A Survey of Web-Based Collective Decision Making Systems

Jennifer H. Watkins[1] and Marko A. Rodriguez[2]

[1] International and Applied Technology
Los Alamos National Laboratory
Los Alamos, New Mexico 87545
jhw@lanl.gov
[2] Digital Library Research and Prototyping Team
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

Abstract. A collective decision making system uses an aggregation mechanism to combine the input of individuals to generate a decision. The decisions generated serve a variety of purposes from governance rulings to forecasts for planning. The Internet hosts a suite of collective decision making systems, some that were inconceivable before the web. In this paper, we present a taxonomy of collective decision making systems into which we place seven principal web-based tools. This taxonomy serves to elucidate the state of the art in web-based collective decision making as well as to highlight opportunities for innovation.

1 Introduction

Collective decision making is the aggregation of individuals' information to generate a global solution. There are a variety of reasons that collective decisions are sought. A collective decision may be desirable to represent the opinions of a group, as in a vote. A collective decision may be desirable to collect the best information available, as in expert elicitation. Or a collective decision may be desirable to produce a new combination of ideas held within the group, as in a brainstorm session. The resulting decision may be employed directly or used as decision support for another process. For the purposes of this paper, mechanisms that elicit decisions from a group of people are called collective decision making systems (CDMSs). This designation is used to represent a departure from group decision support systems (a subfield of computer supported collaborative work) as CDMSs are not necessarily collaborative in nature [1]. In addition, this paper refers exclusively to web-based collective decision making systems, often called social software [2]. The unifying purpose of these systems is to structure individual input in such a way as to generate a meaningful aggregate decision, even if that input is implicitly derived or from asynchronous or anonymous contributions.

The human proclivity to decide in a group is long standing. However, web-based tools for collective decision making have advanced this ability and need to a larger scale. In this article, seven types of popular web-based systems are

R. Nayak, et al. (Eds.): Evolution of WEB in Artificial Intell. Environ., SCI 130, pp. 245–279, 2008.
springerlink.com © Springer-Verlag Berlin Heidelberg 2008

discussed—document ranking, folksonomy, recommender system, vote system, wiki, open source software, and prediction market—within a taxonomy of features. The decision capabilities that determine each type of CDMS are the result of a specific combination of features. These features can be organized into a taxonomy of problem space, implementation, individual features, and collective features. Such a taxonomy serves to distinguish the context under which a particular CDMS can be used and to highlight the similarities between seemingly disparate tools. In addition, this taxonomy reveals combinations within the feature space not utilized by existing systems that could compose a new system and thus a new decision making capability.

The first half of this article describes in detail the history, purpose, and instantiation of the seven types of web-based collective decision making systems in turn. The remainder of the article presents the seven system types within a feature space organized by a taxonomical structure. The aim is to set each system in a broader context while providing a framework to aid in system design. But first, the most fundamental delineation of CDMSs will be outlined—that between the collective and the aggregation mechanism. This dual understanding of CDMSs is the first branch of the taxonomy presented shortly.

All collective decision making systems require a population of participants (i.e., a collective) and a means of aggregating their knowledge into a collective decision (i.e., an aggregation mechanism). For example, deliberation aggregates through conversation; democracy aggregates through voting; a recommender system aggregates through user footprints. The following sections describe these two components.

1.1 The Collective

Collective decision making is founded on the belief that people are not flawless decision makers. An individual is a good, but not ideal, complex problem-solver. Collective decision making utilizes a better one, namely the unit of participants. The typical account of decision making involves an expert who applies his or her knowledge to generate a solution. Through collective decision making, however, it is the collective itself that is considered the expert. The collective can be thought of as a meta-individual that possesses, generates, and decides on knowledge in much the same way an individual does. Like an expert, a collective has more knowledge than other individuals through the combination of information held by each member. Collectives are autopoietic, they have continuity in identity despite changes in membership, allowing us to think of them as persistent individuals [3]. Thus, collective decision making is distributed over numerous processes within the collective, as opposed to contained within a single decisive event.

1.2 The Aggregation Mechanism

A collective without an aggregator is no more powerful than an individual. An aggregation mechanism serves two purposes in eliciting collective decisions. One, it draws out the pertinent information of each individual in the collective. Two,

Table 1. Collective decision making systems and their common aggregation mechanisms

Collective Decision Making System	**Aggregation Mechanism**
document ranking	PageRank
folksonomy	collaborative tagging
recommender system	collaborative filtering
vote system	plurality
open source software	collaborative development
wiki	collaborative editing
prediction market	market scoring rule

it combines that information in such a way as to make it useful. Every CDMS has a variety of web-based aggregation mechanisms. For example, vote systems may employ approval voting, Borda count, or plurality voting. Table 1 lists the CDMSs discussed in this article and their common aggregation mechanisms.

2 Web-Based CDMS

We focus exclusively on web-based collective decision making systems, as opposed to, for example, face-to-face decision making. Web-based decision support systems are not only computer mediated but are made powerful through the vast population of individuals that use the Internet. These individuals are utilized by the aggregation mechanisms, either through tracking the combined behavior of many or through scouting for expertise. This online collective provides two potential benefits. One, such a large, dispersed population captures statistical collective intelligence or the generation of knowledge through the weighted averaging of independent, individual judgments [4]. Most of the web-based systems discussed here require these large numbers of self-interested participants to generate an accurate decision. Two, some systems benefit from the ability to amplify expertise. The Condorcet Jury Theorem, from probability theory, states that if each individual in a collective is more likely than not to be correct, then as the size of the group scales, the probability of the collective decision being correct moves toward certainty [5]. Some of the systems discussed discourage participation by those who are not more likely to be correct and thus enjoy the result of this theorem.

Before developing the taxonomy further, the following sections describe the web-based collective decision making systems of interest to this article.

2.1 Document Ranking

Document ranking, the system that organizes web-pages for the purpose of document retrieval (the matching of records to queries), uses information inferred from the links between documents. The World Wide Web is a network of web-pages connected through hypertext links. A given web-page becomes embedded in the wider network of web-pages when the person publishing the site creates links to other pages and when the site receives incoming links. Its importance in the web, as determined by its location in the network, produces a ranking. Search engines utilize these rankings to retrieve the most relevant documents in response to keyword queries. Document ranking exploits the aggregate of the individual decisions to link to specific pages, interpreting the resulting network as collective informational content.

Document ranking is the method of information retrieval employed by most popular search engines. PageRank, employed by the Google[1] search engine, is perhaps the most well known of the document ranking algorithms. The PageRank algorithm considers not only the number of incoming links (indegree) to a given web-page, but also the incoming links to the originating web-pages in a recursive manner [6]. Thus, a given page can receive a high ranking through a high indegree or through a single incoming link that itself has a high ranking. The ranking must correspond well with an individual's subjective sense of importance. Search engines achieve this correspondence because the choice to link to a page contains latent human judgement about importance. The structure of the web is determined at the local level when an individual chooses to link to another web-page. Globally, this structure can be interpreted in a variety of ways to inform document retrieval.

PageRank was designed to improve the relevancy of the search results returned. It is superior to text-based ranking functions originally applied to the web that simply relied on a full text keyword search. The web is of such massive scale and amorphous organization that these traditional techniques are infeasible. The harnessing of the collective actions of web-page creators generated an improvement in information retrieval techniques. However, Google uses, as do other search engines, a proprietary combination of criteria to determine document relevancy [7].

While the PageRank algorithm is perhaps the most well-known aggregation mechanism for document ranking, there are a variety of other algorithms that utilize collective decisions for the purpose of information retrieval. For example, the HITS algorithm interprets hypertext links as "conferred authority" [8]. Instead of a single ranking metric, HITS utilizes two measures, hub and authority, along which all web-pages are scored. Due to the semantic understanding individuals encode into network structure through linking pages, a pattern emerges where pages with a high hub score densely link to authoritative, high-quality pages (those with a high authority score) on a given subject. It is this relationship that is exploited by HITS to return precise search results.

Because of the necessity of search engines to locate material on the web, a number of techniques have been developed to falsely inflate the ranking of certain pages. These techniques are known as adversarial information retrieval [9]. The Google bomb is a slang term for the coordinated linking to a particular page with a particular key-word phrase, usually for humorous or political intent. Spamdexing is a complimentary technique used to falsely inflate the ranking of a website in order to increase hits, e.g., for commercial gain. A link farm,

[1] URI: http://www.google.com

a specific type of spamdexing involves linking every page in the farm to every other page to increase the rankings of all the pages. The chances for exploitation of search engine results has precipitated search engine optimization, techniques that look to improve the traffic (both volume and quality) to a particular site by orienting it properly for both human and search engine indexing. The result is an algorithmic arms race between search engine companies that strive to maintain relevancy and the search engine optimizers.

Despite the constantly evolving nature of document ranking criteria and algorithms, the essential collective decision core remains the same: individuals create web-pages that link to other web-pages. These links can be perceived as votes of quality. In aggregate, these individual actions sum to an informational content that can be exploited for the purposes of information retrieval. The success of Google's search engine is a testament to the immense utility such unintended footprints can produce. The beauty of document ranking algorithms is that they are able to extract meaning from digitally represented human actions that were made for other purposes. This latent human intent in aggregate forms the data used for our search engines.

2.2 Folksonomy

Web-services such as Flickr[2] and Del.icio.us[3] (Del.icio.us is a domain hack for "Delicious" and will hereafter be referred to as such) allow users to label, or tag, resources with descriptive metadata such that the statistical aggregate of all tags creates a collectively designed index, or folksonomy [10]. The folksonomy is used as a tool for information retrieval connecting users to resources via tags. Tagging is the appending of metadata to a resource, most often for the purpose of description. The user tags resources for their own purposes using their own descriptions. Over time, the same resource will be tagged many times and particular tags will be used repeatedly to describe the same resource. This overlap increases the relevance of the tagged resource for retrieval by the tag as a keyword.

The aggregation of many users' tags to create a folksonomy is achieved through a mechanism referred to as collaborative tagging. Through the combination of multiple users' interpretations and thus tags of a particular resource, a folksonomy is generated that indicates the popularity of a particular term to describe a particular resource. Despite the self-interested and uncoordinated actions of the participants, analysis indicates that users' interact with the system through collaborative tagging in a patterned manner to create a coherent tool [11].

The categorization method of folksonomies is in contrast to traditional centralized methods including ontologies, controlled vocabularies, and thesauri [12]. These methods require the careful construction of a world view into which all current and future resources can be placed. Instead of these traditional indexing methods which is an expert-based and time-consuming effort, folksonomies distribute the indexing over a large population of users [13]. In essence, the tagging

[2] URI: http://www.flickr.com
[3] URI: http://del.icio.us

of objects by a single person is of less use than the formal classification of those resources by an expert. However, in aggregate the result of many individuals' tags can form a folksonomy that is more robust than traditional methods.

The folksonomy also stands in contrast to newer indexing methods that utilize computer automated crawling of resources, as utilized by search engines [13]. The human indexing provides a semantic understanding of the content of each resource that may not be captured by a web-crawler. Folksonomies more closely resemble traditional human-based classification systems in their ability to understand semantic content, but automated systems in their overhead and cost. Because individuals contribute to the folksonomy primarily for their own benefit, the classification value is merely a by-product of a well-designed system.

In practice, folksonomies are used to describe a variety of resource types, from photos to blog entries. The most common use of folksonomies is to describe the information at a particular URL so that the tagger can find the information again later. This practice is called social bookmarking and is an extension of the bookmarking feature included with most web browsers. Bookmarking began as such with the Mosaic browser. This ability required a hierarchical organization of favorite websites that can quickly become unwieldy with lax management. An increase in the speed and precision of search engines led to dynamic bookmarking where an individual simply searches for a favorite site again. This ability is augmented by social bookmarking which refers to the tagging of a web-page with descriptive metadata for ready retrieval. It is browser-independent and allows users to see how URLs were bookmarked by others and to see the bookmarks of a particular user—thus it is social.

The utility of a folksonomy depends on the duplication of tags. Social bookmarking sites provide a feedback mechanism that encourages the convergence of tags [14]. The Delicious capability to see how other users have tagged a given URL provides feedback that encourages the imitation of others' tags. Thus, early tags of a particular URL are the most popular [15]. The most common depiction of the tags for a particular resource is the tag cloud [16]. A tag cloud depicts an alphabetized list of the tags applied to a given resource. The popularity of the tag, the frequency of its use, is indicated through a relatively larger font size. The presentation of a folksonomy is of primary importance for utility as its clarity aids convergence.

2.3 Recommender System

Recommender systems track user behavior, whether implicitly or explicitly, as a means of recommending potentially interesting resources to users in the system. A ranking of the resources not yet seen by a user is produced according to some measure of the user's preferences. The purpose is to filter and organize the overabundance of resources within the system's domain. In other words, recommender systems manage information overload by acting as a search function to provide a personalized subset of the total collection [17]. As one becomes a more finely differentiated individual through interactions with the system, a more individualized filter is developed based on the interactions of other individuals. The purpose is to aid the user in discovering novel and interesting products (i.e., it is primarily a tool implemented for commercial reasons).

The need for a recommender system is based on the typical problem in computer-mediated environments of information overload. Search engines work well when one knows what one is looking for. However, there are situations when this is not the case. Here, a recommender system performs information retrieval without any keyword entry on behalf of the user. Instead, the system infers desires through past interactions with the system.

One class of aggregation mechanism for recommender systems is the suite of collaborative filtering algorithms. Collaborative filtering compares the independent decisions of many users with persistent identity to generate a similarity metric such that users are recommended products they have not accessed but those who are similar to them have [18]. In this sense, the collective decides what will be of interest to the individual. Perhaps the most common technique for establishing similarity is nearest neighbor analysis adapted from pattern recognition research [19]. An alternative content-based algorithm made popular by Amazon.com[4] develops similarity metrics on products instead of users. Products that are similar to a purchased item will be recommended to the user [20].

The persistent use of a particular recommender system is essential to both the individual and the collective as recommendations gain sophistication with more personal- and with more collective-level data. The initial paucity of information with which to infer recommendations is referred to as the cold-start problem [21]. It can become a burden to the user to populate the system with enough information so as to make an accurate recommendation. However, there are a variety of algorithms dedicated to decreasing the impact of this problem so that users will recognize the utility of the system immediately [22, 23]. In addition, a recommender system is a unique web-based application in that it can be implemented so as to work completely unseen to the user. In this effortless instantiation, entry into the collective is automatic when an individual logs in to a site. For example, Amazon.com implicitly tracks user behavior for use in the recommender system. Appropriate recommendations are then inferred from an individual's usage of the site. Other recommender systems require explicit user participation. For example, the Netflix[5] recommender systems requires that the user rate movies they have previously viewed through a simple one through five star interface.

However, the ease of entry into these systems has drawbacks. The use of implicit user tracking technology is seen by some as an invasion of privacy. A user, especially an eclectic user who rates products across many domains, can be identified through the tracking data alone, which can be distributed for use by third parties [24].

[4] URI: http://www.amazon.com
[5] URI: http://www.netflix.com

2.4 Vote System

The vote system is a time-honored means of gathering individual decisions and aggregating them into a single collective decision. As such, vote systems are the hallmark of democratic governance. The feature space of vote systems has been researched extensively resulting in a large number of aggregation mechanisms. Each mechanism specifies two components — the ballot form and the tally method. The ballot determines the way that an individual can express a decision and the tally method determines how those expressions are aggregated into a collective decision. Common aggregation mechanisms include plurality, Borda count, and approval vote. A plurality vote allows the voter to choose only one option on the ballot. A ballot for a Borda count vote allows all options to be ranked in order of preference by the voter. Points are assigned according to the rank. An approval vote allows the voter to choose as many options as are deemed preferences. For all three mechanisms the majority option wins. Note that in vote systems a majority need not imply more than half of the votes. For multi-winner votes, a tally method other than majority rule may be used, most often for proportional representation.

In addition to the plethora of aggregation mechanisms, there are multiple forms of governance within the democracy designation of which direct democracy and representative democracy may be the most familiar. Regardless of the details of the form of governance, the essential element of a vote system is that it be perceived as fair. Online or offline, vote systems are used to determine collective preference. It is this that sets vote systems apart from other decision systems, as there is no objective measure of accuracy outside of perceived fairness. It is to this end that a wide variety of aggregation mechanisms exist. Each aggregation mechanism elicits votes differently to affect different outcomes to satisfy voters.

The fair transference of individual decisions into a collective decision is studied no more rigorously than in voting systems. The determination of the best systems for aggregating preferences is an important pillar of voting theory literature. To aid in this study, a number of rules have been outlined, all of which are criterion for a fair vote [25]. However, Arrow's General Possibility Theorem proved that there can exist no rank-based vote system between three or more alternatives that will satisfy all fairness requirements [26]. Taken loosely, this implies that in any system individual preferences will fail to aggregate into collective preferences. The result of this theorem is that it is necessary to specify which fairness rules must be met and which can be violated before a vote system is implemented.

While academic interest in vote systems has a long and rich history, the introduction of vote systems to the web is in its infancy. Fields of study such as e-democracy and e-government are increasing the interest in implementing online vote systems. Recognizing the unprecedented potential of the web to facilitate communication on a large and distributed scale, e-democracy embraces the notion of wisdom through collective decisions. This sub-discipline is interested in developing web-based tools that support democratic processes to improve the development of policy [27]. In addition, there are a number of studies proposing the use of a comprehensive system that aids the voter in acquiring information about the candidates, making a decision about the best candidate, and then casting that vote [28, 29]. These studies are designed to facilitate the wider aim of political participation.

E-democracy has met with security and reliability challenges in the development of web-based electronic voting (the casting and/or tallying of votes via the Internet) where it is hoped that more members of society will vote than do in traditional elections. Challenges to developing a secure and reliable system include protecting the secrecy of the vote for each voter, network vulnerabilities, and the appropriate implementation of cryptographic techniques [30]. Nevertheless, direct-recording electronic (DRE) voting machines are used in US elections without paper-based, voter-verfiable copies. While electronic poll site voting and kiosk voting, which are supervised by election officials pose security concerns, the concerns deepen for remote voting via the Internet. It is difficult to guarantee in such situations that the voter is who they say, is not being coerced to vote in a particular manner, or is not selling their vote [31].

Despite the interest in moving political elections online, current online vote systems remain confined almost entirely to polling interfaces where no actual decision is affected. There are some interesting exceptions. Estonia became the first country to implement an online electronic voting system for a national election in 2007, where 30,275 Estonians voted through the system.[6] After the Pentagon ruled the system insufficiently secure to implement for soldiers living overseas, the 2004 Michigan democratic caucus elections had the option of using an electronic voting system to register votes.

While the most obvious application of voting systems is to the execution of political government, a vote system need not be limited by this restriction. With the ease of communication and tally functions, web-based voting systems have the possibility to establish new algorithms and methods for producing fair collective decisions. Smartocracy,[7] a social software voting site, utilizes an online social network to spread voting power to those the voter trusts as proxy, not those elected to represent him or her [32].

2.5 Wiki

A wiki is a highly distributed way to gather, create, and share knowledge. A wiki is server software that allows users to freely create and edit web-page content using any web browser [33]. The purpose of a wiki is to capture the collective knowledge held by participants such that the resulting documents transcend the abilities of individual contributors. In a wiki system, any individual can use a simple markup language to create pages and link them to other internal pages. These pages allow content and organizational contributions and edits at any time. Every facet of a wiki web-page, the content, its organization, the links,

[6] "Estonia scores world first with web poll," The Age, March 1, 2007 `http://www.theage.com.au/news/web/estonia-scores-world-first-with-web-poll/2007/03/01/1172338771317.html`

[7] URI: http://smartocracy.net

even its very existence, is alterable by any member of the collective, regardless of original authorship. The result is a network of collaboratively generated documents that contains the authorial wisdom of all its contributors.

A wiki aggregates decisions through the mechanism of collaborative editing. Collaborative editing simply refers to the ability to alter and contribute freely regardless of original authorship. Here a collective produces a document through the melding of asynchronous and independently made individual decisions. Collaborative editing falls under the research of computer-supported cooperative work, a sub-discipline for which wikis are merely one solution [34]. In the past, this sort of group work has required small groups that interact face-to-face. However, through wikis the creation of content takes place on a large-scale and has become a distributed process that often involves strangers. Wikipedia[8] is an online encyclopedia that uses wiki software to allow anyone to contribute. It is a multilingual collaboration to capture the collective's knowledge in encyclopedic format. There are articles written in over 200 languages[9] with the largest collection, the English language version, approaching its 2 millionth article at an exponential rate [35]. Wikipedia has become the de facto source of encyclopedic information on the web with over 50 million queries a day. It is powered using MediaWiki, open source software distributed by the Wikimedia Foundation that is used to host numerous other wikis. Thus, many of the features in Wikipedia are also standard features in other wikis. Despite the predominance of Wikipedia, the wiki system should not be conflated with its most popular example. Wikis are knowledge management and content creation tools that serve projects at multiple scales both public and internal to institutions [36].

Wikipedia, like other social software encourages a sense of community amongst members of the collective. Like most communities, that of Wikipedia divides labor both explicitly through levels of permissions and organically through users identifying a given task and choosing to complete it. This bare bones infrastructure supports over 200,000 edits per day. The community is also governed by a number of rules and guidelines for participation, mostly guiding the type of content that is appropriate. To handle debate and disputes there are talk pages for articles kept separate from the encyclopedic content.

However, the scale and decentralization of Wikipedia leads to inaccuracies, sabotage, vandalism, and exploitation. In order to maintain the open nature of the project, most of these problems are handled by the vigilance of other Wikipedians, a collective-based solution. Other problems require a change in the structure of the wiki itself, an aggregation mechanism solution. For example, some contributors add erroneous links in Wikipedia to another site so as to falsely inflate that site's PageRank through the prestige of Wikipedia. To combat this spam problem, Wikipedia software applies a NOFOLLOW rule to every link on the site. Essentially, this prohibits web-crawling robots from following the links on Wikipedia and inflating the adjacent sites' PageRanks. Despite indicators that Wikipedia would be overcome by malicious and unintentionally poor content, it

[8] URI: http://en.wikipedia.org
[9] statistics from http://meta.wikipedia.org/wiki/Statistics accessed July 14, 2007.

remains a viable source of information on the web. In fact, due to the success of the Wikipedia paradigm, it is only one of a multitude of wiki-based projects hosted by the Wikimedia Foundation.

The use of wikis dramatically changes the ownership practices involved in publishing. Thus, the Wikimedia Foundation embraces the copyright licenses developed to account for the new business practices associated with the web. Creative Commons[10] has developed a number of licensing options that differ from traditional copyright in that only some proprietary rights are reserved. The goal is to protect the rights in which a given producer is interested while making access to the work as available as possible. All written material on Wikipedia is under a creative commons license. In addition, Wikimedia contributors have the option to choose the degree of copyright under which their work is protected.

2.6 Open Source Software

Open source software is an online collaborative development method employed to create computer software [37]. The name refers to the free availability of the source code that composes a piece of software. A programmer can contribute code to, alter, or delete the source code to produce changes in the software. Through the democratic inclusion of a large collective of interested participants, innovative solutions are contributed and bugs in the code are efficiently found and fixed.

As with wikis, open source software aggregates through the process of collaborative editing [38]. Here participants asynchronously construct, maintain, and improve upon a software project. However, code is a precise and complex dependency-based representation that can be fragile to change. Unlike text that can maintain coherency through various perturbations, code can be "broken". Thus, open source projects use a number of strictures and conventions to maintain the integrity of the existing product. One of the most fundamental of these is versioning whereby once the software reaches a certain state it is named and delineated from other code. Software as developed through a cyclical process of writing and testing and versioning allows the use of revision control to track incremental changes. Other practices include the use of a hierarchical permissions structure whereby only a subset of the total number of programmers can move code to a testing phase, incorporate changes into new versions, and declare a version ready for release.

The Linux operating system[11] is a prototypical example of open source software. This operating system was originally written by Linus Torvalds in 1991 as a processor-independent version of Unix, a proprietary program. Torvalds successfully received help to development problems he posted to a programming-oriented list-serv. This spurred him to provide the entire source code to the programming community for further contributions. The community proved to be a willing and productive collective for software development. Since Linux, the

[10] URI: http://creativecommons.org
[11] URI: http://www.linux.org

development of software through open source techniques has blossomed. SourceForge[12] is a repository of open source software available for use, improvement, and critique. It hosts over 100,000 projects and over one million registered users, larger than any other resource of its kind.

Open source software has become a successful CDMS on a number of counts. It is lauded for its low cost, flexibility, reliability, and robustness when compared with proprietary equivalents. This is typified in that companies, perhaps most notably the formerly staunch licensor IBM, that rely on proprietary software for revenue are now incorporating open source business models [39]. Open development is also an exemplar of a new online and collaborative political economy that can fuel innovation [40]. For example, the LAMP quad is the powerhouse behind innumerable online services. This stack of technologies includes an operating system, server software, a database program, and a scripting language interpreter. LAMP stands for Linux, Apache, MySQL, and Perl (alternatively PHP or Python)—all open source software available for free. In addition to providing the tools online applications need to innovate, the open source approach is spreading to other arenas as well, notably to educational materials [41].

The licensing of software for free distribution and modification of the source code is an essential component of the open source paradigm. The Free Software Foundation initiated the GNU General Public License (GPL) for the free modification and distribution of software. This collection of licenses is sometimes referred to as copyleft, a play on the term copyright, as it grants rights to the user as opposed to reserving rights of the producer. The GNU GPL is referred to pejoratively as "viral" as all subsequent uses of the code must also be under a GPL license. In addition, the Open Source Initiative[13] maintains the integrity of the term open source through an industry-accepted definition. The definition allows the commercial use of open source software for those individuals and companies that wish to harness the economics of open source [42].

2.7 The Prediction Market

The prediction market is a forecasting tool where the pertinent information held by each trader is revealed and aggregated through a game-like exchange [43]. As in traditional financial markets, this exchange refers to the buying and selling of contracts (stocks) by participants who choose the price at which they are willing to trade. The contracts in a prediction market represent not shares in a company, but forecasts of specific event outcomes (such as the winner of a political election) and their price reflects the probability that the outcome will take place [44]. The market value of each contract fluctuates according the price at which traders buy and sell it. Thus, a market value is the collective's estimate of the probability of a future event. Contracts in a given event are sold with a value between 0 and 100 (interpretable as a probability). Thus, a market value of 100 for a given contract suggests the event is certain to happen. As in traditional

[12] URI: http://sourceforge.net
[13] URI: http://www.opensource.org

markets, traders who buy low and sell high earn the difference in prices, while those who sell low and buy high lose the difference. Unlike traditional financial markets, traders may also earn money by owning contracts whose forecast of the event outcome was correct. In the most simple payout scheme, this contract would have a value of 100 and the value of contracts in any other outcome would drop to zero.

There are a number of aggregation mechanisms for prediction markets. While in all of them traders leverage their privately held knowledge competitively to out-predict others, the manner in which trades are elicited and affect prices differ. Perhaps the most familiar mechanism of aggregation in traditional financial markets is continuous double auction (CDA) [45]. Here bids (the price for which a trader is willing to buy a given contract) are matched to asks (the price for which another trader is willing to sell the given contract) to complete a trade and update the market price. However, unlike traditional markets, prediction markets have a terminus, usually just before the forecasted event is to take place. This feature encourages other aggregation mechanisms that would not be feasible in traditional markets. A notable option is Hanson's suite of market scoring rules that allow a trader to update the market price at any time without engaging another trader in an auction. The trader can be thought of as updating the market price by compensating the trader whose price was replaced [46]. Another alternative is Pennock's dynamic pari-mutuel market adapting features from pari-mutuel betting into a CDA [47].

Prediction markets have garnered attention for their ability to accurately predict the future. They perform as well or better than traditional prediction techniques such as polling. For example, the Iowa Electronic Markets[14] (IEM) in the 2004 presidential election correctly predicted the number of electoral votes by which Bush would win [48]. The IEM out-predicts polls 75% of the time. The Hollywood Stock Exchange[15] (HSX) in 2007 correctly identified seven out of the eight winners in the most popular Oscar categories as they did in 2006.[16] In 2005, all eight winners were predicted correctly.

Prediction markets are powerful decision tools because they generate an exact probability for each forecast and that probability varies through time as more information is revealed. In 2001, the Defense Advanced Research Projects Agency (DARPA) funded two grants in electronic market-based decision support that came to be known as the Policy Analysis Market. The purpose of the grants was to develop a system that could capture the forecasting accuracy of prediction markets for problems of governmental interest. Problems included political instability, how US policy would affect the instability, and how instability would affect US interests [49]. The area of focus was the Middle East. Intended for public use, these markets would have explored how a combination of events would lead to a future event. However, the project was cancelled before large-scale human subjects testing took place, so little data is available. The markets were

[14] URI: http://www.biz.uiowa.edu/iem/
[15] URI: http://www.hsx.com
[16] HSX Press Release http://www.hsx.com/about/press/070226.htm

closed due to political outrage over the use of gambling for devastating events. The markets were later instantiated by a private company.

Prediction markets, despite their impressive success stories, are not a panacea for all future uncertainties. Typically, prediction markets have only been successful on topics that are of interest to a large number of people (e.g., politics, sports, Hollywood) and thus have a large pool of possible traders from which to draw. A large trading population is important to draw out accurate information by making it worthwhile for informed traders to participate. To some extent the effect of thin participation can be mitigated by specially designed market algorithms, such as market scoring rules, that encourage trade in such a situation [50]. In addition, legality remains a problem for prediction markets in the United States. A prediction market can be construed as a form of gambling where traders place real-money bets on essentially valueless contracts. To avoid legal entanglement, many prediction markets use play-money (and performance rankings) instead of real-money. While money is a universal motivator, it has been found that play-money markets can provide performance comparable to real-money counterparts [51].

A taxonomy of features necessary to define these systems will be detailed next.

3 Taxonomy of Collective Decision Making Systems

All collective decision making systems can be placed within a taxonomy of features that both distinguish systems from each other and that highlight system similarities. The presented taxonomy of CDMSs is organized into four primary classes—problem space, implementation, individual features, and collective features. Each of the seven web-based CDMSs maintain unique signatures within the feature space circumscribed by the taxonomy. The features of each system will be described in the sections that follow. The virtue of the taxonomy is to make apparent places for innovation. While many systems have common features, it is the sum of features that makes a system unique. Essentially, a change in one feature could potentially create a unique system with benefits and drawbacks dissimilar from the original system, thus filling a new niche. Note that Table 2 summarizes what follows in tabular form and that Fig. 1 displays the information as a dendrogram.

3.1 Problem Space

Every collective decision making system is designed to generate a decision for a particular type of problem. This class characterizes that problem space.

Decision Type. Decision type is a primary delineation for classifying decision making systems. While, for example, folksonomies and document rankings do not appear to be similar systems, the purpose of both as defined by decision type is information retrieval. The seven collective decision making systems provide decisions for only four types of problems: information retrieval, governance, content creation, and forecast. None of the four decision types inherently require a collective; however, through a well-designed system the power of a multitude

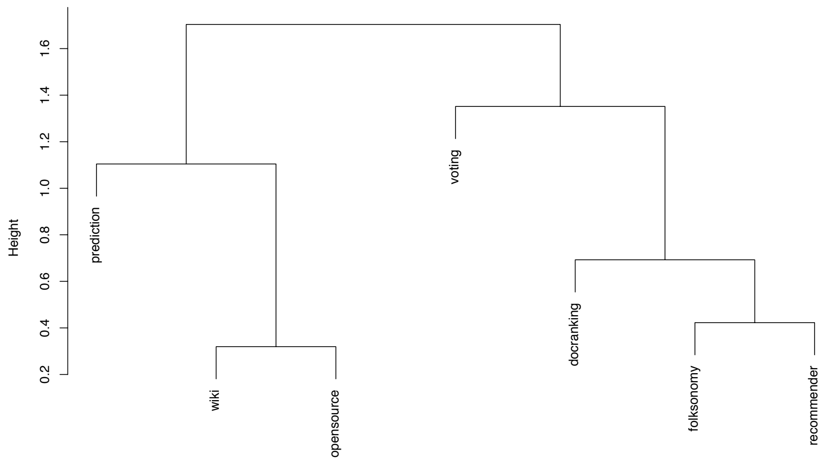

Fig. 1. Taxonomy-based comparison of CDMSs

of thinkers can be harnessed to produce powerful results. These four decision types do not completely fill the problem space, nor do the seven systems yield the complete set of systems that can generate these decision types.

Information retrieval is an interdisciplinary area of research encompassing the science behind the search for resources, whether text, documents, or records in a database. A primary aim is to control information overload, a common occurrence on the web. One way to manage the immense amount of information available online is to rank options via a pertinent algorithm to provide a list of search results. Search engines, such as Google, employ a variety of document ranking algorithms to this end; however, search engines are just one example of information retrieval on the web. Like search engines, folksonomies also follow a query-resource format where keyword queries connect users to applicable resources. In addition, folksonomies aggregate resources at the user level. Who tagged a resource can be as important as the tag itself as one resource in common suggests the possibility of the discovery of additional interesting resources. For example, CiteULike,[17] a social bookmarking site for academic papers, organizes a user's favorite papers into a personal library that any other user can peruse. Thus, every user's library serves as that user's bookmarks as well as an impersonal recommendation list for other users who have liked one or more resources in that library. To facilitate consumerism, most recommender systems use information retrieval techniques to anticipate the resource desires of a user. This anticipation is also an attempt to control information overload through the ranking of products.

Governance, a fond topic among political scientists and philosophers, is the administration of power over a population. This includes both the allocation of decision making rights and the aggregation of those decisions and is therefore

[17] URI: http://www.citeulike.org

a prime application for collective decision making techniques. Since its advent, democratic rule has been executed through the vote. While a multitude of implemented and theoretical voting derivations exist, the form always follows the casting of a vote by a specified population during a predetermined time followed by a tally. The indication of one's wishes through this general form is one of the most fundamental collective decision makings systems. The vote is different from a poll in that at the conclusion of a poll no decision is reached, and thus no governance takes place.

Content creation is a self explanatory term to refer to all works created. While teamwork resulted in content long before the advent of the Internet, web-based systems such as the wiki and open source software has enabled distributed collaboration across time and geography. Anonymous and asynchronous collaboration is the norm online. In addition, this collaboration is taking place on a massive scale; at the time of this writing Wikipedia reports more than 4.8 million registered contributors.

Prediction or forecasting is the estimation of the state of future events. The generation of formal predictions to minimize risk has historically been entrusted to haruspices, augurs, oracles, chartists, and other prophecy experts. However, online systems with the power to aggregate the opinions of many individuals uses collective decision making to reveal what is not readily apparent to individuals. Prediction markets are clearly forecasting tools. Recommender systems, to a lesser extent, also involve prediction, as the system attempts to anticipate the desires of individuals. However, until these systems routinely generate accurate recommendations not attributable to the mere power of suggestion as well as surprising recommendations, they will be excluded from the prediction categorization.

Decision Principle. The decision principle of a system refers to the manner in which one decision is chosen over another, regardless of the algorithm implemented. This is a primary distinguishing factor between systems. The decision principle may differ within a common decision type, thus it is a prime place for innovation. The application of a new decision principle for a decision type could yield a more effective CDMS.

All three information retrieval systems utilize a different decision principle. Document ranking requires the graph theoretic principle of centrality. Centrality measures the importance of a web-page relative to its position in the network. There are multiple measures of centrality and thus multiple algorithms for its determination.

Folksonomies develop through a measurement of frequency. The frequency of a given tag for a particular item represents how well that item is described by the tag. The tag cloud, a means of displaying the relative frequency of words, is often substituted for a ranked list. In addition to indexing items, the frequency of the use of a tag throughout the entire system, as opposed to for a single item, produces a zeitgeist of the user community at a given time. Many social bookmarking sites show a tag cloud of the most popular recent tags aggregating frequency overall users and all URLs.

Recommender systems that utilize collective decision making operate through the decision principle of similarity. Similarity metrics are used to determine the amount of coherence between two people or two items. Specifically, many recommender systems use matrix similarity. Through collaborative filtering, a similarity metric is determined between pairs of individuals in the system. Note that there are many ways to instantiate similarity metrics within the collaborative filtering paradigm.

Like folksonomies, vote systems aggregate through frequency, or a tally of the votes. Just as the higher frequency tags in a folksonomy suggest its popularity, so does the frequency of votes for a given candidate. In political elections, each vote is unweighted; however, online vote systems can implement a weighting system where it is not just the vote that is considered but the context of the vote as well in the form of a weighting [52].

Consensus plays a role in content creation systems as the content remains stable only as long as all participants individually believe it is satisfactory. This definition is more subtle than that of face-to-face meetings where all members of the collective explicitly give their assent. Here, the members may not have seen the most recent document so their assent is implicit in their not having looked at or changed it. Also, a member may join the collective at any time and alter the stable version. In order to keep contributors informed of changes, Wikipedia implements a "watch list" feature on which users can add pages in which they are interested. These pages are monitored automatically and alert the user when changes are made. The two-minute correction time for some types of vandalism in Wikipedia is attributable to the number of people looking out for that page at any given time [53]. Simple features like the watch list help CDMSs to perform optimally.

Trade, the decision principle powering prediction markets, is the most formalized instantiation of the consensus principle. In markets, the traders independently choose when and in what to participate. There are none but the most basic rules to guide trader behavior. Adam Smith's "invisible hand" is a metaphor for decentralization where markets are driven by the forces of supply and demand. In other words, self-interested individuals in a market produce global effects reflected in the prices of contracts. Prediction markets elicit the foreknowledge of individuals as it develops over time and weights and aggregates it. Here, a stable market value suggests that all participants individually believe that the valuation is correct. Thus, the last contribution (the last trade) stands as the current valuation.

Goal. The goal of a decision support system refers to the decision output that will be produced if the system is performing optimally. It is through the statement of a goal that a system's performance can be evaluated. The goal of a system is directly tied to the decision type (or purpose) of the system thus, as there were four decision types, there are four goals we will discuss.

The goal of all information retrieval systems is to perform a quality retrieval of the available and pertinent information. A vote system is unique in that the ultimate goal is a subjective feeling of satisfaction. While in spirit a vote system may be charged with producing a solution that, for example, maximizes public utility, the actual goal is the satisfaction of the population as to its execution. Most vote systems have no recourse for poor alignment between votes and utility, but do have recourse in the form of recounts for tallies that were perceived as unfair. The goal of content creation systems is also a subjective quality best categorized as utility. If the system can generate useful documents or source code then it has reached its goal. Often this utility is tied to comparisons with commercial counterparts. Thus, the accuracy of Wikipedia is compared to Encyclopaedia Britannica and the speed and cost of production of open source versus proprietary software is debated. The goal of a prediction market is to generate an accurate prediction about the state of the future.

Accuracy Metric. All decision making systems can be evaluated and improved if there is a metric by which to judge their accuracy. Again, accuracy is tied to the other features of problem space and thus only four metrics will be discussed.

Information retrieval systems are typically evaluated according to precision and recall. Precision measures the ratio of relevant results to the total number of retrieved results. High recall means that the retrieved results are a comprehensive sample of the relevant results available in the collection. Both metrics are necessary to describe a good information retrieval system as achieving high recall simply by retrieving all documents is not useful in reducing information overload. However, precision and recall are inversely proportional, thus with an improvement in one comes a decline in the other [54]. As search engine users tend to review only the first score of results, document ranking values precision over recall. Although recall is difficult to test in search engines as the space of all relevant documents for a given query is not obvious, numerous studies have compared search engines, and thus various document ranking techniques, in terms of precision and recall [55, 56].

Folksonomies use keyword matching to connect a queried tag to resources that have been labeled with that tag. This is information retrieval and thus precision and recall are important metrics. However, folksonomies serve as an alternative to professionally generated indices and are thus rated for accuracy by comparison. A major criticism of folksonomies is that imprecision and inconsistency in the use of tags produces an index that lacks rigor [57]. This is certainly something that a controlled vocabulary accounts for. However, the robustness of the system to change exceeds that of traditional taxonomies. For example, the Dewey Decimal System is an often used example of a system that has ossified due to the crystallization of the predominant worldview at the time of its inception. Its development by one man, and thus one perspective, stands in stark contrast to the fuzzy categorizations that develop through the myriad contributions of a folksonomy's collective. Thus, the conclusion of one system's superiority over another's is counterproductive as the strengths of one system are the weaknesses of the other. Each system type, whether folksonomy or traditional taxonomy, has an appropriate application based on the given problem and constraints.

Research on collaborative filtering algorithms is well-established. Their importance for commercial applications has made their refinement a priority. To illustrate the potential benefit of such technology to online business, Netflix is offering a million dollar reward for improving upon their current collaborative filtering algorithm by ten percent. It appears, however, that there may be a maximal accuracy bound due to the vagaries of individual ratings [58]. In addition, system improvement is a multi-faceted problem that extends beyond accuracy metrics [59]. The goal is to create a system that reduces the onus of participation on the user while providing unexpected recommendations, known as serendipity in the literature. This notion of serendipity complicates the evaluation of recommender systems. There are a variety of features on which accuracy metrics can be tabulated easily. The most common method is the leave-n-out method whereby a portion of the dataset is removed from view of the system to determine if the system will recommend the missing data. However, this method fails to account for user satisfaction. Users are often more pleased to be recommended novel products than accurate but obvious ones. The disjoint between user satisfaction and accuracy precludes the dismissal of systems that produce serendipitous recommendations but fail accuracy metrics. An additional mechanism to aid user satisfaction is the ability to succinctly explain why a particular recommendation was made to the user so as to decrease user skepticism [58].

The goal of a vote system is more subjective than other systems; their accuracy metric is perceived fairness. Although not a rigorous metric, there is a rich literature centered around the fairness of a particular algorithm for a vote. These arguments center on the desired outcome for the vote based on a number of factors. For example, [60, 61] explore the utility of majority rule compared to other electoral systems while [62] explores the superiority of the Borda count aggregation mechanism. Although, this is a well studied field of social choice theory, there is not yet consensus on which type of algorithm is best for a given situation. Indeed, it depends highly on historical process and public opinion.

Systems that develop content are held to the quality standards of each individual. For example, Wikipedia supports the recent changes patrol. These contributors use the watchlist to review edits to entries to maintain quality and monitor vandalism. In addition, tags can be added to the top of entries to indicate that a dispute needs to be resolved. As with folksonomies, the accuracy of documents and software can be compared to their proprietary counterparts. The journal *Nature* conducted an inquiry into the accuracy of Wikipedia compared with a resource generated in the traditional model—the Encyclopaedia Britannica. Of the science-oriented articles studied, Wikipedia contained 162 errors and Britannica 123 [63]. While this study has been formally and vehemently opposed by Britannica for not taking into account the nature of the errors, the real discrepancy is that the two formats excel under different constraints. Wikipedia is not limited by size as the hosting of a website is significantly cheaper than printed copies. In addition, the army of contributors to Wikipedia are required less for their authority than for their robust response to new information and the diversity of information they possess.

There is no objective utility function to rate the decision output of an open source software system. Any system that satisfies its contributors is a success.

Open source software is an excellent example of the power of collective decision making in that the distributed nature of the collaborative development process exceeds that of commercially developed software in a number of ways. Open source software is in some cases preferred to commercial software when compared by cost of production, time to release, and quality [64].

The vanguards of prediction market research, the Iowa Electronic Markets, demonstrate the superior predictive abilities of their markets to polling organizations using standard error of forecast in their accuracy analysis [65]. This compares the ability of a statistically representative sample to the self-selecting traders. An alternative compares the estimates of experts to that of the market [44]. Businesses are beginning to augment traditional methods of prediction, representative samples and expert elicitation, with prediction markets. Internal prediction markets aid corporations in gathering information that is not typically expressed by the corporate hierarchy and face-to-face meetings. Hewlett-Packard, Google, Yahoo!, Microsoft, and Intel all experiment with prediction markets. The results are mixed but encouraging. Hewlett-Packard designed a proprietary form of market called BRAIN that weights the trades of employees based on past trade successes. They report that price estimates went from 4% error using traditional methods to a 2.5% error with prediction markets using much less time and effort [66].

3.2 Implementation

Implementation refers to the characteristics prescribed by the problem space and system design. This class outlines the specialized skills required of the collective to participate. It is of particular utility when describing web-based systems.

Solution Space. The solution space is the set of all solutions that could be chosen as a decision. As with the problem space class, solution space is a primary defining characteristic of collective decision making systems. Here we discuss the four solution spaces applicable to each decision type.

Information retrieval systems are limited only by the total number of relevant and irrelevant results available in the system—the system's collection. For search engines like Google, the collection encompasses all of the artifacts on the web that are linked to other artifacts. The indexable web was assessed in 2005 at 11.5 billion pages, 8 billion of which Google had indexed [67]. Folksonomies are also concerned with the tagging of this collection. The Netflix collection utilized by their recommender system includes over 81,000 movie titles. It is precisely for these massive solution spaces that information retrieval systems are engineered.

Originally, ballots were blank papers used to write-in the names of candidates running for political office. Most often in vote systems today ballots restrict the solution space by pre-specifying the options. Ballots became necessary to specify the precise option that received the vote, as, for example, multiple people with the same name could claim a written-in vote. Therefore, a pre-printed ballot is used to designate all available options. An additional feature of the pre-printed ballot is secrecy, as eliminating handwriting deters connecting a voter to their vote. Thus, the Australian ballot specifies not only privacy to vote, but a pre-printed ballot as well.

The solution space for content creation documents is not limited in any functional sense. Wikipedia entries are not limited in length the way many offline resources are as the cost is less than in printed counterparts. In addition, entries are not limited to text. The Wikimedia Commons[18] is a database of freely distributable images, sound bytes, and video clips that are combinable with Wikipedia articles to enrich the entries. Open source software is similarly free from length constraints in the solution space where fast, efficient algorithms are the primary concern. However, the need for the compatibility of software does serve to define a solution space for some applications. Perhaps the largest constraint on content creation systems is that the content must be original, un-copyrighted, or properly attributed to be legal.

Prediction markets require the most rigidly defined solution space. A market must have a disjoint set of contracts where the fulfillment of one contract necessarily negates the fulfillment of the others. In addition, the contracts must exhaust the solution space. Every possible future outcome must be accounted for. For example, it is common in election-based prediction markets to see a question with two contracts—1) a Republican wins the election 2) another party wins the election. By not naming the Democratic Party in the second contract, the possibility of a win by a third party is left open and thus it covers the solution space.

A prediction market must be built around a question that has an objective answer once the contracts expire and this answer must clearly refer to a single contract. Otherwise, the results will be nullified and traders will require compensation. For example, TradeSports[19] encountered controversy when the outcome of their North Korea Missile market failed to completely satisfy either outcome. While a test missile was launched in accordance with the prediction of one contract, the launch was not verified by the Department of Defense which suggested that the other contract was more accurate.[20]

Interface Complexity. Interface complexity is particularly important for understanding the population from which decisions are originating and applies most acutely to web-based collective decision making systems. The interface complexity is not restrictive if only standard computer skills are necessary. However, some systems require special skills or a unique context to operate and thus restrict some potential members of the collective from participating. While most online decision systems do not require a specific representative population as do statistical polls, it is worth considering the segments of the population that are excluded due to the demands of the interface.

Interfaces that are not restrictive require only standard computing techniques and web navigation skills. Folksonomy interfaces are not restrictive as they

[18] URI: http://commons.wikimedia.org

[19] URI: http://www.tradesports.com

[20] TradeSports Press Release
http://www.tradesports.com/aav2/news/news_58.html

require simply entering personal tags through the keyboard. Users of browser-based bookmarking tools should feel comfortable using social bookmarking. Recommender systems also have a very low complexity, with participation in some occurring automatically and most only requiring a simple rating system. For e-voting to take hold, the interface must be as non-restrictive as possible. User interfaces that support accurate decision entries is of prime importance to vote systems for political elections and has been the subject of much federally-sponsored and independent research [68, 69]. Electronic vote systems in use today employ an extremely simple point-and-click or touchscreen interface. While the accurate and accessible functioning of the interface is important, the perception of such an interface is essential as vote systems are satisfaction based [70].

Restrictive interfaces require skills that are not yet part of the standard repertoire of general computer users. For example, wikis employ user-friendly interfaces similar to non-restrictive systems. However, MediaWiki powered editing sites such as Wikipedia use a wiki markup language called wikitext that, while very simplified, is more complex than the use of word processing programs. Prediction markets also have an interface that is moderately complex. Unfamiliarity with trade processes makes the prediction market a specialized decision tool that may alienate some potential contributors. There are a number of systems attempting to overcome this hurdle in commercial ventures. Inkling markets,[21] for instance, focuses on ease of use by simplifying the burden on the user to interpret price movement. While the interface itself is simplistic, the market was simplified as well through the implementation of an alternative market design. Here, traders use a scale (e.g., the market price is slightly low, low, or way too low) to input their decision instead of placing a bid or ask.

Some systems are highly restrictive. Document ranking is a unique system type in that the interface to contribute (by linking a web-page) is not controlled by document ranking systems. While every user of the Google search engine has benefited from document ranking, only a subset of the web-using population has contributed to the system. Document ranking requires the control of a web-page so as to link it to other web-pages. While the actual linking of pages is simple (especially through website creation software such as VCOM's Web Easy Pro and Apple's iWeb), the occasion to do so is more restrictive. Open source software is perhaps the most specialized of the systems discussed as they require skill in software development for their evaluation and contribution. While not strictly interface-based, the high complexity designation is due to the specialization of skills needed to work in open source software as it restricts the population eligible for membership in the collective.

Skill Set. To describe the reasoning behind the interface complexity rating, the skill set distinction describes the online actions required. All web-based CDMS require a minimum level of computer and Internet competency to operate. However, none of the systems discussed are intended to require training.

[21] URI: http://inklingmarkets.com

The Scottish Qualifications Authority[22] outlines four areas of knowledge that encompass basic computing skills. These are computer skills (mouse and keyboard operation, opening and closing files, locating files), e-mail skills, word processing skills, and web skills. Folksonomies, recommender systems, and vote systems all have a low interface complexity rating because they require no more than these basic skills. Wikis have a restrictive designation as they require more sophistication in operation than the aforementioned systems. The markup language is not What You See Is What You Get (WYSIWYG) as word processors are. Prediction markets require only basic computer skills, but demand of the interested user understanding of market trading. Sites that host one of these systems usually include tutorials to familiarize the new user with the system. Document ranking is a highly restrictive system as the contributor must have the skills and means to publish a website to the web. Open source software is restrictive in that intelligent participation requires specialized knowledge in computer programming, software debugging, documentation authoring, etc.

Contributor/User. The contributor/user distinction refers to whether the individuals interacting with the system are the only ones to benefit from the decision or if the benefits reach both the contributors and web users in general. While contributor and user have been used interchangeably throughout this article, in this section contributor refers to the participants entering their decisions into the system (the collective) and user refers to others who use the collective decision. Note that, as will be discussed shortly, all systems are of benefit to the contributor, thus no external enticements are usually required for recruitment into the collective. The only question is whether general users benefit as well.

Both document ranking and folksonomies are of utility to general users as well as contributors. Anyone who has used the Google search engine but has never linked a web-page is evidence of this. Recommender systems become more worthwhile to the contributor as they contribute and build up a pattern of behavior. However, they generate little overall value to a first-time user. In addition, recommender systems that use collaborative filtering to find similarity between users provide no utility to those who are not participating in the system. However, content-based recommender systems that compare the similarity between products could be useful to a first-time user.

Vote systems typically generate decisions that affect solely the populace that votes on them. This is a defining characteristics of direct democracy. However, there is a range of other arrangements that could be envisioned—one person deciding for all (dictatorship), a representative group deciding for one (jury), etc. Despite the potential benefits and detriments an individual may receive without participating in a vote as a result of that vote, there is no sense of a general user in a vote system. Therefore, a vote system is of benefit only to the contributors who are given the opportunity to express their views.

Both wikis and open source software are also of utility to both contributors and general users. Both systems utilize a collective to generate a product of

[22] URI: http://www.sqa.org.uk

wide interest. There are over 50 million page requests on Wikipedia everyday, but only 200,000 edits. Prediction markets provide a game-like environment for contributors to elicit information about the future for others. Unlike vote systems, where there is no sense in which a general user can participate in the system, prediction markets are used similarly to polls. Their prices are tracked through time as forecasts for particular events by individuals who do not trade. This service is even sellable; the Hollywood Stock Exchange was the first to produce a commercialization plan where the information generated by those playing in the markets was sold to interested buyers in the entertainment industry.

3.3 Individual Features

The preceding two classes of the taxonomy have dealt with features of the aggregation mechanism. This class and the following pertain to the composition and statistics of the collective. Individuals that compose the collective maintain independent choice in web-based collective decision making systems. Thus, individuals are important to consider when examining the role of the collective.

Motivation. The use of these systems by a large user-base is in many ways inexplicable. The notion of the most highly consulted online encyclopedia,[23] Wikipedia, being written by unpaid volunteers is in complete paradox to standard economic motivational theories. In addition, low voter turnout in national elections suggests that simply being asked for your opinion is not a sufficient motivation for many. Because of the necessity of large collectives to activate the problem solving potential of these systems, engaging motivating factors is an essential feature of every CDMS.

The need for affiliation is a primary motivational factor in human behavior [71, 72]. This need motivates individuals to make connections with those they want to be associated with. The wild popularity of social networking sites demonstrates that this need for affiliation and ability to connect transfers to the web. The structure of the web is set by a similar desire for connectedness, where a hyperlink serves as an affiliative bond. Through linking, the individual is prescribing where the published page fits in the network of web-pages. Therefore, the PageRank algorithm characterizes an incoming link as a vote of quality for that site as the originator of the link chose to associate with it.

Folksonomies are of particular utility for those who wish to organize and index information. Delicious, for example, replaces browser-based bookmarking with online bookmarking accessible from any computer. A contributor generally tags websites they wish to find again with a word that is meaningful to them without regard for others. The result is a personalized sample of the web. Recommender systems provide personalized advice out of an overwhelming number of options to facilitate browsing and purchasing online. A recommender system is able to best choose similar users if each user has a rich history of behavior in the system allowing the systems to "get to know" the user.

[23] According to Alexa global top 500 URI: http://www.alexa.com accessed July 11, 2007.

Vote systems, as a method to elicit the desires of the populace, function by allowing each voter to express their beliefs. A vote system is cooperative in that an individual hopes that others are deciding in the same way they are, thus increasing the likelihood that their desire will be chosen. On the other hand, prediction markets are competitive in that a trader makes the most money if they express a view that most others do not have and they are correct in their prediction. These distinctions affect the way information is shared in the system. For different reasons, participants in each system may be unwilling to share the decision they registered.

Users of both wikis and open source software systems, as forms of content creation, are fundamentally motivated by a desire to impart knowledge to create valuable tools. Each edit in a content creation system is motivated by a criticism of the work in its present form. As in document ranking, contributors are motivated by the existing content.

Expertise. Expertise is the knowledge an individual must have for the system to generate an accurate decision. This is not to be confused with the skill needed to operate the system interface. It is of fundamental importance to distinguish systems that are for experts from ones that operate on more general principles. Systems that do not require experts work simply from a statistical collective intelligence perspective where the more people who participate, the more likely a satisfactory result will emerge. On the other hand, expertise-based systems must elicit information from those specially knowledgeable to generate satisfactory results.

None of the three information retrieval systems require expertise. They function through a process of averaging public opinion, although the algorithm varies in each system. However, all three systems have counterparts that do use experts instead of a collective. Mahalo[24] is a social search engine that ranks web-pages by hand. Individuals contribute the best web-pages for a set of popular search terms. These individuals are selected through an online application based on their frequent and high quality participation in other social software sites, rendering them experts. Folksonomies are often compared to their taxonomic counterparts generated by professional taxonomists. Before Amazon.com, librarians served as the experts in connecting people with their media whims and needs. It is through system design that expert knowledge-keepers can be replaced by an amorphous collective of fallible and untrained individuals.

Like information retrieval systems, vote systems replace the judgment of a single individual with the opinions of the collective. Vote systems do not require expertise as they are held to no accuracy metric other than satisfaction. For this reason, campaigns such as "get out the vote" continue. Every person allowed to vote is so encouraged regardless of their knowledgeability.

On the other hand, content creation systems and prediction markets require expertise. In order to create a work, a participant must be able to provide a useful and unique contribution. Prediction markets will identify the inexpert

[24] URI: http://www.mahalo.com

through his or her dropping portfolio value; however, knowledge of the future state must be present to be amplified in the market.

Membership. Membership refers to the method by which participants become a part of the collective. Almost all of the systems discussed rely on the principle of self-selection. In other words, individuals provide the initial impetus to participate and are not selected upon by the system for fitness in the collective. Document ranking is the only system that does not rely on self-selection. While it is up to each individual to link to whatever web-pages they please, they do not choose to lend this decision to document ranking systems. Instead, the decisions of the collective are co-opted by robots that traverse the web by following these links to determine the structure of the web. It is worth noting that a next generation of search engine designed as social software and typified by Sproose[25] ranks pages based on contributors' explicit votes.

The systems that have a self-selecting collective also require the use of a consistent user name to maintain a persistent identity through time. The log-in serves to organize anonymous and asynchronous interactions with the system into a coherent entity. It also enables the discrete tracking of user behavior for automatic membership. For example, Amazon.com exploits the tracking of a logged-in user to automatically enroll the user in their recommender system. The desires of the user is inferred from past purchases. The fundamental difference between the co-opting of decisions made by individuals for the purposes of document ranking and that of Amazon.com is that the individual on Amazon.com has explicitly engaged in a user relationship with the website by logging in. The confirmation of identity is also important in vote systems where only one vote is allowed by every eligible participant. Problems with the verification of identity is a major impediment to the establishment of online voting systems [73].

Folksonomies, recommender systems, and vote systems take very little care to maintain the quality of their collective. As these systems are not expert-based this is not surprising. However, wikis, open source software, and prediction markets are all systems requiring expertise and thus contain interesting features to ameliorate the impact of unhelpful members. Wikipedia posts a "wanted list" of contributors and IP addresses that have engaged in vandalism to identify those whose edits should be monitored. Persistent vandals are permanently blocked from participating. Open source software provides a hierarchical arrangement where contributions are reviewed before being incorporated into the code. In most cases, packages are signed to provide accountability for poor contributions and to look out for malicious content. To encourage traders to play only if they are reasonably assured of their decision, prediction markets offer incentives based on participants' performance. The monetary and prestige-based incentives encourage one to participate if they desire the reward or not to participate if the consequences are too great. Traders form a self-selecting population where each individual chooses if, when, and the extent of their participation. To participate

[25] URI: http://www.sproose.com

without knowledge may lead to financial losses for the trader, which hinders future participation.

3.4 Collective Features

Statistics regarding the collective in a CDMS may be difficult to interpret as the collective itself is an amorphous and changing collection of individuals. However, any collective decision must consider its aggregation mechanism in conjunction with the facts of its collective.

Size. Size refers to the number of individuals needed in a collective to produce a collective decision of quality. This extremely relative measure is designated either variable or large for our systems of interest. Systems that require a large collective suggests that statistical collective intelligence plays a role in generating quality results. In other words, it is through high participation levels that accuracy develops. Conversely, systems that can handle a variable population size suggests that expertise is required. The only system (of the seven) where this does not hold true is the vote system. A vote system allows a variable population size but does not require expertise. As vote systems are based on the principle of fairness, the vote need only satisfy this requirement. An exception is the requirement of a quorum adopted by some voting bodies. A quorum is the minimum number of people needed to be present to participate in a vote to make it legitimate. In web-based votes where an individual's "presence" during a vote is difficult to guarantee, the institution may require a per-option quorum[26] where an *option* must receive the number of votes equal to the quorum before it can be considered a winner. The per-option quorum protects against a non-monotic situation where the vote cast to reach quorum allows another option to win. Full participation in voting systems can alternatively be simulated when presence to vote is infeasible. For example, the trust-based social network algorithm dynamically distributed democracy (DDD) simulates complete participation in a direct democracy as user participation wanes [74].

Information retrieval systems work best with a large number of contributions. This is because of the reliance on statistical collective intelligence to provide a complete and rich description of the solution space. As more information is contributed through individual interaction with the system a cleaner probability distribution is generated.

Content creation tools allow a variable collective size. The size necessary depends on the complexity of the decision and the distribution of knowledge on the topic. If there are three foremost experts in an area, then others may not be necessary. Open source software systems echo the sentiment of more is better with Torvald's famous quote, "Given enough eyes, all bugs are shallow" [37]. The size of the population of prediction markets necessary to generate an accurate solution is not a well-researched subject. While traditional financial markets operate with thousands of participants a day, prediction markets can handle, but do not

[26] Implemented by Debian URI: http://www.debian.org

require this amount of traffic [75]. Ostensibly, this is an expert-based system, so if the knowledge to predict the future is held between a few, then those are the only ones that need participate. However, the noise trader in traditional markets induces experts to participate by moving prices away from a correct value [76]. In other words, the poor contributions of noise traders allow experts to include relevant information and thus earn money by moving a price back in line. The Iowa Electronic Markets (IEM) advise that 20 to 30 participants can generate accurate predictions.[27]

Diversity. The role of diversity is a well-studied area of collective dynamics [77, 78]. Diversity is the fundamental mechanism behind the emergence of collective decision making. A collective is necessarily diverse, although the ways in which the individuals differ are of importance. Some systems benefit by utilizing a population that represents different pieces of information because the diverse contributions help to cover the solution space. For these systems, it is through diversity and a large collective size that optimal solutions are generated. In other systems, diversity allows an individual to improve upon the contributions of another [79]. A collective is used precisely because only through a large distribution do patterns of consensus become apparent. Thus, all systems balance the exploitation of diversity with the capturing of similarity. The seven systems are classified according to their most prominent use of diversity—coverage of the solution space or incremental improvement upon the current solution.

Information retrieval systems rely on a comprehensive index of the collection that makes up the solution space. Thus, a collective is used to gather information about this space. Large numbers are required in the collective to incorporate enough diversity to cover the solution space. If all users were totally homogenous no general distribution would be required. Recommender systems require participants to have similar preferences, but a diversity of experiences leading to differences in the items they have accessed.

Diversity is not always a desired characteristic in collective systems. For example, in a vote system, it would be best if every participants' views were in total accord. As long as the vote system properly delegates the favorable position, the system will be regarded as universally fair. While debate is a cornerstone of democracy, consensus is ideal for the vote system. In such a case, all votes would return a unanimous decision. Thus, neither type of diversity is desired in a vote system.

Both content creation systems and prediction markets require diversity to produce incremental improvements in the system. To generate a collective decision in these systems, it is important that each person has a different skill set, element of knowledge, or critique to contribute. In prediction markets, diversity is the impetus for trade. It is the individually different valuations of contract prices that initiate trades. The market aggregates the incremental movements of contracts toward an accurate prediction. The competitiveness of prediction markets, where a trader succeeds at anothers failure, encourages the contribution of

[27] IEM FAQ http://fluprediction.uiowa.edu/fluhome/FAQ.html

diverse prediction-relevant information. Before each participant chooses to trade in a market, they must evaluate the uniqueness of their information. A trader has an opportunity to perform the best if they have unique information. In other words, if the market price does not already reflect a trader's information he or she can earn money by buying or selling shares to bring the actual price closer to their estimation.

Interaction. Interaction is a property of the collective that refers to the amount of feedback experienced by the contributors from other members of the collective. For our purposes, this feature is broken down into three types of interaction. Imitative refers to a level of interaction that urges a normative response in the user. Strategic refers to the expression of decisions based on a strategic analysis of options. Stigmergic refers to the indirect communication left by individuals in a shared space [80].

Document ranking is not inherently interaction-based. Contributors simply choose to link to other web-pages and in aggregate this produces a connected network. Recommender systems do no require direct interaction between others in the system. In face, the lack of transparency connecting past preferences to recommendations leads some users to "test" the system to try to reveal why a given recommendation was made. To counteract this behavior, some sites now explain their recommendations [58]. For example, Amazon.com explains that a given item was recommended based on a specific item that was either viewed or purchased by the user. Folksonomies also do not require interaction; however, the convergence of tags to produce a coherent system depends upon individuals choosing to tag as others have. The popularity of tags that were originally used for a document and other patterns in tagging behavior suggest imitative interaction [15, 11].

A vote system may require strategic interaction with others in the system. In nearly contemporaneous papers, Gibbard and Satterthwaite presented a theorem of broad circumstances in which voters have an incentive to strategically vote in a manner that does not reflect their true preferences [81, 82]. For example, in some systems, if an individual votes on an option that is not in serious contention, a third party vote for example, that is considered a wasted vote. The voter would have better expressed their desires, if they knew that there would not be strong support for their first choice, by choosing a more likely contender. The best strategy is dictated by the aggregation algorithm employed. Prediction markets, like all financial markets, also involve strategic interaction with the system as they are game-like. Specific strategies for each aggregation mechanism of prediction markets have been researched both to aid in strategy implementation and understand their effects on system accuracy [83, 84].

Content creation systems have a high level of stigmergic interaction as the work itself functions as the feedback within which users interact. The large number of contributors that participate in these systems extends our pre-Internet notions of the size of collaboration. The scale of collaboration in wikis and open source software is reminiscent of insect colonies. Thus, it is apt that the tools used to facilitate this collaboration are similar to those of insect colonies [85].

Table 2. Collective decision making systems in the feature space

	Document Ranking	**Folksonomy**	**Recommender**	**Vote**	**Wiki**	**Open Source**	**Prediction Market**
Problem Space							
Decision Type	information retrieval	information retrieval	information retrieval	governance	content creation	content creation	prediction
Decision Principle	centrality	frequency	similarity	frequency	consensus	consensus	trade
Goal	quality retrieval	quality retrieval	quality retrieval	satisfaction	document utility	code utility	predictive accuracy
Accuracy Metric	precision recall	precision recall	precision recall	fairness	usability	usability	forecast standard error
Implementation							
Solution Space	number of artifacts	number of artifacts	number of artifacts	ballot	creative output	creative output	disjoint + exhaustive
Interface Complexity	very restrictive	not restrictive	not restrictive	not restrictive	restrictive	very restrictive	restrictive
Skill Set	web-page design	basic skills	basic skills	basic skills	wikitext syntax	programming	market trading
Contributor/User	both	both	contributors	contributors	both	both	both
Individual Features							
Motivation	connectedness	organization	personalized advice	cooperative	critical	critical	competitive
Expertise	unnecessary	unnecessary	unnecessary	unnecessary	necessary	necessary	necessary
Membership	co-opted	self-selecting	auto/self-selecting	self-selecting	self-selecting	self-selecting	self-selecting
Collective Features							
Size	large	large	large	variable	variable	variable	variable
Diversity	coverage	coverage	coverage	none	improvement	improvement	coverage + improvement
Interaction	none	imitative	none	strategic	stigmergic	stigmergic	strategic

Both use the environment to leave information that communicates to others. Wikis improve efficiency in this communication process by assembling a list of pages that need to be written (essentially the links looking for articles) and open source software often employs postings of known problems to focus the efforts of myriad contributors [86]. In sum, these features facilitate the high interaction levels of these productive systems that might otherwise be overwhelmed by the chaos of so many contributors.

4 Conclusion

The move to web-based collective decision making systems has precipitated an enhanced ability to gather useful information from individuals as well as aggregate this information using scalable techniques for a variety of outcomes. The taxonomy presented defines each system by the unique combination of their features and highlights similarities between the systems. Unexplored combinations suggest a potential for the development of additional systems to meet our decision-making needs. It is left to future work to examine the feature space of web-based collective decision making systems to determine the unexploited options and the unexplored combinations to design new tools. Each system has its own particular benefits, specific applications in the problem space, and disadvantages. If the variations between the systems are explored and the best system for a particular problem is determined, then CDMSs will have reached the extent of their abilities to facilitate decisions.

References

1. Desanctis, G., Gallupe, R.B.: A foundation for the study of group decision support systems. Management Science 33(5), 589–609 (1987)
2. Tepper, M.: The rise of social software. netWorker 7(3), 18–23 (2003)
3. Luhmann, N.: Soziale Systeme. Suhrkamp Verlag, Frankfurt am Main (1984)
4. Atlee, T.: Forms of collective intelligence (May 2006), Retrieved January 20, 2007 from, `http://www.community-intelligence.com`
5. Condorcet, M.d.: Essai sur l'application de l'analyse a la probabilite des decisions rendues a la pluralite des voix, Imprimerie Royale, Paris (1785)
6. Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank citation ranking: Bringing order to the web. Technical report, Stanford Digital Library Technologies Project (1998)
7. Brin, S., Page, L.: The anatomy of a large-scale hypertextual web search engine. Computer Networks 30(1–7), 107–117 (1998)
8. Kleinberg, J.M.: Authoritative sources in a hyperlinked environment. In: Proceedings of the ACM-SIAM Symposium on Discrete Algorithms, pp. 668–677. ACM Press, New York (1998)
9. Castillo, C., Chellapilla, K., Davison, B.D. (eds.): Proceedings of the 3rd International Workshop on Adversarial Information Retrieval on the Web. ACM Press, New York (2007)

10. Mathes, A.: Folksonomies - cooperative classification and communication through shared metadata. Computer Mediated Communication - LIS590CMC (graduate course) (December 2004)
11. Cattuto, C., Loreto, V., Pietronero, L.: Semiotic dynamics and collaborative tagging. Proceedings of the National Academy of Science 104(5), 1461–1464 (2007)
12. Hammond, T., Hannay, T., Lund, B., Scott, J.: Social bookmarking tools (1). D-Lib Magazine 11(4) (April 2005)
13. Voss, J.: Tagging, folksonomy, and co.—renaissance of manual indexing. In: 10th International Symposium for Information Science, Cologne (January 2007)
14. Udell, J.: Collaborative knowledge gardening. InfoWorld (August 2004)
15. Golder, S.A., Huberman, B.A.: Usage patterns of collaborative tagging systems. Journal of Information Science 32(2), 198–208 (2006)
16. Steinbock, D., Pea, R., Reeves, B.: Wearable tag clouds: Visualizations to support new collaborations. In: Conference on Computer Supported Collaborative Learning, Camden, New Jersey (2007)
17. Resnick, P., Varian, H.: Recommender systems. Communications of the ACM 40(3), 56–58 (1997)
18. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering 17(6), 734–749 (2005)
19. Samet, H.: Similarity searching: Indexing, nearest neighbor finding dimensionality reduction, and embedding methods for applications in multimedia databases. In: International Conference on Pattern Recognition, Cambridge, UK (August 2004)
20. Vucetic, S., Obradovic, Z.: A regression-based approach for scaling-up personalized recommender systems in e-commerce. In: Workshop on Web Mining for E-Commerce at 6th ACM-SIGKDD, Boston, MA (2000)
21. Maltz, D., Ehrlich, K.: Pointing the way: Active collaborative filtering. In: Proceedings of the CHI 1995 Human Factors in Computing Systems, ACM, New York (1995)
22. Ahn, H.J.: A hybrid collaborative filtering recommender system using a new similarity measure. In: Proceedings of the 6th WSEAS International Conference on Applied Computer Science, Hangzhou, China (April 2007)
23. Middleton, S.E., Shadbolt, N.R., De Roure, D.C.: Ontological user profiling in recommender systems. ACM Transactions on Information Systems 22(1), 54–88 (2004)
24. Ramakrishnan, N., Keller, B.J., Mirza, B.J., Grama, A.Y., Karypis, G.: Privacy risks in recommender systems. IEEE Internet Computing, 54–62 (2001)
25. Straffin Jr., P.D.: Topics in the Theory of Voting. Birkhauser, Boston (1980)
26. Arrow, K.: Social Choice and Individual Values, 2nd edn. John Wiley and Sons, Chichester (1963)
27. Thornton, A.: Does internet create democracy? Ecquid Novi 22(2) (2001)
28. Robertson, S.P., Wania, C.E., Park, S.J.: An observational study of voters on the Internet. In: 40th Annual Hawaii International Conference on System Sciences (HICSS 2007), Waikoloa, Hawaii (January 2007)
29. Robertson, S.P.: Voter-centered design: Toward a voter decision support system. ACM Transactions on Computer-Human Interaction 12(2), 263–292 (2005)
30. Kohno, T., Stubblefield, A., Rubin, A.D., Wallach, D.S.: Analysis of an electronic voting system. In: IEEE Symposium on Security and Privacy, pp. 27–40 (May 2004)
31. Schyren, G.: How security problems can compromise remote Internet voting systems. Electronic Voting in Europe, 121–131 (2004)

32. Rodriguez, M.A., Steinbock, D.J., Watkins, J.H., Gershenson, C., Bollen, J., Grey, V., deGraf, B.: Smartocracy: Social networks for collective decision making. In: 40th Annual Hawaii International Conference on Systems Science (HICSS 2007), Waikoloa, Hawaii (2007)
33. Leuf, B., Cunningham, W.: The Wiki Way: Quick Collaboration on the Web. Addison-Wesley, Reading (2001)
34. Oster, G., Urso, P., Molli, P., Imine, A.: Data consistency for P2P collaborative editing. In: Proceedings of the 2006 Conference on Computer Supported Cooperative Work, pp. 259–268 (2006)
35. Almeida, R.B., Mozafari, B., Cho, J.: On the evolution of Wikipedia. In: Proceedings of ICWSM, Boulder, CO (2007)
36. Ebersbach, A., Glaser, M., Heigl, R., Dueck, G., Adelung, A.: Wiki: Web Collaboration. Springer, New York (2005)
37. Raymond, E.S.: The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary. O'Reilly, Sebastopol (1999)
38. Yamauchi, Y., Yokozawa, M., Shinohara, T., Ishida, T.: Collaboration with Lean Media: How open-source software succeeds. In: Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, pp. 329–338 (2000)
39. Samuelson, P.: IBM's pragmatic embrace of open source. Communications of the ACM 49, 21–25 (2006)
40. Weber, S.: The Success of Open Source. Harvard University Press, Cambridge, MA (2004)
41. Baldi, S., Heier, H., Mehler-Bicher, A.: Open courseware and open source software. Communications of the ACM 46, 105–107 (2003)
42. Lerner, J., Tirole, J.: Some simple economics of open source. Journal of Industrial Economics 52, 197–234 (2002)
43. Berg, J., Rietz, T.A.: Prediction markets as decision support systems. Information Systems Frontiers 5(1), 79–93 (2003)
44. Wolfers, J., Zitzewitz, E.: Prediction markets. Journal of Economic Perspectives 18(2), 107–126 (2004)
45. Smith, E., Farmer, D.J., Gillemot, L., Krishnamurthy, S.: Statistical theory of the continuous double auction. Quantitative Finance 3(6), 1469–7688 (2003)
46. Hanson, R.: Logarithmic market scoring rules for modular combinatorial information aggregation. Journal of Prediction Markets 1(1), 3–15 (2007)
47. Pennock, D.M.: A dynamic pari-mutuel market for hedging, wagering, and information aggregation. In: ACM Conference on Electronic Commerce, New York (2004)
48. McCrory, G.: Iowa electronic markets forecasted bush win in presidential election. University of Iowa News Service (November 2004) Retrieved January 29, 2007, `http://www.news-releases.uiowa.edu`
49. Hanson, R.: Policy analysis market archive. George Mason University (2003) Retrieved January 29, 2007, `http://hanson.gmu.edu/policyanalysismarket.html`
50. Abramowicz, M.B.: The hidden beauty of the quadratic market scoring rule: A uniform liquidity market maker, with variations. Technical report, The George Washington Law School (March 2007)
51. Servan-Schrieber, E., Wolfers, J., Pennock, D.M., Galebach, B.: Prediction markets: Does money matter? Electronic Markets 14(3) (2004)
52. Rodriguez, M.A.: Social decision making with multi-relational networks and grammar-based particle swarms. In: 40th Annual Hawaii International Conference on System Sciences (HICSS 2007), Waikoloa, Hawaii (2007)

53. Viegas, F.B., Wattenberg, M., Dave, K.: Studying cooperation and conflict between authors with history flow visualizations. In: Proceedings of SIGCHI Conference on Human Factors in Computing Systems, Vienna, Austria, pp. 575–582 (2004)
54. Cleverdon, C.: On the inverse relationship of recall and precision. Journal of Documentation 28, 195–201 (1972)
55. Shafi, S.M., Rather, R.A.: Precision and recall of five search engines for retrieval of scholarly infomation in the field of biotechnology. Webology 2(2) (August 2005)
56. Clark, S., Willett, P.: Estimating the recall performance of search engines. ASLIB Proceedings 49(7), 184–189 (1997)
57. Guy, M., Tonkin, E.: Folksonomies: Tidying up tags? D-Lib Magazine 12(1) (January 2006)
58. Herlocker, J.L., Konstan, J.A., Terveen, L.G., Riedl, J.T.: Evaluating collaborative filtering recommender systems. ACM Transactions on Information Systems 22(1), 5–53 (2004)
59. McNee, S.M., Riedl, J., Konstan, J.A.: Being accurate is not enough: How accuracy metrics have hurt recommender systems. In: Extended Abstracts of the 2006 ACM Conference on Human Factors in Computing Systems, Montreal, Canada (April 2006)
60. Blais, A.: The debate over electoral systems. International Political Science Review 12, 239–260 (1991)
61. Norris, P.: Choosing electoral systems: Proportional, majoritarian and mixed systems. International Political Science Review 18, 297–312 (1997)
62. Reilly, B.: Social choice in the south seas: Electoral innovation and the Borda count in the pacific island countries. International Political Science Review 23(4), 355–372 (2002)
63. Giles, J.: Internet encyclopaedias go head to head. Nature, 900–901 (December 2005)
64. Fitzgerald, B.: A critical look at open source. Computer 37, 92–94 (2004)
65. Berg, J., Nelson, F., Rietz, T.A.: Accuracy and forecast standard error of prediction markets. Technical report, University of Iowa, College of Business Administration (July 2003)
66. Fine, L.: HP Prediction Markets. Confab.Yahoo! (December 2006)
67. Gulli, A., Signorini, A.: The indexable web is more than 11.5 billion pages. In: 14th International World Wide Web Conference, Chiba, Japan (May 2005)
68. Wiklund, M.: Procuring a user-centered voting system. Technical report, United States of America Federal Election Commission (October 2003)
69. Bederson, B.B., Lee, B., Sherman, R.M., Herrnson, P.S., Niemi, R.G.: Electronic voting system usability issues. In: Proceedings of SIGCHI Conference on Human Factors in Computing Systems. ACM Press, New York (2003)
70. Nass, C.: Perceptions of the voting process: Measurement and relationship to interface design. In: Proceedings of the Workshop on Election Standards and Technology, American Association for the Advancement of Science, Washington, D.C. (January 2002)
71. Murray, H.A.: Explorations in Personality. Oxford University Press, New York (1938)
72. McClelland, D.C.: The Achieving Society. Van Nostrand, Princeton, NJ (1961)
73. Jefferson, D., Rubin, A.D., Simons, B., Wagner, D.: A security analysis of the secure electronic registration and voting experiment (SERVE) (January 2004), Retrieved June 15, 2007 from, `www.servesecurityreport.org`

74. Rodriguez, M.A., Steinbock, D.J.: A social network for societal-scale decision-making systems. In: Proceedings of the North American Association for Computational Social and Organizational Science Conference (NAACSOS 2004), Pittsburgh, PA, USA (2004)
75. Hanson, R.D.: Decision markets. IEEE Intelligent Systems 14(3), 16–19 (1999)
76. Wolfers, J., Zitzewitz, E.: Five open questions about prediction markets. Technical report, National Bureau of Economic Research (February 2006)
77. Johnson, N.L.: Diversity in decentralized systems: Enabling self-organizing solutions. Technical report, Los Alamos National Laboratory (November 1999)
78. Page, S.E.: The Difference: How the Power of Diversity Creates Better Groups, Teams, Schools, and Societies. Princeton University Press, Princeton (February 2007)
79. Hong, L., Page, S.E.: Problem solving by heterogeneous agents. Journal of Economic Theory 97(1), 123–163 (2001)
80. Grasse, P.P.: La reconstruction du nid et les coordinations interindividuelles, la theorie de la stigmergie. Insectes Sociaux 6, 41–84 (1959)
81. Gibbard, A.: Manipulation of voting schemes: A general result. Econometrica 41(4), 587–601 (1973)
82. Satterthwaite, M.A.: Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. Journal of Economic Theory 10, 187–217 (1975)
83. Dimitrov, S., Sami, R.: Non-myopic strategies in prediction markets. In: ACM Conference on Electronic Commerce, San Diego (June 2007)
84. Nikolova, E., Sami, R.: A strategic model for information markets. In: ACM Conference on Electronic Commerce, San Diego (June 2007)
85. Elliot, M.: Stigmergic collaboration: The evolution of group work. M/C Journal 9(2) (May 2006)
86. Heylighen, F.: Why is open access development so successful?: Stigmergic organization and the economics of information. In: Lutterbeck, B., Barwolff, M., Gehring, R.A. (eds.) Open Source JahrBuch, Lehmanns Media, Berlin (2007)