or the "Response Viewer" slider, the new ratings are automatically saved to the database. This is in contrast to earlier versions of Opinion Space, where the user was forced to click "Save" buttons in order to save changes. In some of the earlier versions, a "shadow" point and "shadow" sliders were displayed when the user moved his sliders without saving, as shown in Figure 33, which indicated the user's last saved positions. If the user clicked anywhere else in the space, his or her point would snap back to the location of

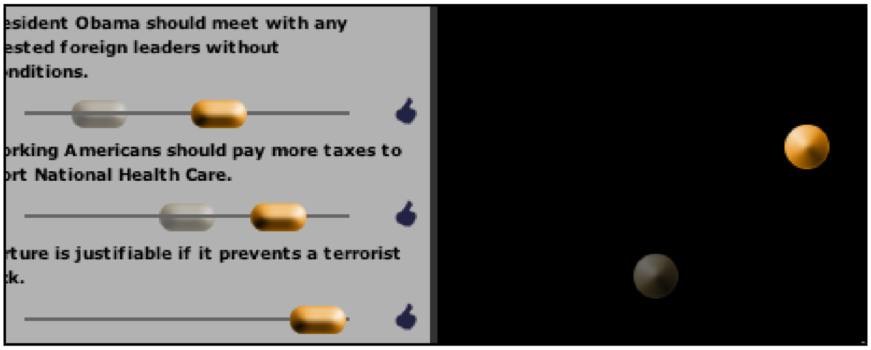


Figure 33: Screenshot of the "shadow" point and sliders in earlier versions of Opinion Space.

the "shadow" point. The goal of the "shadow" point and sliders was to remind the user to click "Save," and the reason we wanted the user to manually click "Save" was because we were worried about random exploration. That is, we were worried that users might move their sliders in various different directions in order to explore the layout of the space, without actually intending to change their opinions. However, all of this confused our early users; as a result, we decided to implement automatic saving. Section 9.4 shows that users did not, in general, change their ratings after they saw the space; thus, the impact of user exploration on the data has been minimal.

Registration In early versions of Opinion Space, users did not enjoy being asked for their email addresses and passwords for registration. Thus, the current system simply asks for email addresses and automatically assigns temporarily passwords. Users can reset their passwords by checking their email and clicking on appropriate think, and users also have the option of clicking "Register" in the navigation bar and inputting a password.

9.2.3 Scoring System

Users tend to be biased when it comes to opinions on controversial issues, which is what inspired the "devil's advocate" version of Opinion Space, the goal of which was to filter out unbiased responses to the discussion question. In determining a metric for scoring responses in Opinion Space 1.0, we take note of this bias and further assume that users will tend to rate points that are near them positively and points that are far away negatively.

We define a "compelling" response as one that is rated positively by a diverse set of users, rather than just the users with points that are nearby. In order to highlight the most compelling responses, we weight ratings as follows, as described by Bitton et. al. [4]:

Let x_i be user i's ratings vector of the Opinion Space propositions, and let $r_{ij} \in [-1, 1]$ be user i's rating of user j's response, where larger-valued ratings correspond to higher degrees of agreement. We define d_{max} to be the greatest possible Euclidean distance between the ratings vectors of two users. If there are m propositions, then:

$$d_{max} = ||r_{max} - r_{min}|| = \sqrt{4m} \quad (7)$$

When user i rates user j's response with r_{ij}, the *score* of the rating r'_{ij} is computed as follows. Let d_{ij} be the Euclidean distance (vector norm) between user i's and user j's rating vectors ($0 \leq d_{ij} \leq d_{max}$). Then:

$$r'_{ij} = \begin{cases} r_{ij}\ d_{ij} & \text{if } r_{ij} \geq 0 \quad (8) \\ |r_{ij}|\ (d_{ij} - d_{max}) & \text{otherwise} \quad (9) \end{cases}$$

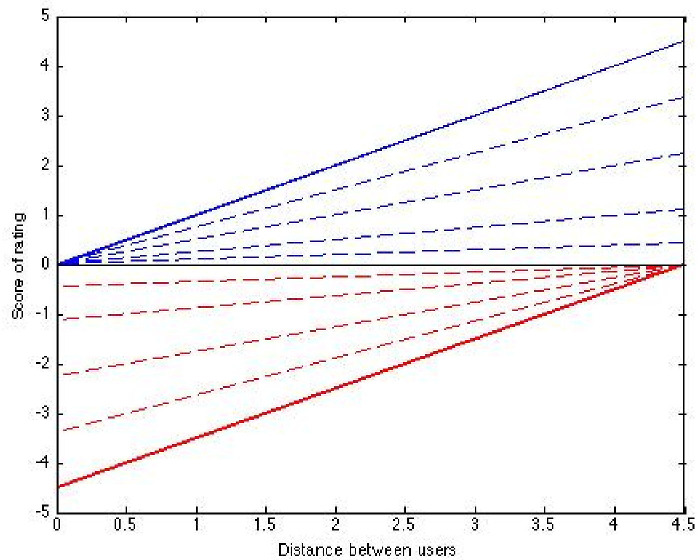


Figure 34: Plot of the rating score as a function of the distance between users i and j. The different lines correspond to different raw ratings, where the slope is the absolute value of the raw rating. The upper half of the plot corresponds to positive raw ratings and the lower half corresponds to negative raw ratings.

Figure 34 shows a plot of the scoring function. For any response, the scores r'_{ij} of all its ratings are aggregated together to form an overall score, which indicates whether or not the response is "compelling." As discussed in Section 9.2.1, the size of the point is determined by this score, as opposed to early versions of the system where point sizes simply reflected average rating.

9.2.4 Populating the System

The system currently features a single "space" that contains five propositions pertaining to US Politics. At the time of writing they are as follows:

1. Gas at $0.99 a gallon would be good for Americans.

2. Congress should establish a "Truth Commission" to investigate the Bush-Cheney administration.
3. President Obama should meet with any interested foreign leaders without preconditions.
4. Working Americans should pay more taxes to support national health care.
5. Torture is justifiable if it prevents a terrorist attack.

The optional discussion question is: "The U.S. economy is in turmoil. Nobel Prize winning economist Paul Krugman warned of a 'crisis in confidence' over a year ago. Do you have a personal experience that illustrates this crisis in confidence? And what strategies might be effective to restore the confidence of American citizens?" We plan to switch the discussion question every few weeks, and to add additional spaces on topics such as the economy, education and the environment in the near future.

In choosing the five propositions and the discussion question, we considered a few factors. First, we wanted them to be engaging so that we could drive users to the website and collect as much data as possible. Second, we wanted them to be controversial so as to spread users out in the space. Finally, we wanted to ensure that a user's response to the discussion question would tend to relate to his or her ratings of the five propositions so as to uphold our assumptions for the scoring system described above.

9.3 Opinion Space 1.0 Data Collected

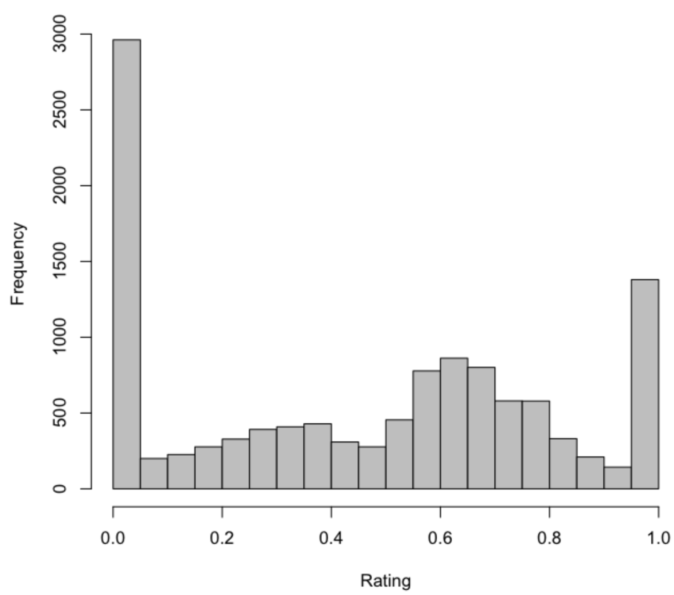


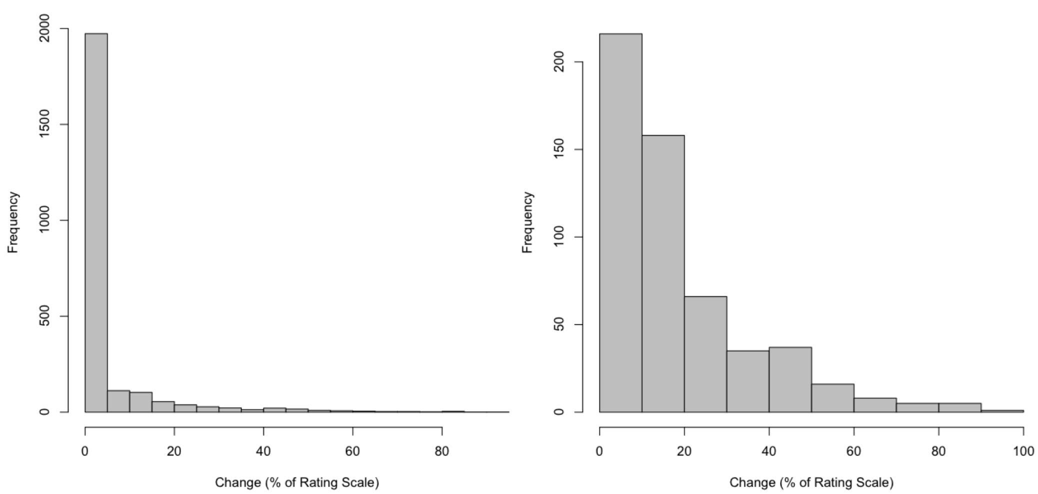
Figure 35: Histogram of Opinion Space 1.0 ratings as of early May 2009 (in the interval [0, 1]).

As mentioned in Section 9.2, Opinion Space 1.0 has collected over 18,000 opinions from over 3,700 users as of May 2009. Users tended to rate responses either very positively or very negatively, with very negative ratings being approximately twice as common as very positive ratings. The histogram of all ratings collected is shown in Figure 35. Note that this histogram only includes explicit ratings: users must actually drag the slider. If a user leaves the slider in the middle by default, but does not drag it at all, that rating is not included in this histogram.

9.4 Opinion Space 1.0 Bias Analysis

We were curious whether presenting a visualization to users would bias their ratings: perhaps they might want to move closer to clusters of other users, to "landmark" users, etc. Such a bias could be problematic in displaying opinions accurately. In order to measure this bias, we compare the ratings of users before they view the visualization with their most recent ratings.

At a point in early May 2009, Opinion Space had seen 3,083 registered users. Of those, 2,416 had rated at least one proposition, and only 547 had changed their opinions at all between their initial ratings of the five propositions and their most recent ratings. For the users who rated at least one proposition, the mean distance change of all users through five-dimensional space (as a percentage of the maximum distance possible) was 4.2%, with a histogram of the distance change per user shown in Figure 36a. For the users who *did change* their stated opinions, the mean distance change was 18.5%, with a histogram of the distance change per user shown in Figure 36b. Thus, the data collected indicate that the bias due to visualization is minimal.



(a) Users who rated at least one proposition

(b) Uses who changed their ratings

Figure 36: Histogram of the distance change per user.

10 Conclusion and Future Work

We have presented several mathematical tools to utilize and support the wisdom of crowds, and we have found that incorporating them into fun and engaging systems allows for the collection of a great deal of data. That data can then be used to improve or enhance the systems and tools.

There is ample room for future work with all of the algorithms, models and systems presented.

Eigentaste 2.0 A key extension for Eigentaste is the implementation of Incremental PCA, which allows for principal components to be calculated as updates [57]. Once that is implemented it would be important to investigate the new directions for Eigentaste that it may allow. Another important project would be the comparison of Eigentaste with more collaborative filtering algorithms.

Jester Possible extensions to Jester include allowing the user to go back and re-rate jokes, showing a list of jokes that the user has previously rated, and adding a moderation system to allow users to submit their own jokes automatically. Jester could also be able to support audio, video and images in the future, in addition to text jokes. A key feature would be linking Jester to an existing online joke database.

Donation Dashboard Future versions of Donation Dashboard would benefit greatly from an automatic payment system so that users could instantly transfer funds to the organizations in their portfolios. The addition of more organizations (and smaller organizations) to the database and the creation of local instances of Donation Dashboard would increase the reach of the system and make it more applicable to users' daily lives. For example, a "San Francisco Bay Area Donation Dashboard" would only recommend organizations in that area. Another area for future work is portfolio generation and manipulation: users could be allowed to manually adjust their portfolios prior to sharing with their peers. In terms of data analysis, it is important to analyze the extreme negativity of a large number of ratings in Donation Dashboard 1.0.

Eigentaste 5.0 Eigentaste 5.0 would benefit from a more robust method of determining similarity between items when data is very sparse. Substituting average ratings for null data dilutes the actual ratings obtained for the item, and the clustering algorithm is more likely to place the sparse items in the same item cluster. For our experiments, we only considered users that had rated all of the items in the set in order to avoid the sparsity issue. Further experimentation is possible with generalizations of the adaptive aspect of Eigentaste 5.0, which could recommend items by cycling through the top n item clusters for a user as opposed to just one. Doing so would introduce more diversity among recommendations and may further reduce portfolio effects. Another possibility is to give users the ability to modify how much item similarity they desire.

Recommending Weighted Item Portfolios Using Relative Ratings Another area for future work is the implementation of Donation Dashboard 2.0, which will utilize the the framework for recommendation weighted item portfolios using relative ratings. It will allow for the comparison of prediction error with that of Donation Dashboard 1.0 as well as the effect of altering presentation orders. The development of improved algorithms that take advantage of our graphical model would also be beneficial.

Eigentaste Security Framework Extensions of the Eigentaste Security Framework might include the consideration of attackers who obfuscate their attacks, i.e. attackers who realize that the raw Love/Hate attack, while most effective when undetected, is easily detected. An example of obfuscation would be to add some randomness to the Love/Hate ratings instead of only using r_{min} and r_{max}. Eigentaste applications would also benefit from an analysis of Eigentaste-specific attacks such as an attack that we call the "Cluster Shift attack": a huge amount of attacker profiles can be inserted into one specific region of the eigenplane, such that all clusters need to be devoted to them; then, all legitimate users are forced into the remaining cluster and the system degenerates to a sorted list by average rating. It would be extremely useful to evaluate the amount of damage that can be inflicted by any of these attacks. Another area to explore is the possibility of detecting attacks by looking at how items move around the ranks within clusters, instead of looking at individual profiles.

Opinion Space There are many potential extensions and improvements to Opinion Space. One important problem discussed earlier is that people don't understand what their position in the space actually *means*. This is not a trivial question by any means, and there is ample research opportunity in trying to aid users in understanding a position that has no semantic meaning. One key feature that would help with this problem is "lenses": upon turning on a particular lens, points in the space would be colored according to certain attributes. For example, a "Political Party" lense might color Democrats blue and Republicans red. These "lenses," especially if they could be overlaid on top of each other, would give the space meaning by giving more meaning to the points in the space. Another important feature to go along with "lenses" would be incorporating Facebook Connect into the system, as this would make available a large amount of demographic information.

In order to help the user understand the directionality of the space, the system could overlay arrows or gradients on top of the user's point indicating the directions that it could move. As the user moves his or her point, it could also highlight points that become similar to the user's point with respect to a *set* of the propositions as opposed to all of them. Zooming in Opinion Space should also not be taken for granted, as it is simply one means of avoiding "information overload." Another option, for example, might be clustering points together as meta-points.

A separate but very important extension would allow users to create their own Opinion Spaces by submitting sets of five propositions. This would allow users to create spaces featuring any topic.

In terms of engaging users, one very engaging possibility would be showing *live* movement of other points; however, this is difficult in terms of infrastructure and would require a large number of users changing their opinions at the same time. A similar but more feasible feature would be the creation of a slider that moves the points through time. Thus, users would be able to watch the points evolve from the formation of the space to the present. Adding a game-like scoring element to the system likely result in much more user engagement. Subtle, random motion of the points might also be helpful.

Another direction is to increase the amount of social features in Opinion Space. The ability to respond to discussion question responses would enable dialog in the system, although a proper interface would need to be developed. Private messaging might be a very useful feature for users who would like to meet other like-minded people, or perhaps people who are not so like-minded! One simple idea would be to show lines emerging from the selected point and connecting to the other users who rated it.

Acknowledgements

I first thank my advisor, Professor Ken Goldberg: he pushed me to my potential, taught me what research *is*, and showed me *how* to be a researcher. His leadership and and mentorship have directed me towards success, and his ideas were pivotal in each of the projects discussed. He has been more than just an advisor, though: he has been a friend, always looking out for my best interest even when it conflicted with his. It has been a pleasure working with him these past three years.

I also thank Professor Kimiko Ryokai for her contributions to the design of Opinion Space and her excellent advice and feedback on this report. Ephrat Bitton also deserves a great deal of thanks: she has had a pivotal role in the development of each of the algorithms and models presented and has been a great friend and valuable colleague over the past three years. I would not have been able to achieve nearly as much without her work and her support.

Next, I thank David Wong and Alex Sydell, who wrote the code for Opinion Space with me. The project would not be anywhere near where it is today without their hard work and insights. Also deserving much thanks are Professor Gail de Kosnik, Meghan Laslocky, Zach Blas, Elizabeth Goodman, and Siamak Faridani, who contributed regularly to the design and content of Opinion Space, as well as the directions taken. Further, Justin Woo provided fantastic technical support for Opinion Space.

I thank Jeff Tang, who contributed a great deal to the Eigentaste Security Framework, early versions of Opinion Space and technical support for Jester and Donation Dashboard. Both his work and his friendship have been extremely valuable to me over the past three years.

Ke Xu wrote some of the earliest code for Donation Dashboard that I built upon, and I am very grateful for his hard work and dedication. Designers also played a key role in Donation Dashboard and Opinion Space, and I thank Gil Gershoni for his incredible design work on Donation Dashboard, and Connie Wang for her extremely valuable designs for both Donation Dashboard and Opinion Space.

Professor Ron Fearing has been very helpful over the past few years as an academic advisor, and I very much appreciate his support and advice over the years. I also thank Professors John Canny, David Wagner, Martin Wainwright and Michael Jordan, as well as Percy Liang and Oleg Mayba, for their valuable feedback. Further, I thank Jeremy Schiff, who aided in the process of writing a large report and has been a friend and valuable colleague over the past three years. Simon Tan also deserves much thanks for giving a great deal of advice pertaining to the development of Opinion Space and data analysis.

None of these systems would have been possible without the work of the original Jester team: I thank Dhruv Gupta, Hiro Narita, Mark DiGiovanni, Chris Perkins, Robert Hennessy and Theresa Roeder for their work that I built upon.

I thank Jim Buckmaster and craigslist.org for their ideas that formed the foundation of Donation Dashboard and their support of the development of both Donation Dashboard and Opinion Space. I also thank Jay Walsh and the Wikimedia Foundation for their fantastic ideas and support of these projects.

Every member of the Automation Sciences Lab also deserves thanks for their ideas, support and help in testing. I also thank everyone who helped in testing and everyone who gave feedback; both testing and feedback were extremely valuable in the process of developing these projects. Finally, I thank the *users* of these systems: they are ones who provided all the valuable data, and I am extremely grateful for their time.

These projects have been supported in part by the Berkeley Center for New Media and the NSF Graduate Research Fellowship Program.

A Selected Jokes from Jester

Warning: some of these jokes have ethnic, sexist, or political content (some have all three). Although we tried our best, it is hard to eliminate all such factors and retain a sense of humor.

A.1 Top Five Highest Rated Jokes

One Sunday morning William burst into the living room and said, "Dad! Mom! I have some great news for you! I am getting married to the most beautiful girl in town. She lives a block away and her name is Susan."

After dinner, William's dad took him aside. "Son, I have to talk with you. Your mother and I have been married 30 years. She's a wonderful wife but she has never offered much excitement in the bedroom, so I used to fool around with women a lot. Susan is actually your half-sister, and I'm afraid you can't marry her."

William was heart-broken. After eight months he eventually started dating girls again. A year later he came home and very proudly announced, "Dianne said yes! We're getting married in June."

Again his father insisted on another private conversation and broke the sad news. "Dianne is your half-sister too, William. I'm awfully sorry about this."

William was furious! He finally decided to go to his mother with the news.

"Dad has done so much harm.. I guess I'm never going to get married," he complained. "Every time I fall in love, Dad tells me the girl is my half-sister."

His mother just shook her head. "Don't pay any attention to what he says, dear. He's not really your father."

A couple of hunters are out in the woods in the deep south when one of them falls to the ground. He doesn't seem to be breathing, and his eyes are rolled back in his head.

The other guy whips out his cell phone and calls 911. He gasps to the operator, "My friend is dead! What can I do?"

The operator, in a calm and soothing voice, says, "Alright, take it easy. I can help. First, let's make sure he's dead."

There is silence, and then a gun shot is heard.

The hunter comes back on the line. "Okay. Now what??"

A radio conversation between a US naval ship and Canadian authorities...

Americans: Please divert your course 15 degrees to the North to avoid a collision.

Canadians: Recommend you divert YOUR course 15 degrees to the South to avoid a collision.

Americans: This is the captain of a US Navy ship. I say again, divert YOUR course.

Canadians: No. I say again, you divert YOUR course.

Americans: This is the aircraft carrier USS Lincoln, the second largest ship in the United States' Atlantic Fleet. We are accompanied by three destroyers, three cruisers and numerous support vessels. I demand that you change your course 15 degrees North, that's ONE FIVE DEGREES NORTH, or counter-measures will be undertaken to ensure the safety of this ship.

Canadians: This is a lighthouse. Your call.

A group of girlfriends is on vacation when they see a 5-story hotel with a sign that reads: "For Women Only." Since they are without their boyfriends and husbands, they decide to go in.

The bouncer, a very attractive guy, explains to them how it works. "We have 5 floors. Go up floor by floor, and once you find what you are looking for, you can stay there. It's easy to decide since each floor has a sign telling you what's inside."

So they start going up and on the first floor the sign reads: "All the men on this floor are short and plain." The friends laugh and without hesitation move on to the next floor.

The sign on the second floor reads: "All the men here are short and handsome." Still, this isn't good enough, so the friends continue on up.

They reach the third floor and the sign reads: "All the men here are tall and plain."

They still want to do better, and so, knowing there are still two floors left, they continued on up.

On the fourth floor, the sign is perfect: "All the men here are tall and handsome." The women get all excited and are going in when they realize that there is still one floor left. Wondering what they are missing, they head on up to the fifth floor.

There they find a sign that reads: "There are no men here. This floor was built only to prove that there is no way to please a woman."

An explorer in the deepest Amazon suddenly finds himself surrounded by a bloodthirsty group of natives. Upon surveying the situation, he says quietly to himself, "Oh God, I'm screwed."

The sky darkens and a voice booms out, "No, you are NOT screwed. Pick up that stone at your feet and bash in the head of the chief standing in front of you."

So with the stone he bashes the life out of the chief. He stands above the lifeless body, breathing heavily and looking at 100 angry natives...

The voice booms out again, "Okay....NOW you're screwed."

A.2 Top Five Lowest Rated Jokes

Jack Bauer can get McDonald's breakfast after 10:30.

Person 1: Hey, wanna hear a great knock-knock joke?
Person 2: Sure, What is it?
Person 1: Okay, you start.
Person 2: Knock-knock.
Person 1: Who's there?
Person 2: ...
Person 1: Hah!

How many feminists does it take to screw in a light bulb?

That's not funny.

Q. What's O. J. Simpson's web address?

A. Slash, slash, backslash, slash, slash, escape.

Q. What is orange and sounds like a parrot?

A. A carrot.

A.3 Top Five Highest Variance Jokes

Chuck Norris' calendar goes straight from March 31st to April 2nd; no one fools Chuck Norris.

Person 1: Hey, wanna hear a great knock-knock joke?
Person 2: Sure, What is it?
Person 1: Okay, you start.
Person 2: Knock-knock.
Person 1: Who's there?
Person 2: ...
Person 1: Hah!

How many teddy bears does it take to change a lightbulb?

It takes only one teddy bear, but it takes a whole lot of lightbulbs.

Why are there so many Jones's in the phone book?

Because they all have phones.

When most people claim to be "killing time", it's only an expression. When Chuck Norris kills time, the minutes actually cease to exist.

B Selected Organizations from Donation Dashboard

B.1 Top Five Highest Rated Organizations

Doctors Without Borders, USA

Delivering emergency aid

Description:

- An international medical humanitarian organization that delivers emergency aid to people in nearly 60 countries.
- Provides health care, runs hospitals and clinics, performs surgeries, carries out vaccination campaigns and constructs wells and provides shelter materials.
- Doctors, nurses, logisticians, water-and-sanitation experts, administrators and other medical and non-medical professionals perform nearly 47,000 aid assignments each year.

For more information, visit this non-profit at: http://www.doctorswithoutborders.org/

Public Broadcasting Service

Description:

- A non-profit public broadcasting television service with 354 member TV stations in the United States, with some member stations available over the air and by cable in Canada.
- Mission is to inform, to inspire, and to educate.
- PBS stations are commonly operated by non-profit organizations, state agencies, local authorities (e.g., municipal boards of education), or universities in their community of license.
- Distributes popular children's educational programs such as Sesame Street, Reading Rainbow, and Mr. Rogers' Neighborhood.

For more information, visit this non-profit at: http://www.pbs.org

Kiva

Loans that change lives

Description:

- The only online micro-lending platform.
- Partners with existing microfinance organizations to identify rising entrepreneurs spanning 42 countries in the developing world. Donors (or lenders) review the entrepreneurs' profiles and choose which they wish to contribute funds. Kiva collects repayments over time and provides updates on progress the businesses make. The funds are returned to the lender, which can either be withdrawn or re-loaned.
- Has funded over 35,888 loans across the globe, with a current repayment rate of 99.88%.
- The average time it takes for a Kiva Loan to be funded is approximately 20.75 hours. The average size of a loan for funding is $546.86.

For more information, visit this non-profit at: http://www.kiva.org/

Planned Parenthood Federation of America

Your trusted provider of health information and services

Description:

- Believes in the fundamental right of each individual, throughout the world, to manage his or her fertility, regardless of the individual's income, marital status, race, ethnicity, sexual orientation, age, national origin, or residence.
- Works to improve women's health and safety, prevent unintended pregnancies, and advance the right and ability of individuals and families to make informed and responsible choices.
- Helps prevent more than 630,000 unintended pregnancies each year; provides more than 1.1 million Pap tests and more than 840,000 breast exams each year; provides more than 2.6 million tests and treatments for sexually transmitted infections; affiliates provide educational programs to 1.3 million young people and adults each year.

For more information, visit this non-profit at: http://www.plannedparenthood.org/

Engineers Without Borders

Description:

- Partners with developing communities across the globe to improve quality of life. Involves implementing sustainable engineering projects that the communities can own and operate. Places an emphasis on education in order to equip developing communities with technical, managerial, and entrepreneurial skills.
- Projects include building water, wastewater, sanitation, energy, and shelter systems.

For more information, visit this non-profit at: http://ewb-usa.org/

B.2 Top Five Lowest Rated Organizations

The NRA Foundation

Teach freedom

Description:

- Mission is to raise tax-deductible contributions in support of the wide range of firearm-related public interest activities of the National Rifle Association of America and other organizations that defend and foster the Second Amendment rights of all law-abiding Americans.
- Grants benefit programs such as youth education, law enforcement training, hunter education, conservation, and firearms training and safety.
- Has awarded more than $100 million since 1990.

For more information, visit this non-profit at: http://www.nrafoundation.org/

The Heritage Foundation

Leadership for America

Description:

- A research and educational institution whose mission is to formulate and promote conservative public policies based on the principles of free enterprise, limited government, individual freedom, traditional American values and a strong national defense.
- Communicates messages to members of Congress, key congressional staff members, policymakers in the executive branch, the news media and the academic and public policy communities.
- Launched an educational campaign called "What Would Reagan Do", helped set the stage in 1987 for a rollback of welfare state, fought health care reform in 1992 and fought against amnesty for illegal immigrants in 2007.

For more information, visit this non-profit at: http://www.heritage.org/

PETA (People for the Ethical Treatment of Animals)

Fighting for the rights of all animals

Description:

- The largest animal rights organization in the world with 1.8 million members and supporters.
- Focuses attention on four areas in which the largest numbers of animals suffer the most intensely: on factory farms, in laboratories, in the clothing trade and in the entertainment industry.
- Works through public education, cruelty investigations, research, animal rescue, legislation, special events, celebrity involvement and protest campaigns to focus attention on the ethical treatment of animals.

For more information, visit this non-profit at: http://www.peta.org/

Boy Scouts of America

The nation's foremost youth program

Description:

- Purpose is to provide an educational program for boys and young adults to build character, to train in the responsibilities of participating citizenship, and to develop personal fitness.
- Adheres to the Scout method to teach typical Scouting values such as self-esteem, citizenship and outdoorsmanship through a variety of activities such as camping, aquatics and hiking.

For more information, visit this non-profit at: http://www.scouting.org

Prison Fellowship

Reaching out to prisoners, ex-prisoners, and their families

Description:

- Partners with local churches across the country to minister to a group that society often scorns and neglects: prisoners, ex-prisoners, and their families.
- Reaches out both as an act of service to Jesus Christ and as a contribution to restoring peace to our cities and communities endangered by crime.
- Promotes the principles and practices of restorative justice Ñ focusing on healing broken relationships, repairing the damage done by crime, and restoring the offender to a meaningful role in society.
- Church-based volunteers provide post-prison support.

For more information, visit this non-profit at: http://www.prisonfellowship.org/

B.3 Top Five Highest Variance Organizations

National Public Radio

Creating a more informed public

Description:

- An internationally acclaimed producer and distributor of noncommercial news, talk and entertainment programming.
- Serves a growing audience of 26 million Americans each week in partnership with more than 800 independently operated, noncommercial public radio stations.
- Each NPR Member Station serves local listeners with a distinctive combination of national and local programming, original online content and audio streaming.

For more information, visit this non-profit at: http://www.npr.org

The Humane Society of the United States

Promoting the protection of all animals

Description:

- Works to reduce suffering and to create meaningful social change for animals by advocating for sensible public policies, investigating cruelty and working to enforce existing laws, educating the public about animal issues, joining with corporations on behalf of animal-friendly policies, and conducting hands-on programs that make ours a more humane world.
- Are the lead disaster relief agency for animals, providing direct care for thousands of animals at our sanctuaries and rescue facilities, wildlife rehabilitation centers, and mobile veterinary clinics.
- Most recent accomplishments include shutting down the last three horse slaughtering facilities in the US, leading an effective boycott against Canada's commercial seal hunt, succeeding in a six-year campaign to make illegal animal fighting a federal crime, and deploying disaster responders to help rescue animals caught in the 2007 Southern California wildfires.

For more information, visit this non-profit at: http://www.hsus.org/

The Wikimedia Foundation

Description:

- Operates online collaborative projects including Wikipedia for the purpose of collecting and developing educational content and to disseminate it effectively and globally.
- Develops and maintains open content, wiki-based projects and provides the full contents of those projects to the public free of charge.

For more information, visit this non-profit at: http://www.wikimedia.org/

American Society for the Prevention of Cruelty to Animals

We are their voice

Description:

- Attempts to alleviate the injustices animals face today by providing national leadership in anti-cruelty, humane education, government affairs and public policy, shelter support and animal poison control.
- The NYC headquarters houses a full-service animal hospital and adoption facility.
- The first animal humane organization in the Western hemisphere.

For more information, visit this non-profit at: http://www.aspca.org/

St. Jude Children's Research Hospital

Finding cures. Saving children.

Description:

- Mission is to advance cures, and means of prevention, for pediatric catastrophic diseases through research and treatment.
- No one pays for treatment beyond what is covered by insurance, and those without insurance are never asked to pay.
- All of the knowledge gained from research is freely shared with the rest of the world.
- Research is focused specifically on cancers, acquired and inherited immunodeficiencies, infectious diseases and genetic disorders. Conduct clinical research in bone marrow transplantation, gene therapy, chemotherapy, the biochemistry of normal and cancerous cells, radiation treatment, blood diseases, resistance to therapy, viruses, hereditary diseases, infectious diseases, and psychological effects of catastrophic illnesses.

For more information, visit this non-profit at: http://www.stjude.org

C Donation Dashboard Organization Legend

1: Bill and Melinda Gates Foundation
2: American National Red Cross
3: Henry Ford Health System
4: American Cancer Society
5: Goodwill Industries International
6: Amnesty International
7: Young Men's Christian Association (YMCA)
8: World Vision
9: Boy Scouts of America
10: American Heart Association
11: Public Broadcasting Service
12: Big Brothers Big Sisters
13: March of Dimes
14: Special Olympics
15: Academy for Education Development
16: Wildlife Conservation Society
17: United Negro College Fund
18: The Carter Center
19: Make-A-Wish Foundation
20: Alzheimer's Association
21: World Wildlife Fund
22: Arthritis Foundation
23: Project HOPE
24: American Civil Liberties Union Foundation
25: FINCA International
26: Puppies Behind Bars
27: One Laptop Per Child
28: National Public Radio
29: Disabled American Veterans Charitable Service Trust
30: Locks of Love
31: Doctors Without Borders, USA
32: United States Fund for UNICEF
33: CARE
34: AmeriCares
35: Teach For America
36: Peace Corps
37: National Park Foundation
38: America's Second Harvest
39: X-Prize Foundation
40: PETA (People for the Ethical Treatment of Animals)
41: The Wikimedia Foundation
42: American Society for the Prevention of Cruelty to Animals
43: The Heritage Foundation
44: Electronic Frontier Foundation
45: Habitat for Humanity
46: Fisher House Foundation
47: Downtown Women's Center
48: The Leukemia and Lymphoma Society
49: United Way of America
50: Planned Parenthood Federation of America
51: Prison Fellowship
52: St. Jude Children's Research Hospital
53: The Humane Society of the United States
54: Center for Economic Policy Research
55: William J. Clinton Foundation
56: NAACP Legal Defense and Educational Fund
57: Guide Dog Foundation for the Blind
58: Marine Toys for Tots Foundation
59: Save the Children
60: Conservation Fund
61: American Jewish World Service
62: Ashoka
63: International Medical Corps
64: AVID Center
65: Kiva
66: Asha for Education
67: Engineers Without Borders
68: Donors Choose
69: The NRA Foundation
70: American Foundation for Suicide Prevention

D Selected Discussion Question Responses from Opinion Space

D.1 Top Five Highest Scoring Responses

I don't have a "crisis of confidence" experience, but rather a few "crisis of quality" stories. I'm a big believer that if you make great stuff, people will buy it. Even though Americans are still technically savvy, creative, and innovative (though not the world leader anymore, sadly) we are not making the quality products that drive manufacturing, retail, and trade our way. When faced with buying a Toyota or a Saturn...the choice is simple. Regardless, our country was founded and flourished on the resiliency of our population. Once we start reinvesting in the education of our youth...particularly in math, science, and engineering...we will reap the benefits.

I'd be reassured if I knew that the days of ridiculous bonuses on Wall St. were over forever. No one is worth that kind of money, and certainly not when the financial sector is such a mess.

The lack of transparency on the part of financial professionals, and their shady practices. Coupled with the prevailing view of their moral ambiguity in relation to their practices cast a large shadow on the economy, and eventually dragged the economy along with them downwards. Now that we have a clearer understanding of such practices, it is prudent that we install new regulation that protects the economy.

I do not have a personal illustration of the crisis, but do have some idea about restoration strategies. The root of the problem is lack of confidence in governance. Congress should not accept campaign money from special interest groups. The apparent conflict of interest is overwhelmingly bad and is a fundamental problem.

become the leader in new technologies like renewable power

D.2 Top Five Lowest Scoring Responses

shoot them all

The party's over here!!

kill the bankers, burn the records then set up a truth and reconciliation commission

No.

xasddsa

E Early Opinion Space Mockups and Screenshots

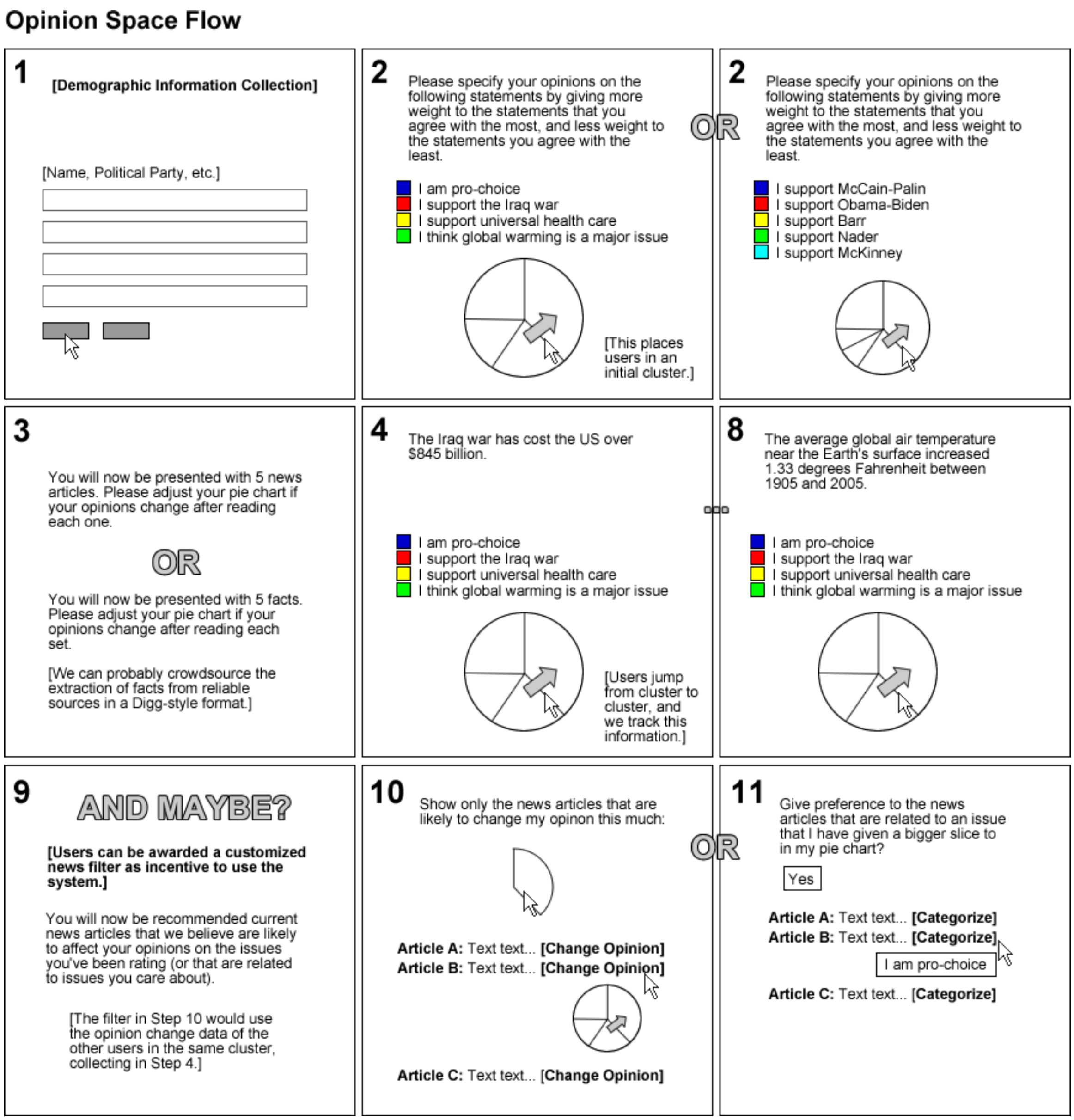


Figure 37: Mockup of Opinion Space as a news filter.

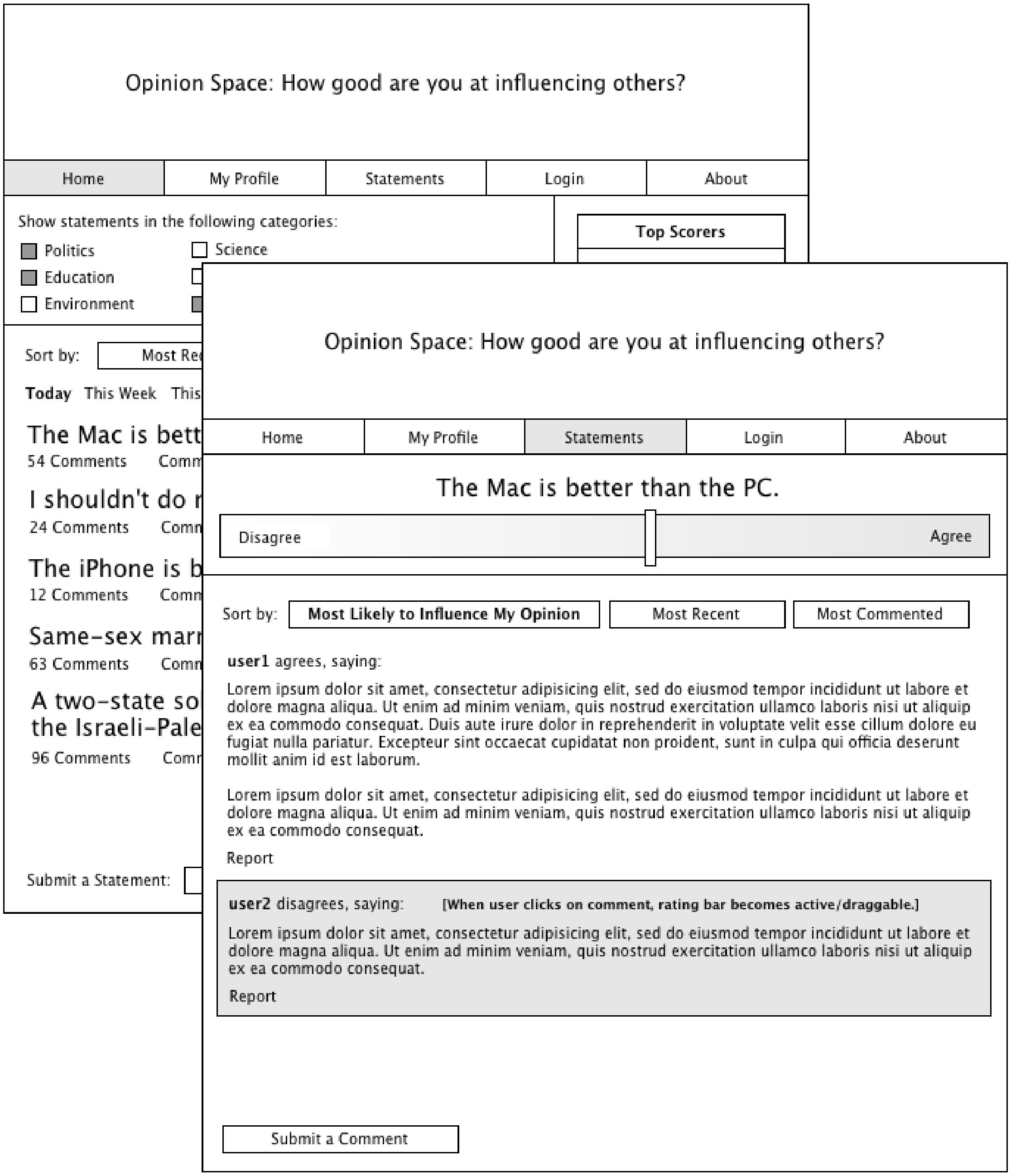


Figure 38: Mockup of Opinion Space as a "devil's advocate" system.

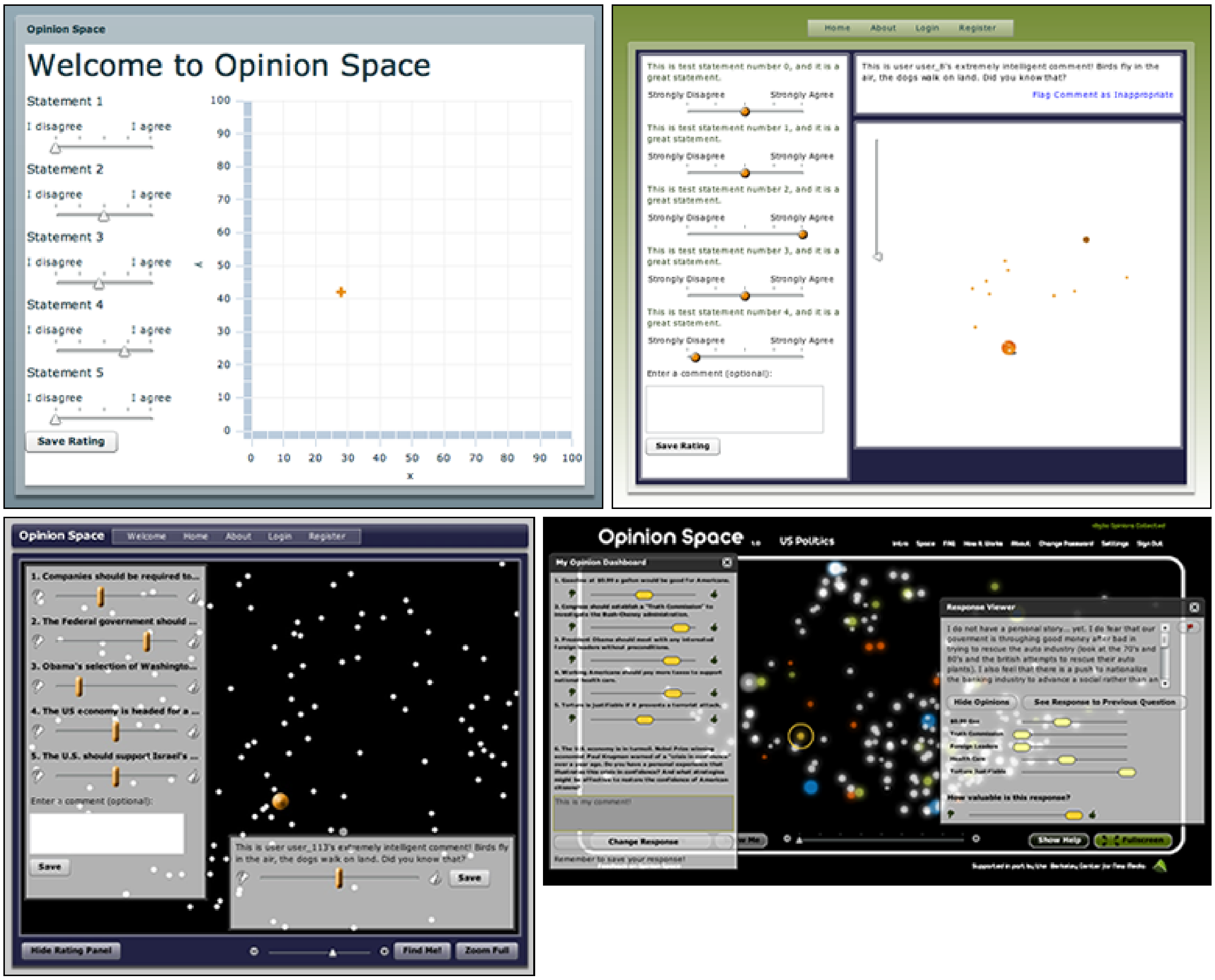


Figure 39: Screenshots of Opinion Space during development, where the fourth image is of Opinion Space 1.0.

F Selected Press Coverage

The pages to follow include selected press coverage of our various projects.

Slashdot NEWS FOR NERDS. STUFF THAT MATTERS.

Inside UC Berkeley's High Tech Joke Recommender

102

Posted by CmdrTaco on Mon Mar 31, 2008 12:41 PM
from the much-better-than-my-bone-rolling-system **dept.**

alphadogg writes

> "Every day is something like April Fools' Day at the University of California, Berkeley joke recommendation site, dubbed Jester. Now on Version 4.0, the site tosses visitors a handful of jokes to rate on a scale of "less funny" to "more funny." It then recommends jokes based on the user's taste (or lack thereof), dynamically making recommendations based on the user's most recent ratings. Jester's more than a joke jukebox though. Underlying it is a Berkeley-patented "collaborative filtering algorithm" dubbed Eigentaste , now on Version 5.0. The more people who use the system and rate jokes, the more data Berkeley researchers have to advance their understanding of recommendation systems, like those used by Amazon.com and other Web sites."

THE CHRONICLE OF HIGHER EDUCATION

The Wired Campus

Education-technology news from around the Web

April 1, 2008

Web Site at U. of California at Berkeley Tailors Jokes to Individual Tastes

What one person finds funny, another might find bland, even offensive. With that in mind, Ken Goldberg, an engineering professor at the University of California at Berkeley, and Tavi Nathanson, a Berkeley computer-science student, have created Jester 4.0, a Web site that claims to reliably tickle users' funny bones. Here's how it works: a visitor goes to Jester and rates eight jokes on a scale of "less funny" to "more funny." A computer algorithm—which Berkeley patented in 2003—then uses the data to determine which jokes will most appeal to the user. Jester forwards these jokes to the user.

When I visit the site I'm first asked to rate this joke:
Q: How many feminists does it take to screw in a light bulb?
A: That's not funny.

According to a February article in the *Berkeleyan,* entrepreneurs are interested in making the site commercial.—*Andrea L. Foster*

Posted on Tuesday April 1, 2008 | Permalink |

Berkeleyan

Appearances aside, Ken Goldberg, left, and Tavi Nathanson view humor through the prism of rigorous mathematics. *(Peg Skorpinski photo)*

So an EECS prof and an undergrad walk into a computer lab.

Jester 4.0 may seem like a lot of laughs, but to Ken Goldberg and Tavi Nathanson, jokes about dumb blondes and Chuck Norris are merely the setup for adventures in higher math

By Barry Bergman, Public Affairs | 13 February 2008

"Lemme understand this . I'm funny how?" demanded Joe Pesci, playing the impish sociopath Tommy DeVito in the 1990 mob movie *Goodfellas*. "I mean, funny like I'm a clown? I amuse you? I make you laugh?.What do you mean funny, funny how? How am I funny?"

Improve the diction, lose the badgering tone, and delete the expletives hiding amid the ellipses - on second thought, leave the expletives - and you've got a mission statement for a Berkeley research project. The punchline - wait for it - is algorithms.

And no, that's not funny. Or is it?

"Funny how?" - or, rather, "How funny?" - is the question being asked by Ken Goldberg, Tavi Nathanson, and a so-called recommender system called "Jester 4.0: Jokes for *Your* Sense of Humor." Goldberg, an engineering professor known for his machine-assisted art, and Nathanson, a senior majoring in electrical engineering and computer sciences, are employing a sophisticated algorithm to usher humor into the 21st century, even if some of the jokes are a bit stale. (Though Clinton jokes, they note, are mounting a comeback.)

Nathanson, an unabashed fan of the emergent genre of Chuck Norris jokes, joined the Jester project in the spring of 2006, when Goldberg was looking for someone to help update the program he first launched in 1998. By 2003, some 75,000 laugh-lorn computer users had visited the site, which has attracted interest from entrepreneurs wanting to take it commercial. The premise is simple enough: First-time visitors rate eight jokes on a continuum from "less funny" to "more funny." Jester then assesses the user's sense of humor, and serves up jokes to match.

The gags range from bland to bawdy - if you're offended by casual profanity, or by gentle pokes at ethnicity and religion, Leno's a safer bet - but their comedy value is up to each individual user. Humor, it turns out, is a funny thing. And that's the point. Where standups crave laughs, Goldberg and Nathanson want data. The jokes are bait. The visitors' ratings - 4 million-plus and counting - are the fish.

Thus do a pair of Berkeley intellects wade intrepidly into the murky waters of Polish jokes and dumb-blonde gags once trolled by the likes of Henny Youngman and Bazooka Joe.

A sample joke from Jester:

"Two kindergarten girls were talking outside. One said, 'You won't believe what I saw on the patio yesterday. a condom!'

"The second girl asked, 'What's a patio?'"

Rim shot. But seriously.

Take my project.please

In the spring of 2006 Nathanson, then completing his sophomore year, e-mailed Goldberg to inquire about possible research projects. They met, and Goldberg - a man renowned for cutting-edge art and creative Internet-based experiments - told him what he had in mind.

"I thought, 'Huh, that's. interesting,' " Nathanson recalls, laughing. "A joke recommender - it's just not something I would have expected to work on." But he also thought it sounded like fun, "which is not something you can say for most research projects you could get involved with." The fact that Jester was a web application, something he actually *was* interested in, sealed the deal.

He worked on the project independently over the summer. "I gave Tavi the code and said, 'See what you can do with it,' " Goldberg says. The undergrad not only made "a huge amount of progress" in overhauling the program - as Nathanson describes it, "basically taking the old stuff and rewriting it using modern languages" - but wound up deciding to pursue his master's in computer science at Berkeley. That, says his professor, will let him "do more than the interface and database and graphics, but really get into the mathematics of it."

The mathematics, indeed, are daunting. Beneath its jokey exterior, Jester matches your taste in humor to that of other users via a process known as collaborative filtering, which also drives the recommendations of sites like Amazon and Netflix. Jester is built on a complex algorithm called Eigentaste, which the campus patented in 2003.

Goldberg came up with Eigentaste after mentions in *Wired* and elsewhere sent more than 10,000 users to the original Jester, overwhelming the site. "It went crazy," Goldberg says. "It completely melted down our system." To make the program more scaleable, he borrowed a technique from the field of pattern recognition called "principal-component analysis," and Eigentaste - the name is a tip of the hat to eigenvalues and eigenvectors - was born.

And though the algorithm has grown more sophisticated, the underlying idea remains the same. While commercial sites boast recommender systems for books, CDs, and movies, Goldberg seized on jokes because he needed data, and because they're copyright-free, popular, and easy for users to evaluate. Jester itself - the name notwithstanding - is as mirthless as Dick Cheney at an ACLU fundraiser.

"The computer doesn't know anything about the jokes - every joke is just a black box, the same as it would be for a movie or a film," Goldberg explains. "It just says, OK, these are just numbers, items. What it does is ask humans to rate these items. And then, depending on how they rate them, it looks for statistical clusters, patterns. And that helps you identify people with similar tastes.

"Once you've classified users, then you can start to say, OK, someone in your taste cluster thought this other joke or movie or book was good, so I'll recommend that to you. That's the idea in a nutshell."

The statistical need for questionable gags - that is, those whose actual humor content is debatable - is also a reason Jester users continue to find so many "high variance" jokes, including some from the ever-growing oeuvre of Chuck Norris jokes. For the uninitiated, here's a favorite of Nathanson's: "Chuck Norris' tears cure cancer. Too bad he never cries."

"I find that funny. A lot of young people find that funny," he says, piercing the heavy silence that greets his delivery. "Clearly you don't."

His audience of one, it's true, belongs to the elephant-joke generation. Not that there's anything wrong with that.

It's not just jokes, folks

The current version of Jester randomly places users into one of two iterations of Eigentaste. Feedback from those who draw the newer algorithm, Eigentaste 5.0, influences the jokes they get for as long as they go on rating them. Recommendations for the rest, for the time being, are based solely on how they rate the initial set of eight jokes.

That's slated to improve, as are the quality and variety of the gags themselves. Yet while math may not be easy - to paraphrase a show-biz truism - comedy is hard. And the hardest part is finding the jokes.

"There isn't a great central joke site right now. There's no Google for jokes," laments Goldberg, who's received a few from his brother-in-law, comedian-filmmaker Albert Brooks. "We have a place on the site that lets you submit new jokes, and most of them are terrible, so we can't rely on that."

But jokes were never the end, merely a means to perfecting faster, more serious - and, arguably, more socially useful - applications. "Anything where you feel deluged, this kind of tool could be very useful in helping you pick through the huge haystacks of information," Goldberg says.

Adds Nathanson, "I always have in the back of my mind, What applications can we come up with for this?" They're already working on a charity program - tentatively dubbed "Donation Dashboard" - that will recommend portfolios of nonprofit organizations to would-be supporters based upon their interests, desired giving levels, and other preferences. As with Jester, the site will employ collaborative filtering to cluster users, in this case like-minded donors, and match them to the most suitable suite of organizations.

Goldberg, who also directs the campus's Center for New Media, has had discussions with executives at craigslist - which recently gave the center $1.6 million for an endowed chair - about promoting Donation Dashboard, which he hopes to launch in March.

Meanwhile, the pair continues to build what Goldberg calls "one of the largest data sets for research in this field," all of it available for others to mine and sift through their own algorithms. Nathanson recently gave a presentation on Eigentaste to a conference on recommender systems, and a number of published papers on Eigentaste similarly attest to the rigorous scholarship behind it.

And speaking of science. here's Jester: "What does an atheist say during an orgasm? 'Oh Darwin! Oh Darwin!'"

Read and rate jokes by visiting Jester at eigentaste.berkeley.edu/user/index.php (http://eigentaste.berkeley.edu/user/index.php) . To receive an e-mail alert of Donation Dashboard's launch, sign up at dd.berkeley.edu (http://dd.berkeley.edu) .

Search the CAA Web site:
Go
HOME ALUMNI STUDENTS CAL NEWS & LINKS DISCOUNTS & SERVICES
Join CAA | Make a Gift | Volunteer | Events | Site Map | Alumni Services
May 18, 2009
– CAA Program Links –
You are Here: Home > California

CALIFORNIA
Home
About us
Advertise
Donate
Subscribe
Contact

Past Issues
-- After 1990 --
-- Before 90s --

2009 March / April

PRAXIS

Foolproof funny

BY MARY KOLE
ILLUSTRATION BY MELINDA BECK

Better humor through computer science

For those with a refined punchline palate, Jester, the Online Joke Recommender, is a dream come true. The website's premise is simple: Read eight jokes and rate them according to how funny you find them. After that, Jester will begin suggesting new material tailored to your tastes. Whether you're a fan of shaggy-dog yarns or snappy one-liners, you'll find plenty of material for your next social gathering or paper presentation.

Ken Goldberg, robotics and engineering professor and director of Berkeley's Center for New Media, is the mastermind behind Jester. His patented algorithm, Eigentaste, now in version 5.0, specializes in what is called "collaborative filtering." If you've ever been recommended books while browsing on Amazon or movies on Netflix, you're already familiar with collaborative filtering. As Tavi Nathanson, Professor Goldberg's graduate researcher, explains it, "You take a user, you find users who are similar, you recommend items based on what those similar users said they liked."

That certainly sounds straightforward enough, but under the hood, it's all hardcore mathematics: The recommendations are based on numbers, not the content of the jokes. "That's a key point," stresses Nathanson. "The system works purely based on ratings. It's this clean statistical approach."

Goldberg's algorithm gets its name from something called "eigenvectors." Don't ask. All you need to know is that "eigen" in German indicates membership in a group—fitting, since Eigentaste finds people with the same taste in things and groups them accordingly. Once you've evaluated the initial jokes, Eigentaste places you into one of 64 different categories of joke lovers—your specific humor profile. Why 64? Nathanson explains that the cluster count could have been any power of 2, but their "gut feeling" was that 64 was broad enough to present a wide array of humor profiles but not so many as to split whiskers.

Unlike the collaborative filters used by Netflix and Amazon, Jester is designed for cases where users can rate a set of items on-the-fly. That doesn't work so well with books or movies or toaster ovens, which take time to evaluate. But it's perfect for jokes.

The immediacy of jokes makes Jester addictive, or "sticky" in Web lingo—a much sought-after quality in websites. On average, about 20 users daily rate about 46 jokes on Jester in a sitting. "Jokes have a naturally magnetic property," says Professor Goldberg. "People are happy to sit down and read them, and you can evaluate [our set of 8] in about a minute." Happily, that has provided Goldberg's team with beaucoup data to work with; there are over 4 million ratings so far.

Jester is not without drawbacks. Perhaps the biggest problem is that jokes quickly grow stale; they're never as funny the second time around. The system, which currently has 128 jokes, has a built-in field where users can submit their own gut-busters, but precious little new humor makes the cut.

So is Jester accurate in matching jokes to visitors' tastes? It's a tough question but maybe not the best one to ask, argues Nathanson. He points out that if a user is drawn to, say, Chuck Norris jokes, the safest recommendation would probably be to offer more of the same. Riskier recommendations, on the other hand, while bound to hurt accuracy, "would most likely make the system more useful or enjoyable." That's undoubtedly true. Just don't tell Chuck Norris.

Mary Kole is a San Francisco writer who knows that the sky is only blue because it's Chuck Norris's favorite color.

About CAA Contact Us Update your Address

CAA Career Opportunities Privacy Policy

©2009 California Alumni Association. All Rights Reserved
For questions about CAA: info@alumni.berkeley.edu
Technical inquiries: web@alumni.berkeley.edu

Site design by: emdesign studio

Site construction by: M&I Technology Consulting

Alumni House
Berkeley, CA 94720-7520
Toll-Free: (888) CAL-ALUM
Phone: (510) 642-7026
Fax: (510) 642-6252

Miss Sunday's Bay to Breakers? Catch a replay and highlights of the webcast!

LOCAL NEWS

Website guides intelligent charity giving

Monday, June 09, 2008 | 10:12 AM

By Terry McSweeney

BERKELEY, CA (KGO) -- Researchers at UC Berkeley are developing a way for you to give away money more intelligently, but they need your help. They've come up with a website that guides kind-hearted people to the charity best suited to their interests and their budgets. Donation Dashboard doesn't want your money, just a little time.

Christian Ramirez will contribute a portion of the $300 billion Americans will give to charity this year. Like many, he sometimes bases his decision on a flyer he receives in the mail.

"Like kids with cancer, that kind of stuff - that is the charities that I give to. So I give to direct mail," said Ramirez.

Story continues below

Advertisement

Others give to phone solicitors, but Vivi Fissekidou has stopped doing that.

"You never really know who they are or what they are and what they're using the money for. So I think it would make more sense to research an issue," said Fissekidou.

Enter Donation Dashboard. You rate fifteen charities, from not interested to very interested, enter your contribution amount, and voila - a non-profit portfolio, including your chosen charities and some similar ones you may never have thought of, all based on statistical analysis.

It's the work of the UC Berkeley Center for New Media and director Ken Goldberg.

"Essentially it uses the wisdom of crowds, when you put in your information, you are helping this greater service by collectively letting this statistical engine be more intelligent," said Goldberg.

The spirit and mission of Donation Dashboard is what attracts contributor and Craigslist CEO Jim Buckmaster.

"Once this thing is up and running on more than one server and has more nonprofits and gets rid of the rest of the bugs - I'll be interested in presenting it to Craigslist users who I think will really enjoy using it," said Buckmaster.

Ephrat Bitton had to choose the initial 70 charities out of the one million available.

"Charities that people were familiar with, but also nonprofits that were doing creative work that people might not have heard of, but people would be interested in donating to," said Ephrat Bitton, UC Berkeley Center for New Media.

"The highest rated ones are Doctors without Borders, KIVA, NPR," said Tavi Nathanson, UC Berkeley Center for New Media.

Those are the most popular right now, but San Francisco based nonprofit Youth Speaks isn't even on the list.

Artistic director Marc Joseph believes, when Dashboard grows up, Youth Speaks will be there.

"Much like MySpace or Facebook or other social networking sites that began with only a few members that are now in the billions," said Marc Joseph, Youth Speaks.

Philanthropist Andrew Carnegie once said it's more difficult to give money away intelligently than it is to earn it in the first place. Well, Donation Dashboard, as the name suggests, wants to put you in control of your contributions so you can make an intelligent choice.

(Copyright ©2009 KGO-TV/DT. All Rights Reserved.)

MarketWatch

JENNIFER OPENSHAW

Jan 21, 2009, 7:22 p.m. EST

The 15-Minute Tip: New tool for charitable giving

Donation Dashboard suggests charitable groups that best suit you

By Jennifer Openshaw

NEW YORK (MarketWatch) -- When you want a new jacket, you go to the store to buy that new Calvin Klein one you've been eyeing. You want that cool Coldplay album? You go to the store to buy the CD or to iTunes to download it. But how do you find a charity worth your money?

That seems a little more complicated, doesn't it? Andrew Carnegie once said, "It's more difficult to give money away intelligently than it is to earn it in the first place" -- and he didn't have to deal with the onslaught of information we all face nowadays.

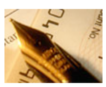
Well, the folks behind Donation Dashboard say it can be just as easy. Their online tool recommends a donation portfolio that is tailored specifically to your own interests. Sure, you already give money to the groups you know and like, but don't you think there could be others you might be interested in? Donation Dashboard can help you figure that out. See the site.

Ken Goldberg, founder of Donation Dashboard and a professor at U.C. Berkeley, said he created the tool because "people are feeling economically pinched, but still want to be able to share in an effective way."

Here's how it works: First, you take an online survey asking you to rate some non-profit organizations. Then, using a scale, you indicate how interested you are in, say, the ACLU, Habitat for Humanity or your local symphony. The tool then offers recommendations that suit your preferences and available funds.

Developed by Goldberg and several students as a project for the Berkeley Center for New Media, Donation Dashboard extends techniques used by commercial Web sites to recommend movies, music and books. It goes far beyond existing charity-ranking sites by statistically combining your ratings with the ratings entered by other good Samaritans to compute a portfolio customized to your interests.

But why should you do this? Three reasons:



Check out Personal Finance

From real estate to retirement, MarketWatch covers the topics that are vital to your pocketbook. Don't miss these recent winners:

- What the stimulus holds for you
- Falling prices may be fleeting
- To buy a home or not to buy
- Four ways to shore up nest eggs
- Ten investment ideas in 2009

Get our free PF Daily newsletter

1. Broaden your philanthropic reach. Not only do you learn more about yourself and your interests, but you'll discover organizations that match your philosophy but don't have the marketing dollars to get your attention.
2. Allocate your scarce dollars best. You don't have as much to give as you'd like. Nobody does. Donation Dashboard can help you put money towards the issues and organizations about which you feel strongest.
3. Feel like a big-shot philanthropist, if only temporarily. It's fun to see what it would be like to give the sum of your choice to the causes you believe in. Go ahead, type in a million dollars!

Donation Dashboard doesn't endorse any particular charity nor take a percentage of gifts. It really isn't a fundraising tool, either. It just opens your mind to the possibilities of charitable giving. Now that's a fun ride.

Jennifer Openshaw, author of " The Millionaire Zone ," is co-founder and president of WeSeed, whose mission is to help people use what they know to learn about stocks, make smarter money decisions and take control of their financial destinies. You can reach her at jopenshaw@weseed.com .

Copyright © 2009 MarketWatch, Inc. All rights reserved.
By using this site, you agree to the Terms of Service and Privacy Policy.

Intraday data provided by Interactive Data Real Time Services, a division of Interactive Data Corp. and subject to terms of use. Historical and current end-of-day data provided by Interactive Data Pricing and Reference Data. More information on NASDAQ traded symbols and their current financial status. Intraday data delayed 15 minutes for Nasdaq, and 20 minutes for other exchanges. Dow Jones Indexes(SM) from Dow Jones & Company, Inc. SEHK intraday data is provided by Comstock and is at least 60-minutes delayed. All quotes are in local exchange time. Real-time last sale data provided by NASDAQ.

SIGN IN OR CREATE ACCOUNT

SUGGEST A LINK | ARCHIVES | SUBSCRIBE | MARK | CORY | DAVID | XENI | JOHN | MODERATION POLICY

SEARCH

boingboing
A Directory Of Wonderful Things

Donation Dashboard: collaborative filter-enhanced charity

POSTED BY DAVID PESCOVITZ, APRIL 21, 2008 8:39 AM | PERMALINK

Donation Dashboard is a new project from UC Berkeley's Center for New Media to match non-profits with individual donors. Developed by Ken Goldberg and his colleagues, it's based on collaborative filtering, the same technique used by Amazon, for example, to recommend books based on the "wisdom" of the crowds. The notion is that people who agreed in the past will likely agree in the future about certain things. From the project page:

> Here's how it works: you are presented with brief descriptions of non- profit institutions and asked to rate each in terms of how interested you are in donating to it. The system analyzes your ratings in light of others' ratings and does its best to allocate your available funds in proportion to your interests. Your customized "donation portfolio" is presented in an easy-to-understand pie chart that you can save at the site for future reference.
>
> The Donation Dashboard website is a pilot system that includes information on 70 non-profit institutions. If the system is successful, the developers hope to expand it with other features and partner with a third party that can streamline collecting and distributing funds.
>
> "There's strength in numbers; the system should improve over time as the number of ratings increases, in this sense each person who visits the site contributes to the collective wisdom about good causes," notes UC Berkeley Professor Ken Goldberg, who is developing the system with graduate students Tavi Nathanson and Ephrat Bitton at UC Berkeley, with conceptual input from Jim Buckmaster at craigslist.

Link

posted in: ENVIRONMENT

FAVORITE THIS! (4) SEND digg it

OLDER
JESSICA RABBIT "UNTOONED"

NEWER
STRANGE FOODS FROM EDIBLE.COM

Discussion

The Philanthropy 400
The Raiser's Razor
Too Busy to Fundraise

Prospecting
News and tips on fund raising

September 08, 2008

New Web Site Recommends Charities to Donors

A new Web site enables individual donors to get a customized list of charities that match their interests.

Created by the Berkeley Center for New Media at the University of California, the Donation Dashboard site works by using software similar to what Amazon.com and other commercial sites use to recommend books, movies, and other products to customers, based on their previous purchases.

Visitors to Donation Dashboard first rate 15 charities — each presented with a brief description of its mission, accomplishments, and the percentage of its budget spent on charitable programs – -in terms of how interested they are in giving to the organization. Based on donors' ratings and statistics from other users whose preferences are similar, the site provides a list of charities likely to meet a donor's interests.

Ken Goldberg, who heads the Berkeley center unit that developed Donation Dashboard, says that as more and more donors use the site over time, its ability to match people with causes will improve as information on people's preferences is entered, and the interests of their peers can be factored into its recommendations.

Since Donation Dashboard made its debut in May, 1,600 people have provided 25,000 ratings of the 70 charities in the site's listings. Mr. Goldberg says the site may expand the number of charities that it can recommend if it can attract enough interest and software-development support.

— Holly Hall

Monday September 8, 2008 | Permalink

Commenting is closed for this article.

Previous: Use Your Annual Report to Promote Donations
Next: How Charities Use Technology Tools to Find Donors

Copyright © 2009 The Chronicle of Philanthropy

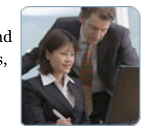
Print • Close Window

PND PHILANTHROPY NEWS DIGEST
A SERVICE OF THE FOUNDATION CENTER

CONNECTIONS

Posted on February 9, 2009

Donation Dashboard



The Berkeley Center for New Media recently launched a Donation Dashboard to help donors choose causes and organizations to support. Similar to commercial Web sites that recommend movies, music, and books to users, the site allows visitors to rate nonprofits, analyzes the ratings, and generates giving suggestions based on a donor's preferences. The suggestions are provided in a pie chart that can be saved for future reference. Currently in beta, the site includes information on approximately seventy organizations.

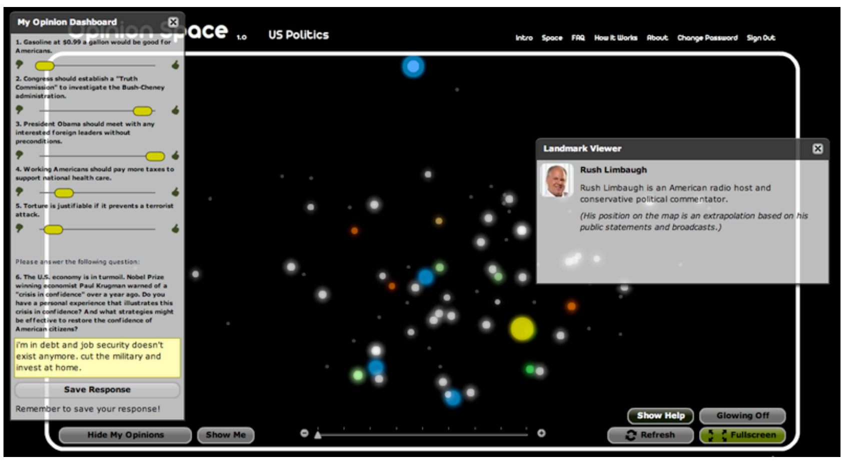
foundationcenter.org

©2009 Foundation Center
All rights reserved.

Epicenter
The Business of Tech

New Tool Plots Online Comments Like Stars in Constellations

By Ryan Singel April 22, 2009 | 5:28 pm | Categories: Uncategorized



Participatory media may be the future, but a look at most comment threads shows that technology hasn't figured out a good way to force humans to act like citizens instead of fifth graders.

UC Berkeley's Center for New Media hopes it has a way to fix that mess in its Opinion Space visualization tool, which provides a planetarium view of users opinions.

Opinion Space, which launched Wednesday, is quite pretty, mildly addictive and full of rich possibilities for visualizing a community's opinions. Wired.com would love to have such a tool at its disposal, though its almost certainly going to be an addition to, rather than an substitute for, traditional comment systems.

The center built the tool as a response to President Barack Obama's call for greater civic participation, which it says is "designed to go beyond one-dimensional polarities such as left/right, blue/red to actively encourage dialogue between people with differing viewpoints."

In a test now live on the Berkeley website, the system plots each commenter on a 2D space — your screen — based on answers to five questions. Most respondents so far land somewhere near the large nebula representing perpetual, left-wing presidential candidate Ralph Nader and far from the bright sun representing right-wing radio provocateur Rush Limbaugh.

Each respondent glows in turn on the map, and their star rises or falls depending on how others view their written answer to a sixth question.

For those who want to know how five answers get plotted on a flat surface, it's got alot to do with where you put the lamp and how the shadows fall.

For those who don't care, but like toys, head over to Opinion Space to check it out, and let us know what you think in our boring old, linear comment section.

See Also:

- Spam Clutters Environment, Not Just Inboxes
- Mysterious Anti-Obama Text Spam Slams Cellphones
- Can 'Encouraged Commentary' Bring Conversations Back to the Blog?
- Spammers Clog Up the Blogs

Post Comment | Permalink

Comments (2)

Posted by: ZuDfunck Dude | 04/23/09 | 7:54 am

Not smart enough to make it work
I hate them in football too!
My ma went there though
Should be more respectful
I am acting like a 5th grader
Aren't I

Posted by: jag | 04/23/09 | 12:11 pm

The site has a lot of potential.
Even better, was the associated site" Jester: The Online Joke Recommender" at
http://eigentaste.berkeley.edu/user/index.php

After a short training session, the program will be making you laugh 'till you puke.

---

blue of california Rates as low as $52* per month get a quote
*LEGAL

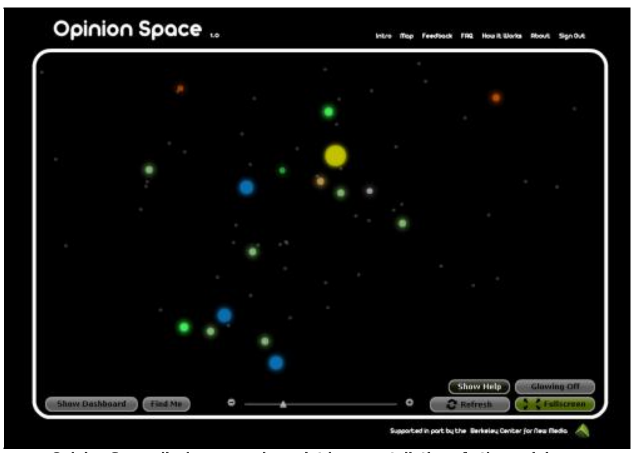
SFGate.com Quick Search GO

HOME • NEWS • SPORTS • BUSINESS • ENTERTAINMENT • FOOD • LIVING • TRAVEL CLASSIFIEDS • JOBS • REAL ESTATE • CARS

MAIN | CONTACT

« Another year of... | Main | U.S. cybersecurity... »

Opinion Space -- new tool maps your viewpoint



Opinion Space displays your viewpoint in a constellation of other opinions.

Everybody has an opinion. But not all opinions fall neatly into one of two sides -- left or right, blue or red, Mac or PC (well, OK, maybe the last example).

To help highlight the diversity of opinions on any given topic, the Berkeley Center for New Media is launching today a new online visualization tool called Opinion Space. Described as an "experimental new system for group discussion," the goal of Opinion Space is to move away from oversimplified, two-sided debates that can limit dialogue and provide a tool that encourages interaction between those with differing views.

To get started, participants are asked to drag a slider to rate five propositions on the chosen topic and type their initial response to a discussion question. Then using Principal Component Analysis from advanced mathematics, Opinion Space plots your overall opinion as a yellow point (or star) in the constellation of other viewpoints. People with similar opinions will be close in proximity.

Opinion Space's layout is determined completely by the data entered by participants. So if your yellow dot is located on the left, that doesn't mean you're more liberal. Opinion Space is designed to move beyond the usual left-right linear spectrum.

CONTRIBUTORS
Suzanne Herel (moderator)
Tom Abate
John Batteiger
Deborah Gage
Ryan Kim
Verne Kopytoff
Bios »

RECENT ENTRIES
- Craiglist CEO waiting for an apology
- Stanford's iPhone class hits 1 million downloads on iTune U
- New Web attacks redirect Google searches
- Video game industry muddles through tough April
- Warranty robo-call telemarketing companies sued by FTC

More »

CATEGORIES
Blogging (13)
Apple (231)
Ask.com (2)
Broadband (1)
CES 2007 (44)
CES 2008 (49)
CES 2009 (29)
China (1)
China Chronicles 2007 (14)
Chips (5)
Chronicle Podcasts (34)
Citizen Journalism (18)
Clean Tech (12)
cloud computing (2)
Craigslist (4)
Current (1)
D5: All Things Digital conference 2007 (4)

spectrum.

Opinion Space also includes "landmarks" (blue dots) that represent the opinions of public figures based on "educated extrapolation." In the demo that I tried out, the discussion topic was U.S. politics, and the "landmarks" were Rush Limbaugh, Arnold Schwarzenegger, Nancy Pelosi and Ralph Nader. I was nowhere near Rush, but somewhat surprisingly, closest to the governor in my views.

Ken Goldberg, director of the Berkeley Center for New Media and one of the UC Berkeley professors who developed the system, sees this tool being applied to just about any context, from something as local as a discussion on neighborhood zoning to a company wanting to survey its customers about a new product. Regardless of the topic, he said the point is to "get out of the polarities that limit discussion."

One of my initial takes was that I thought the visualization was intriguing -- what if you applied it to dispute resolution, or job interviews, or finding the perfect mate? But I also wanted more ability to actually engage in a discussion and an embed option.

Goldberg said that this is just the start. Like many things on the Web, Opinion Space will evolve and add features, such as giving users more control of the discussion and incorporating it into other sites (maybe Facebook, Craigslist, Wikipedia).

What type of features would you like to see? How might you use this tool?

Posted By: Marcus Chan (Email) | April 22 2009 at 12:15 AM
Listed Under: New Media

(9)

Add Your Comment Sort Comments by: **Oldest First** | Newest First

frank_emery 4/22/2009 7:29:06 AM

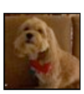
This is very cool - and gives you a chance to express some opinions in a text box that other users can read (this is anonymous) and to see if you are out on the fringe (of course OF respondents, not necessarily real) - and NO TEA PARTY RHETORIC as far as I could see - I urge you to take a few minutes and run it...after all, this may be the election technology of the future...

Recommend: (4) (1) [Report Abuse]

nightridingnow 4/22/2009 7:29:27 AM

No thanks, I'd rather get out and meet and incite people to do hands on and real expressions of opinion like boycotts, picketing, education and sometimes even more.

Recommend: (4) (3) [Report Abuse]



gotstugots 4/22/2009 7:35:47 AM

Fascinating. As for the visual graphic, rather than left/right, the spectrum of opinions seems to run top to bottom, with Rush at the top, Arnold in the middle, and Pelosi lower down. Much like the author, I am closer to Arnold than I may have thought. Now, you lefties, don't scream when all the Dittoheads find a hidden meaning that Limbaugh is oriented at the 'top' of the universe!

E-commerce (1)
E-recycling (3)
E3 2008 (17)
eBay (5)
EFF (1)
Entertainment (28)
Facebook (11)
Firefox (1)
Gadgets (38)
Google (132)
Hardware (45)
Health (10)
Hewlett-Packard (5)
Innovation (120)
Macworld 2007 (9)
Macworld 2008 (11)
Macworld 2009 (1)
Maps (1)
Microsoft (14)
Misc. (126)
Mobile (1)
New Media (129)
News (142)
Obama (1)
Omidyar (1)
Online advertising (1)
OpenTable (1)
Privacy (12)
Pwnage (5)
Search (1)
Security (42)
Silicon Valley law (12)
Social networking (23)
Software (28)
Stock options (11)
SXSW 2007 (15)
Tech Culture (173)
TechCrunch50 (5)
Ted 2007 (12)
TED 2008 (9)
TED 2009 (12)
Telecom (39)
TheStreet.com (1)
Tokoni (1)
Twitter (3)
Valleywag (2)
Venture capital (51)
Video games (125)
Visions of India (12)
Web 2.0 (168)
Webby Award (2)
Wireless (90)
Yahoo (66)
YouTube (4)
Zazzle (1)

ARCHIVES

« MAY 2009

SUN	MON	TUE	WED	THU	FRI	SAT
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16

Sponsored by:

NETWORKWORLD

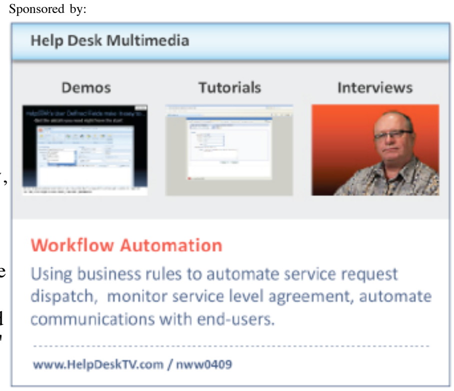
This story appeared on Network World at
http://www.networkworld.com/news/2009/042209-uc-berkley-opinion-space.html

UC Berkeley's new Opinion Space Web site visualizes your viewpoints

Creators of Jester online joke recommender now want to know what you think
By Bob Brown , Network World , 04/22/2009

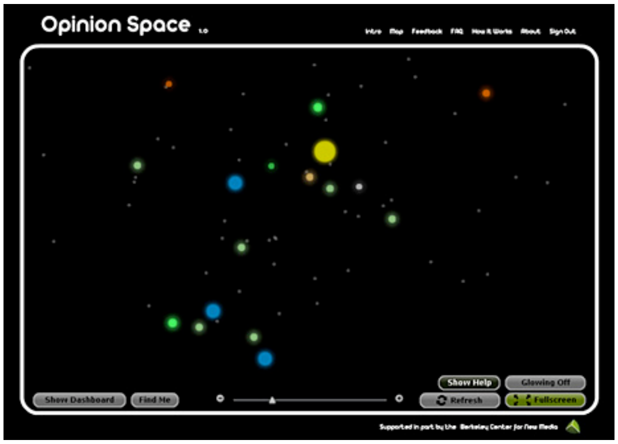
You might have a good sense of how your political opinions stack up against those of Arnold Schwarzenegger, Nancy Pelosi and Ralph Nader, but a new Web site from the University of California-Berkeley researchers can also show you how your viewpoints stack up against many others.

Sponsored by:



Opinion Space, which formally launches on Wednesday, is the brainchild of researchers at the Berkeley Center for New Media, the same outfit that brought us the Jester joke recommendation system and the Donation Dashboard, which aids people in giving to charities. The team behind Opinion Space calls it "an experimental model of opinion and dialogue… designed to go beyond one-dimensional polarities such as left/right, blue/red..." They have been working on the site since December.

Like earlier systems from Berkeley Center for New Media, Opinion Space relies on a technology called collaborative filtering to plot your opinions against others' ideas. Collaborative filtering combines "continuous spatial models with statistical techniques (such as Principal Component Analysis) to efficiently represent gradations of input," according to the Opinion Space site.



Click to see: Screen shot of opinion space

The Web site gives a 3D sort of perspective on a user's opinions, depicting points of view as a constellation of stars glowing brighter or appearing nearer or farther depending on how alike or different they are. Users can check out the opinions of others (including extrapolations of Schwarzenegger's, Pelosi's and Nader's based on their public statements and policies) and change their own opinions over time.

Users are asked to use an on-screen slider to show how much they agree or disagree with a series of statements. The initial series of questions focuses on U.S. politics and includes propositions such as "Gasoline at $0.99 a gallon would be good for Americans" and "President Obama should meet with any interested foreign leaders without preconditions." It also includes an open-ended question: "The U.S. economy is in turmoil. Nobel Prize winning economist Paul Krugman warned of a 'crisis in confidence' over a year ago. Do you have a personal experience that illustrates this crisis in confidence? And what strategies might be effective to restore the confidence of American citizens?"

Down the road, other sets of questions will be posed on topics such as education and the environment.

"We definitely plan to develop future Opinion Spaces on techie topics and we think it could be applied in many contexts, from community groups, classes and companies asking for input on their products and services," says Ken Goldberg, a UC Berkeley professor and director of the Berkeley Center for New Media.

Goldberg says Opinion Space was inspired by "frustration with the overload of Facebook and discussion forums, and more importantly by Barack Obama's exciting thoughts on 'participatory democracy'."

Django 1.0 is the Web framework used on Opinion Space's back end and Flex 3 is used for the front end. The system is running with a MySQL database on a Linux server that's using Apache, says Tavi Nathanson, a UC Berkeley electrical engineering and computer science graduate student who is part of the Opinion Space team.

Follow Bob Brown on Twitter and via his Alpha Doggs network research blog.

All contents copyright 1995-2009 Network World, Inc. http://www.networkworld.com

References

[1] K. Ali and W. van Stam. TiVo: making show recommendations using a distributed collaborative filtering architecture. *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 394–401, 2004.

[2] Amelie Anglade, Marco Tiemann, and Fabio Vignoli. Complex-network theoretic clustering for identifying groups of similar listeners in p2p systems. In *RecSys '07: Proceedings of the 2007 ACM conference on Recommender systems*, pages 41–48, New York, NY, USA, 2007. ACM.

[3] R. Bhaumik, C. Williams, B. Mobasher, and R. Burke. Securing collaborative filtering against malicious attacks through anomaly detection. In *Proceedings of the 4th workshop on intelligent techniques for web personalization (ITWP06), Held at AAAI*, 2006.

[4] Ephrat Bitton, Tavi Nathanson, and Ken Goldberg. Opinion space: A spatial model for collaborative filtering of comments in an online discussion forum. Submitted to the 2009 ACM conference on Recommender systems, May 2009.

[5] J.S. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. *Learning*, 9:309–347, 1992.

[6] Derek Bridge and Francesco Ricci. Supporting product selection with query editing recommendations. In *RecSys '07: Proceedings of the 2007 ACM conference on Recommender systems*, pages 65–72, New York, NY, USA, 2007. ACM.

[7] Kenneth Bryan, Michael O'Mahony, and Pádraig Cunningham. Unsupervised retrieval of attack profiles in collaborative recommender systems. In *RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*, pages 155–162, New York, NY, USA, 2008. ACM.

[8] R. Burke, B. Mobasher, and R. Bhaumik. Limited knowledge shilling attacks in collaborative filtering systems. In *Proceedings of the 3rd IJCAI Workshop in Intelligent Techniques for Personalization*, 2005.

[9] R. Burke, B. Mobasher, R. Bhaumik, and C. Williams. Segment-Based Injection Attacks against Collaborative Filtering Recommender Systems. In *Data Mining, Fifth IEEE International Conference on*, pages 577–580, 2005.

[10] D. Cosley, S.K. Lam, I. Albert, J.A. Konstan, and J.T. Riedl. Is seeing believing?: how recommender system interfaces affect users' opinions. *Proc. of the Annual SIGCHI Conf.*, pages 585–592, 2003.

[11] M. Deshpande and G. Karypis. Item-based top-N recommendation algorithms. *ACM TOIS*, 22(1):143–177, 2004.

[12] Y. Ding, X. Li, and M.E. Orlowska. Recency-based collaborative filtering. *Proc. of the 17th Australasian Database Conf.*, 49:99–107, 2006.

[13] L.C. Freeman. Visualizing Social Networks. *Journal of Social Structure*, 1(1):4, 2000.

[14] T. George and S. Merugu. A scalable collaborative filtering framework based on co-clustering. *Proc. of the 5th IEEE Int'l Conf. on Data Mining*, pages 625–628, 2005.

[15] D. Goldberg, D. Nichols, B.M. Oki, and D. Terry. Using collaborative filtering to weave an information tapestry. 1992.

[16] K. Goldberg, T. Roeder, D. Gupta, and C. Perkins. Eigentaste: A Constant Time Collaborative Filtering Algorithm. *Information Retrieval*, 4(2):133–151, 2001.

[17] Dhruv Gupta, Mark Digiovanni, Hiro Narita, and Ken Goldberg. Jester 2.0 (poster abstract): evaluation of an new linear time collaborative filtering algorithm. In *SIGIR '99: Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, pages 291–292, New York, NY, USA, 1999. ACM.

[18] Tarik Hadzic and Barry O'Sullivan. Critique graphs for catalogue navigation. In *RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*, pages 115–122, New York, NY, USA, 2008. ACM.

[19] J. Heer and d. boyd. Vizster: Visualizing Online Social Networks. In *Proceedings of the 2005 IEEE Symposium on Information Visualization*, pages 33–40, 2005.

[20] J.L. Herlocker, J.A. Konstan, and J.T. Riedl. An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms. *Information Retrieval*, 5(4):287–310, 2002.

[21] J.L. Herlocker, J.A. Konstan, L.G. Terveen, and J.T. Riedl. Evaluating collaborative filtering recommender systems. *ACM TOIS*, 22(1):5–53, 2004.

[22] D.S. Hochbaum and A. Levin. Methodologies and Algorithms for Group-Rankings Decision. *Management Science*, 52(9):1394, 2006.

[23] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6):417–441, 1933.

[24] IT Jolliffe. *Principal component analysis*. Springer, 2002.

[25] D. Kim and B.J. Yum. Collaborative filtering based on iterative principal component analysis. *Expert Systems with Applications*, 28(4):823–830, 2005.

[26] J. Konstan, S. McNee, C.N. Ziegler, R. Torres, N. Kapoor, and J. Riedl. Lessons on Applying Automated Recommender Systems to Information-Seeking Tasks. *Proceedings of the Twenty-First National Conference on Artificial Intelligence*, 2006.

[27] Shyong K. Lam and John Riedl. Shilling recommender systems for fun and profit. In *WWW '04: Proceedings of the 13th international conference on World Wide Web*, pages 393–402, New York, NY, USA, 2004. ACM.

[28] D. Lemire. Scale and translation invariant collaborative filtering systems. *Information Retrieval*, 8(1):129–150, 2005.

[29] G. Linden, B. Smith, and J. York. Amazon. com Recommendations: Item-to-Item Collaborative Filtering. *IEEE INTERNET COMPUTING*, pages 76–80, 2003.

[30] Paolo Massa and Paolo Avesani. Trust-aware recommender systems. In *RecSys '07: Proceedings of the 2007 ACM conference on Recommender systems*, pages 17–24, New York, NY, USA, 2007. ACM.

[31] L. McGinty and B. Smyth. Comparison-Based Recommendation. *Proceedings of the 6th European Conference on Advances in Case-Based Reasoning*, pages 575–589, 2002.

[32] B. Mobasher, R. Burke, R. Bhaumik, and C. Williams. Effective attack models for shilling item-based collaborative filtering systems. In *Proc. of WebKDD*, pages 21–24, 2005.

[33] B. Mobasher, R. Burke, R. Bhaumik, and C. Williams. Toward trustworthy recommender systems: An analysis of attack models and algorithm robustness. 2007.

[34] B. Mobasher, R. Burke, and J.J. Sandvig. Model-Based Collaborative Filtering as a Defense against Profile Injection Attacks. In *Proceedings of the National Conference on Artificial Intelligence*, volume 21, page 1388. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006.

[35] Bamshad Mobasher, Robin Burke, Runa Bhaumik, and Chad Williams. Toward trustworthy recommender systems: An analysis of attack models and algorithm robustness. *ACM Trans. Internet Technol.*, 7(4):23, 2007.

[36] D.L.F. Nilsen. *International Journal of Humor Research*. Walter de Gruyter Inc., New York.

[37] B. Pang and L. Lee. Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2):1–135, 2008.

[38] S.T. Park, D. Pennock, O. Madani, N. Good, and D. DeCoste. Naïve filterbots for robust cold-start recommendations. *Proc. of the 12th ACM SIGKDD Int'l Conf.*, pages 699–705, 2006.

[39] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical magazine*, 2(6):559–572, 1901.

[40] D.M. Pennock, E. Horvitz, and C.L. Giles. Social choice theory and recommender systems: Analysis of the axiomatic foundations of collaborative filtering. *Proceedings of the Seventeenth National Conference on Artificial Intelligence*, pages 729–734, 2000.

[41] D.M. Pennock, E. Horvitz, S. Lawrence, and C.L. Giles. Collaborative filtering by personality diagnosis: a hybrid memory and model-based approach. *Proc. of the 16th Conf. on Uncertainty in Artificial Intelligence*, pages 473–480, 2000.

[42] J. Pitman. *Probability*. Springer, 1993.

[43] T.K. Quan, I. Fuyuki, and H. Shinichi. Improving accuracy of recommender system by clustering items based on stability of user similarity. *IAWTIC'2006 Proc.*, pages 61–61, 2006.

[44] A.M. Rashid, I. Albert, D. Cosley, S.K. Lam, S.M. McNee, J.A. Konstan, and J.T. Riedl. Getting to know you: learning new user preferences in recommender systems. *Proc. of the 7th Int'l Conf. on Intelligent User Interfaces*, pages 127–134, 2002.

[45] A.M. Rashid, S.K. Lam, G. Karypis, and J.T. Riedl. ClustKNN: a highly scalable hybrid model-& memory-based CF algorithm. *Proc. of WebKDD 2006*, 2006.

[46] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In *In Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*, 1994.

[47] Paul Resnick and Rahul Sami. The information cost of manipulation-resistance in recommender systems. In *RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*, pages 147–154, New York, NY, USA, 2008. ACM.

[48] B. Sarwar, G. Karypis, J.A. Konstan, and J. Reidl. Item-based collaborative filtering recommendation algorithms. *Proc. of the 10th Int'l World Wide Web Conf.*, pages 285–295, 2001.

[49] B. Satzger, M. Endres, and W. Kießling. A Preference-Based Recommender System. *E-Commerce and Web Technologies. Springer Berlin/Heidelberg*, 2006.

[50] A.I. Schein, A. Popescul, L.H. Ungar, and D.M. Pennock. Methods and metrics for cold-start recommendations. *Proc. of the 25th Annual Int'l ACM SIGIR Conf.*, pages 253–260, 2002.

[51] J. Thornton. Collaborative filtering research papers.

[52] Paolo Viappiani, Pearl Pu, and Boi Faltings. Conversational recommenders with adaptive suggestions. In *RecSys '07: Proceedings of the 2007 ACM conference on Recommender systems*, pages 89–96, New York, NY, USA, 2007. ACM.

[53] F.B. Viégas and J. Donath. Social network visualization: Can we go beyond the graph. In *Workshop on Social Networks, CSCW*, volume 4, pages 6–10, 2004.

[54] E.G. Vozalis and K.G. Margaritis. Recommender systems: an experimental comparison of two filtering algorithms. *Proc. of the 9th Panhellenic Conf. in Informatics*, 2003.

[55] M. Vozalis and K. Margaritis. On the combination of user-based and item-based collaborative filtering. *International Journal of Computer Mathematics*, 81(9):1077–1096, 2004.

[56] J. Wang, A.P. de Vries, and M.J.T. Reinders. Unifying user-based and item-based collaborative filtering approaches by similarity fusion. *Proc. of the 29th Annual Int'l ACM SIGIR Conf.*, pages 501–508, 2006.

[57] J. Weng, Y. Zhang, and W.S. Hwang. Candid covariance-free incremental principal component analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):1034–1040, 2003.

[58] DC Wilson, B. Smyth, and DO Sullivan. Sparsity reduction in collaborative recommendation: a case-based approach. *Int'l Journal of Pattern Recognition and Artificial Intelligence*, 17(5):863–884, 2003.

[59] Hu Wu, Yongji Wang, and Xiang Cheng. Incremental probabilistic latent semantic analysis for automatic question recommendation. In *RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*, pages 99–106, New York, NY, USA, 2008. ACM.

[60] G.R. Xue, C. Lin, Q. Yang, W.S. Xi, H.J. Zeng, Y. Yu, and Z. Chen. Scalable collaborative filtering using cluster-based smoothing. *Proc. of the 28th Annual Int'l ACM SIGIR Conf.*, pages 114–121, 2005.

[61] Mi Zhang and Neil Hurley. Avoiding monotony: improving the diversity of recommendation lists. In *RecSys '08: Proceedings of the 2008 ACM conference on Recommender systems*, pages 123–130, New York, NY, USA, 2008. ACM.

[62] C.N. Ziegler, S.M. McNee, J.A. Konstan, and G. Lausen. Improving recommendation lists through topic diversification. *Proceedings of the 14th international conference on World Wide Web*, pages 22–32, 2005.

[63] A. Ziv. *Personality and sense of humor*. Springer Pub Co, 1984.