MIT Sloan School of Management

MIT Sloan School Working Paper 4679-08
12/31/2007

The MIT Collaboratorium: Enabling Effective Large-Scale Deliberation for Complex Problems

Mark Klein

The MIT Collaboratorium: Enabling Effective Large-Scale Deliberation for Complex Problems

Mark Klein
MIT Center for Collective Intelligence
m_klein@mit.edu

Abstract. While current online discussion tools such as email, chat, wikis, and web forums have been enormously successful at enabling unprecedented global knowledge sharing, they face significant limitations from the perspective of enabling effective large-scale deliberation around complex and controversial issues such as climate change. This paper describes the design and rationale of a system, called the Collaboratorium, which was developed to transcend these limitations by supporting large-scale on-line argumentation.

The Challenge – Large-Scale Deliberation for Complex Problems

Humankind now finds itself faced with a range of what we can call *systemic* problems, i.e. vastly complex challenges, like climate change, that affect every one of us and are affected by every one of our actions. Such problems call for us to be able to engage in effective deliberations on a global scale. The spectacular emergence of the Internet has enabled unprecedented opportunities for such interactions - via email, instant messaging, chat rooms, blogs, forums, wikis, podcasts, video and picture and idea sharing sites, and the like - on a scale that was impossible a few short years ago. To date, however, such large-scale interactions have been incoherent and dispersed, contributions vary widely in quality, and there has been no clear way to converge on well-supported decisions concerning what actions humanity should take to solve their most pressing problems. Can we do a better job of harnessing the vast collective intelligence now potentially available to us? This paper explores this question, reviewing the weaknesses of current technologies, exploring the challenges involved in developing more effective solutions, and presenting a large-scale argumentation system aimed at addressing these challenges.

Limitations of Current Technology

Let us define large-scale deliberation, for this context, as a process wherein many individuals collectively explore and come to decisions over responses to a complex challenge (Walton and Krabbe 1995). Effective deliberation requires that:

- All important issues are considered
- The broadest possible range of high-quality solution ideas are identified
- The strongest arguments for and against each idea are captured
- People can distinguish good from bad arguments
- Individual select solutions rationally, i.e. they consider all the important issues and ideas, and make selections that are consistent with the arguments they most trust
- The aggregate results fairly represent the "wisdom of the people"

How well does current technology support this? Following (Moor and Aakhus 2006), we can divide existing collective interaction ("social software") technologies into three categories: *sharing* tools, wherein individuals share content of value to the wider community; *funneling* tools, wherein group opinions are consolidated into aggregate judgments, and *argumentation* tools, wherein groups identify the space of issues, options, and tradeoffs for a given challenge.

By far the most commonly used technologies, including wikis like wikipedia, media sharing sites like youtube facebook and flickr, open source efforts such as Linux Mozilla and Apache, idea markets such as innocentive, and web forums such as digg and Slashdot, fall into the *sharing* category. While such tools have been remarkably successful at enabling a global explosion of idea and knowledge sharing, they face serious shortcomings from the standpoint of enabling large-scale deliberation around complex and controversial topics (Sunstein 2006) (also see http://en.wikipedia.org/wiki/Criticism_of_Wikipedia):

- *low signal-to-noise ratio*. The content captured by sharing tools is notorious for being voluminous and highly repetitive. This is a self-reinforcing phenomenon: since it can be difficult to find out whether an idea has already been proposed in a large existing corpus, it's more likely that minor variants will be posted again and again by different people. People may also, conversely, decide not to enter an idea, assuming it already appears somewhere. This low signal-to-noise ratio makes it difficult to uncover the novel ideas that inspire people to generate creative new ideas of their own.
- *unsystematic coverage* caused by bottom-up volunteer-based contributions. It's hard to tell what areas are covered in detail and which are not, since there is no compact overview available and no one 'in charge'. There is, as a result, no guarantee that the key issues, best ideas, and strongest arguments have been systematically identified.
- *balkanization:* Users of sharing systems often tend to self-assemble into groups that share the same opinions – there is remarkably little cross-referencing, for example, between liberal and conservative web blogs and forums – so they tend to see only a subset of the issues, ideas, and arguments potentially relevant to a problem.
- *dysfunctional argumentation*: Sharing systems do not inherently encourage or enforce any standards concerning what constitutes valid argumentation, so postings are often bias- rather than evidence- or logic-based, and spurious argumentation is common. Users with divergent opinions often engage in forum "flame wars" (wherein discussions degrade into repetitive arguments and ad hominem attacks) and wiki "edit wars" (where users attempt to "win" by removing each other's perspectives from an article). Such phenomena have forced the shutdown of many forums as well as the locking of many wikipedia pages (so they can not be edited by the public). All these effects seriously degrade the community's ability to fully consider a problem.
- *hidden consensus*: Sharing tools do not provide any explicit means for identifying a group's consensus on a given issue.

Funneling technologies, which include e-voting and prediction markets address the 'hidden consensus' challenge by aggregating individual opinions to determine the most widely/strongly held view among a pre-defined set of options. Such systems, however, often aggregate *biases* as opposed to reasoned logic- and evidence-based *judgments* (though prediction markets can potentially do better since they provide an immediate financial incentive for people to 'get it right'). Funneling systems also do not help people discuss and share what the important issues are, what the alternatives should be, or why one alternative should be selected over any other.

Argumentation (also known as issue networking and rationale capture (Kirschner, Shum et al. 2005) (Moor and Aakhus 2006) (Walton 2005)) tools address this last challenge, by helping groups define networks of *issues* (questions to be answered), *options* (alternative answers for a question), and *arguments* (statements that support or detract from some other statement):

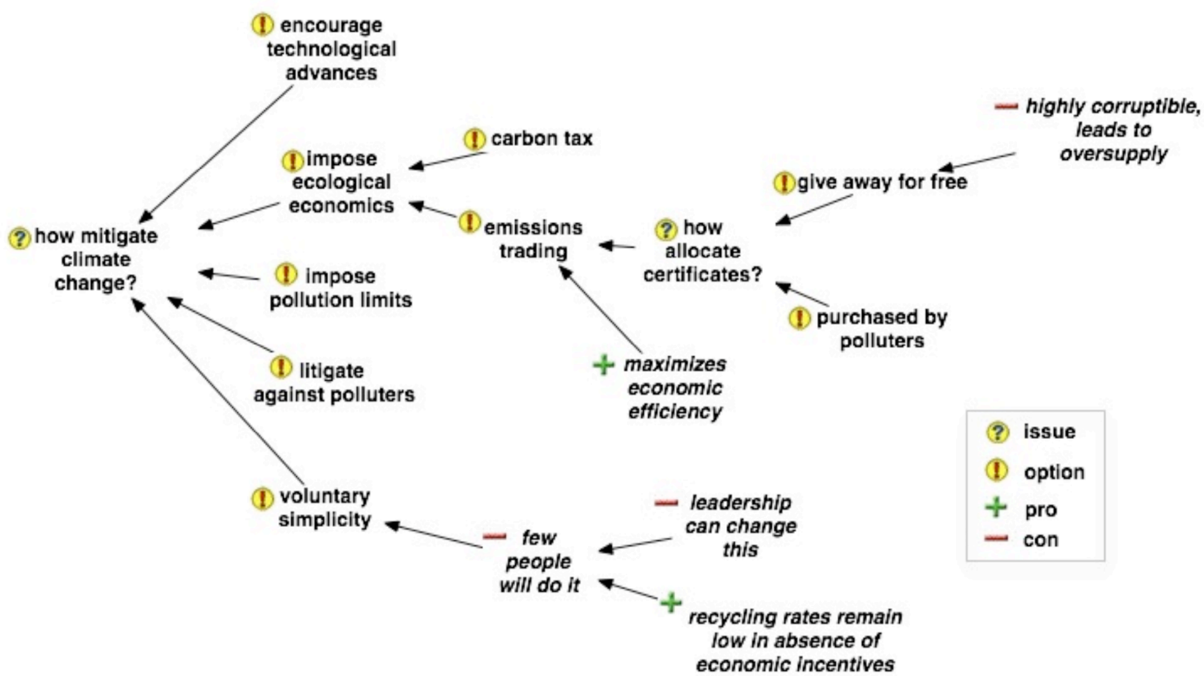

Figure 1. An example argument structure.

Such tools help make deliberations, even complex ones, more systematic and complete. The central role of argument entities encourages careful critical thinking, by implicitly requiring that users express the evidence and logic in favor of the options they prefer. The results are captured in a compact form that makes it easy to understand what has been discussed to date and, if desired, add to it without needless duplication, enabling synergy across group members as well as cumulativeness across time.

One striking initial example of the power this approach is the argument map created on debatemapper.com to summarize the debate surrounding a speech (made on June 12 2007 by former British PM Tony Blair) on the changing relationship between politics and the media. While the majority of commentaries (e.g. news media articles and blogs) individually engaged in relatively shallow and narrow reasoning, each making only a few points, the argument map representing the aggregation of these commentaries "constituted a mature and reasoned response to the Prime Minister's lecture and developed the debate significantly beyond the case he outlined" (see "Mapping the Prime Minister's Media Debate" on opentopersuasion.com).

Current argumentation systems do face some important shortcomings, however. A central problem has been ensuring that people enter their thinking as well-formed argument structures – a time and skill-intensive activity - when the benefits thereof often accrue mainly to *other* people at some time in the future. Most argumentation systems have addressed this challenge by being applied in physically co-located meetings where a single facilitator captures the free-form deliberations of the team members in the form of an commonly-viewable argumentation map (Shum, Selvin et al. 2006). Argumentation systems have also been used, to a lesser extent, to enable non-facilitated deliberations, over the Internet, with physically distributed participants (Jonassen and Jr 2005) (Chklovski, Ratnakar et al. 2005) (Lowrance, Harrison et al. 2001) (Karacapilidis, Loukis et al. 2004) (Heng and de Moor 2003). The scale of these efforts has been small, however[1], involving on the order of 10 participants working together on any given task, far less than what is implied by the vision of large-scale deliberation sought in this work.

Our Vision – A Scenario

The goal of our work is to develop a new class of web-mediated technology that transcends the limitations of existing deliberation-support systems. In order to make our vision more concrete, let us consider an illustrative scenario for how such a system could work for the challenge of climate change:

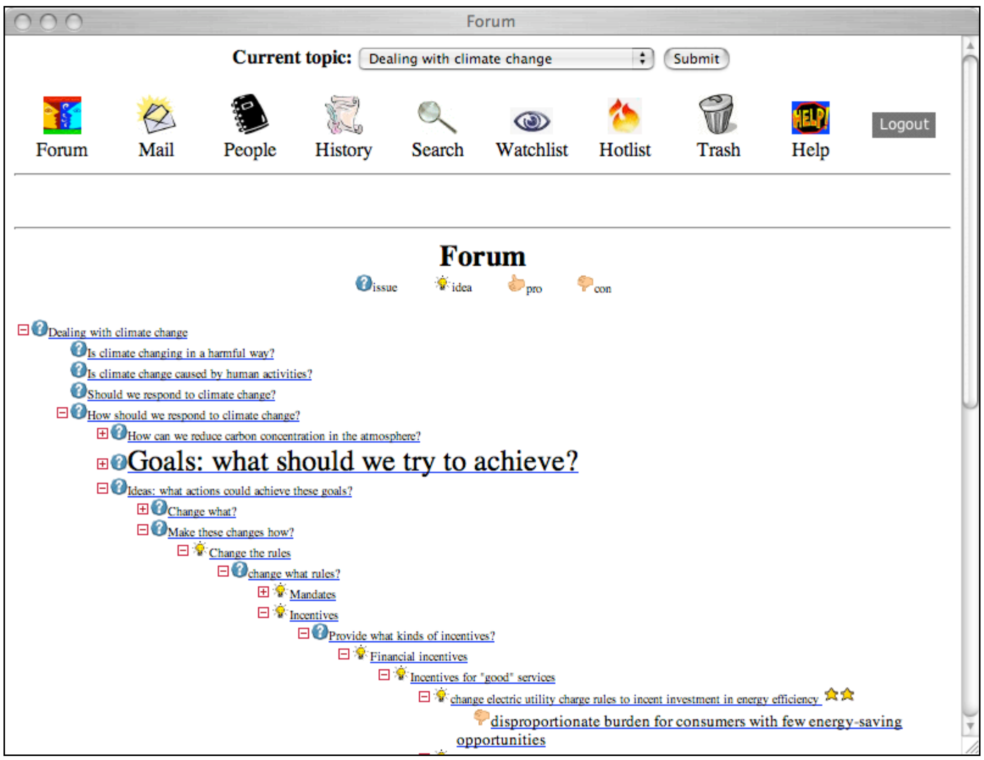

Figure 2. The Collaboratorium discussion forum.

Kara logs in to the Collaboratorium and enters the topic "Dealing with Climate Change". She encounters an indented listing of titles on different topics (figure 2). She's seen the short video

[1] The one exception we are aware of (the Open Meeting Project's mediation of the 1994 National Policy Review (Hurwitz 1996)) was effectively a comment collection system rather than a deliberation system, since the participants predominantly offered reactions to a large set of pre-existing policy documents, rather than interacting with each other to create new policy options.

tutorial, so she knows that articles are organized topically as an interleaved tree of issues and ideas, interspersed with arguments pointing out the strengths and weaknesses of each idea. The font size of the article titles represents how actively the users have been viewing and editing that article lately.

At first she is tempted to browse through the articles that other people have been interested in, but she has a specific purpose – to learn more about something she read in the newspaper, about a plan to add iron to the oceans to help with climate warming somehow. So she uses the search function to look for articles that mention iron:

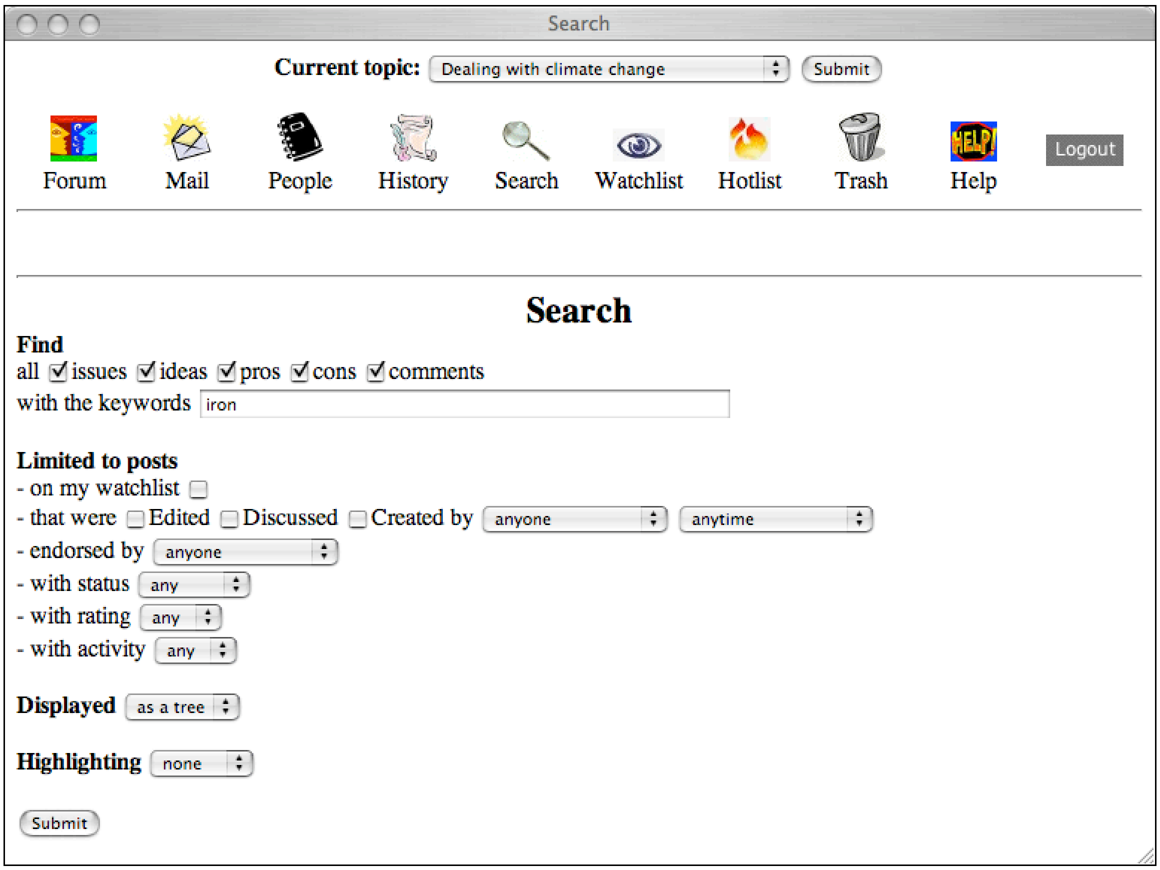

Figure 3. The Collaboratorium search screen.

and finds a relevant match:

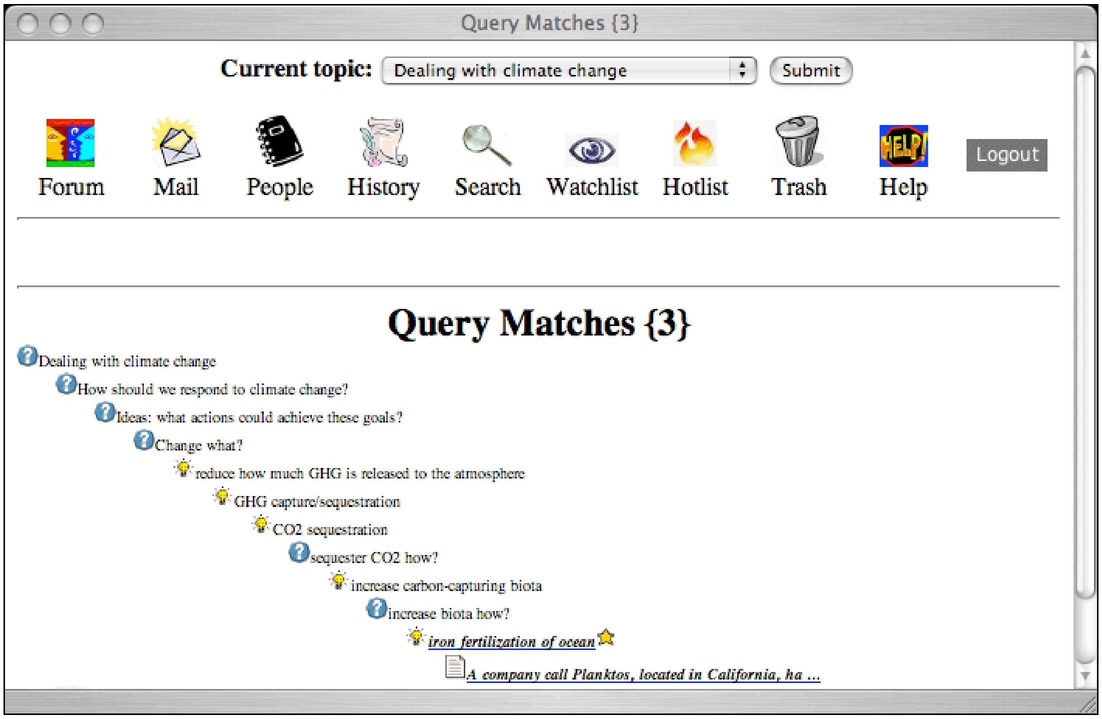

Figure 4. Results from the search function.

She clicks on the article hotlink to see what it says:

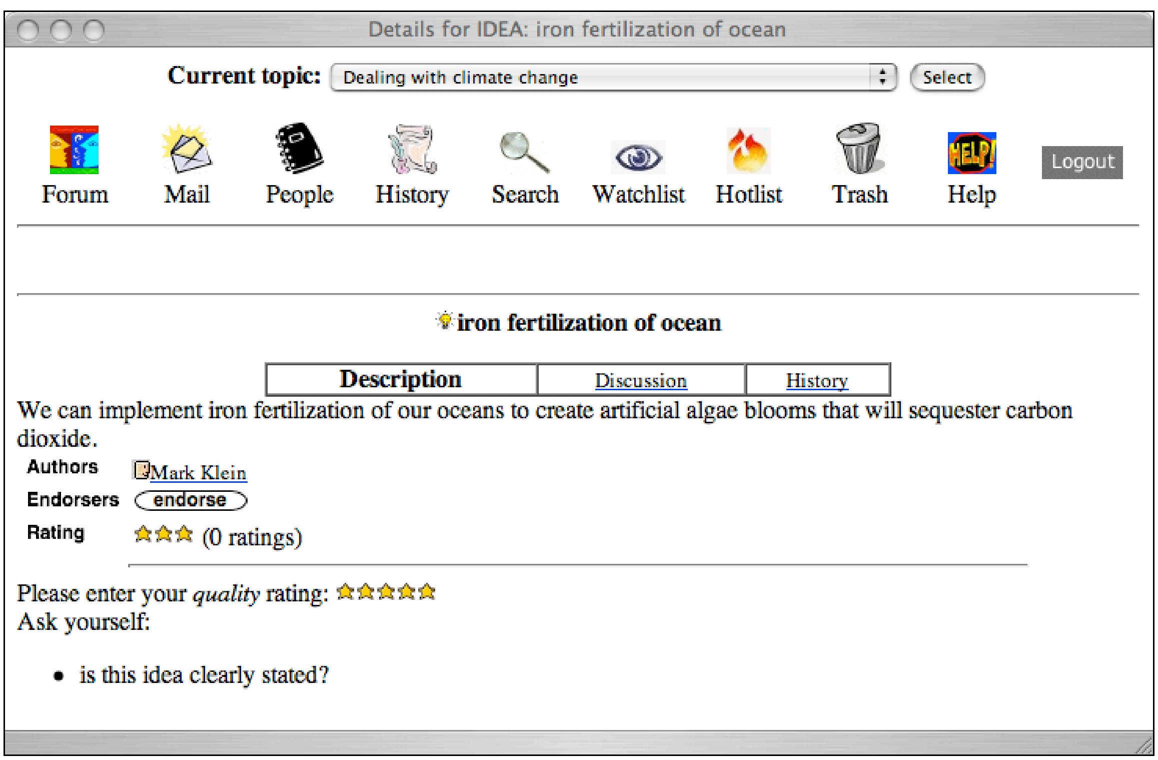

Figure 5. A Collaboratorium article.

Now she can see why people would consider putting iron in the ocean: it encourages the growth of living things that sequester carbon dioxide, a greenhouse gas. Interested in learning more, she clicks on the discussion tab to see what comments other people have left on the topic:

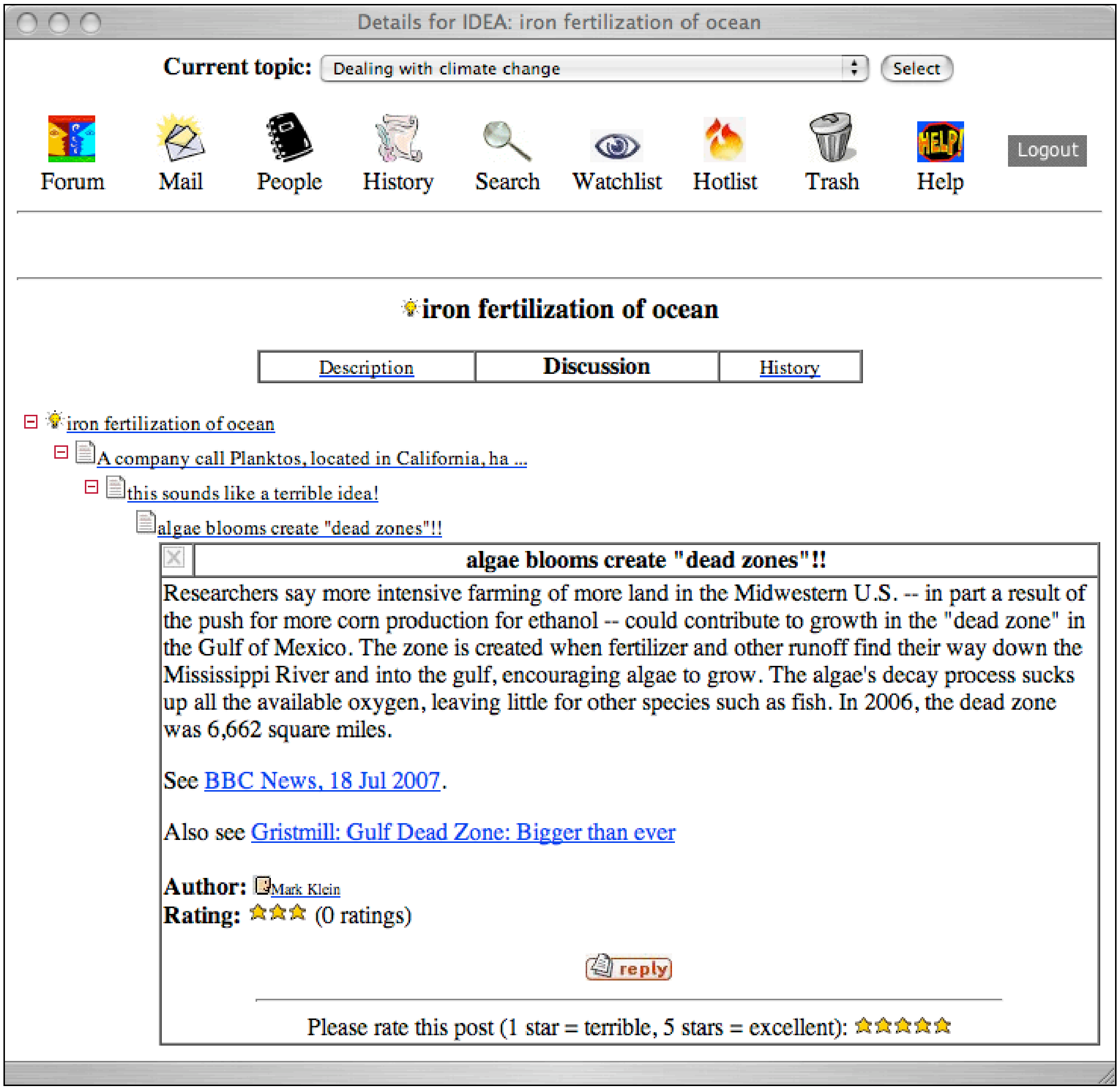

Figure 6. Threaded discussion for a Collaboratorium article.

She can see that iron fertilization is a controversial response to climate change. It’s time to step back and see if the same goals can be achieved with fewer negative effects. By looking at nearby articles in the topic tree, she can see that there are other ways to sequester greenhouse gases by encouraging plant growth, such a planting trees, no-till farming, and so on. These other sequestration approaches also appear to have fewer negative points, judging from the “pro” and “con” articles linked to them, than ocean fertilization:

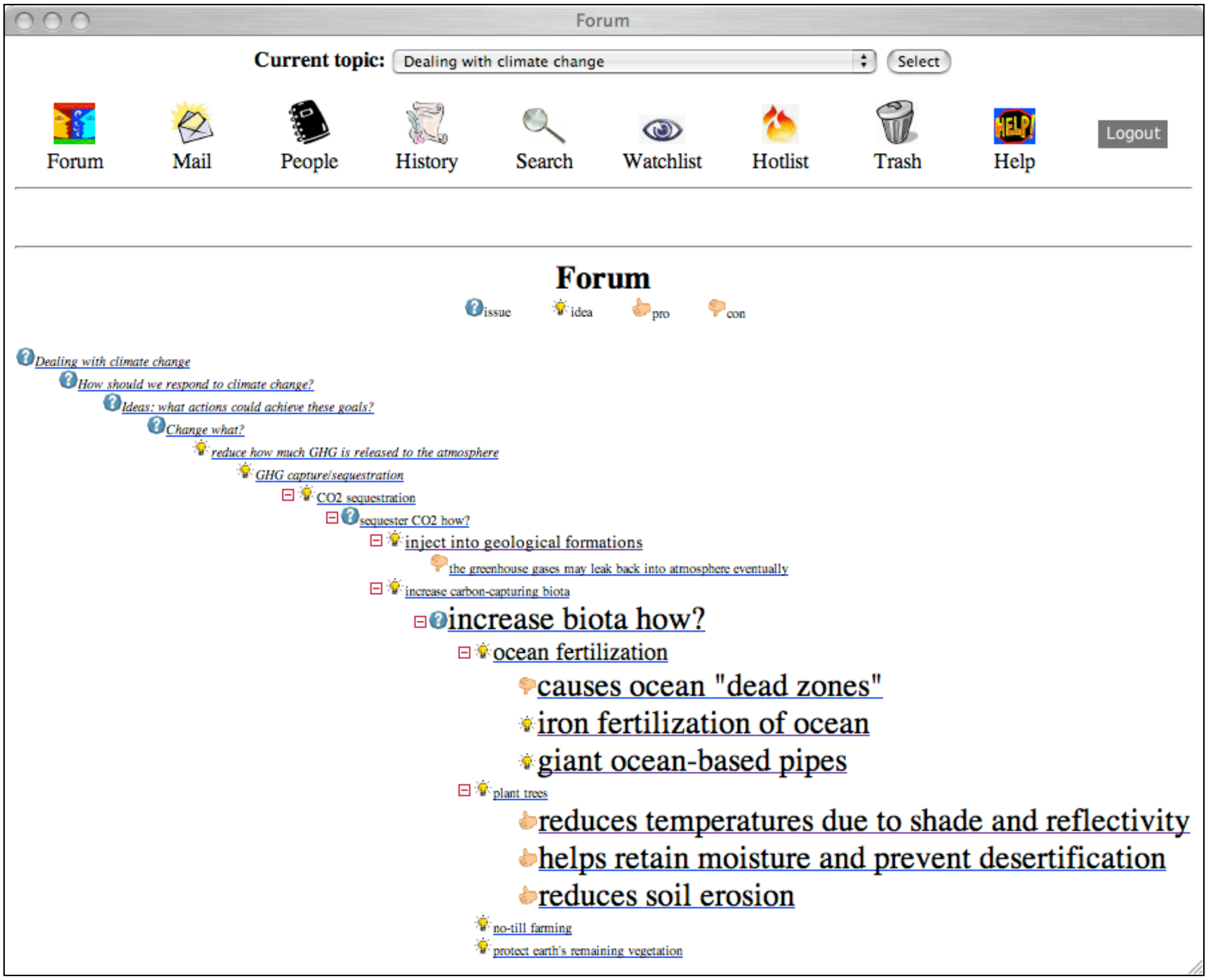

Figure 7. Using the forum to find related articles.

Kara sees that this question has even been put up to a vote: people have weighed in concerning their preferred alternatives for the issue "increase biota how?". She added her own votes, based on her review of the information and arguments she found in the argument map:

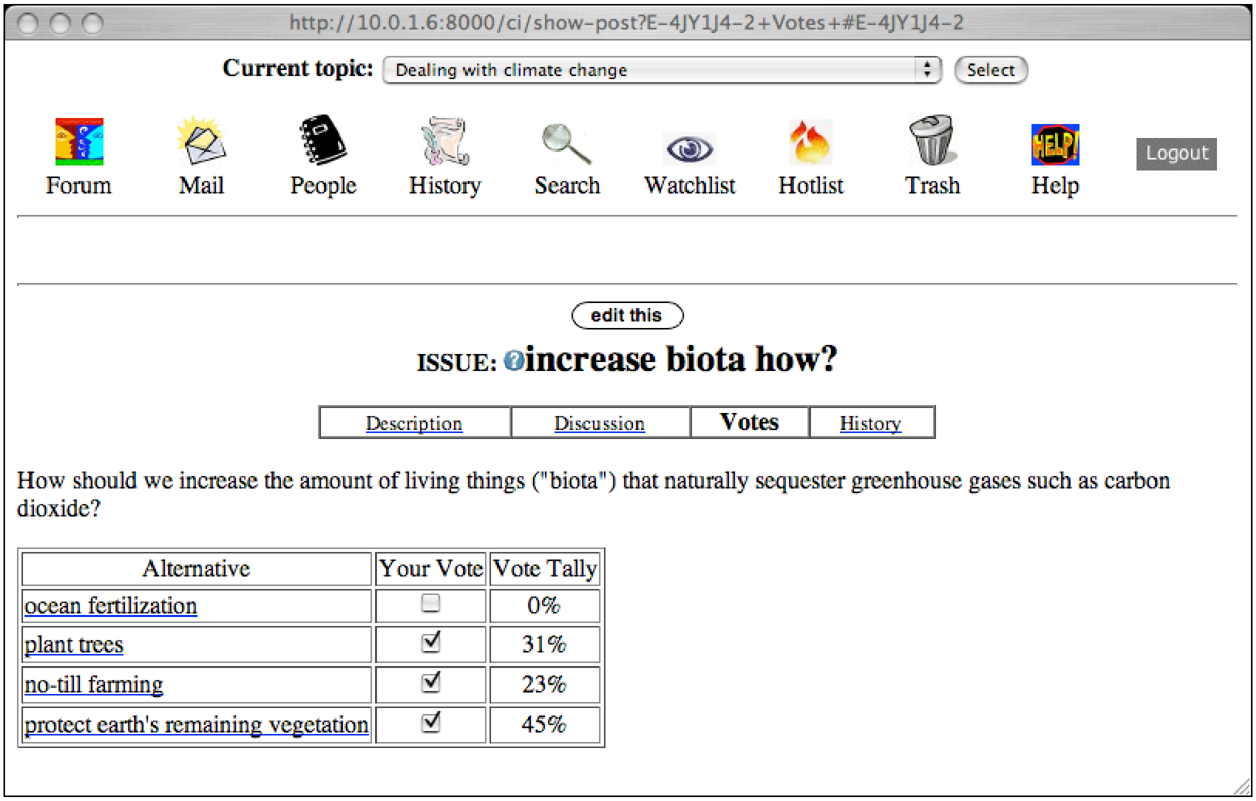

Figure 8. Voting on Collaboratorium ideas.

Kara found that she was particularly influenced by articles that were authored, or endorsed, by organizations she felt she can trust. Every user has a Collaboratorium home page where you can learn about their interests, as well as see what they have contributed. When Kara found an article she liked, she checked out the author's home pages to learn more about them. She could see, for example, that the Climate Change Center at MIT has authored many highly-rated articles on climate change topics: she'll keep her eye out for their work in the future.

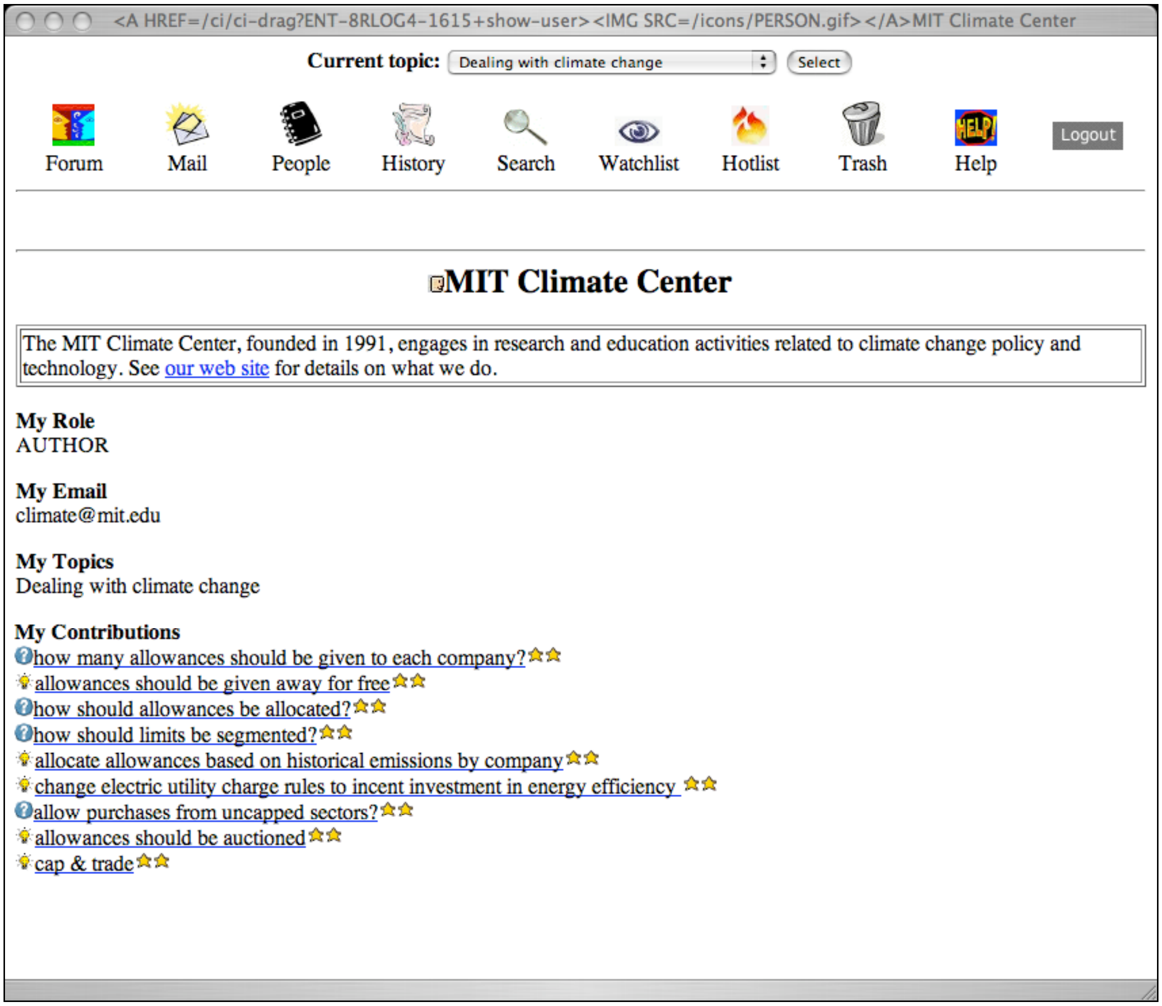

Figure 9. Home page for the (fictional) MIT Climate Center.

Kara now feels that she has a better handle on the issues surrounding iron fertilization and carbon sequestration more generally, which will help her better evaluate what she reads in the paper, as well as what she hears from political candidates concerning their plans to address climate change. Perhaps this new knowledge will influence her purchase decisions, whom she votes for in the upcoming elections, where she sends her charitable contributions, and so on.

Kara also became curious about how this collection of articles is created. She's heard of web-based systems like wikipedia and youtube and digg, where vast amounts of often excellent content is created and evaluated by thousands of volunteers throughout the globe. She watched the video tutorial for potential contributors and learned that there is just such a vibrant community behind the Collaboratorium , involving people from all walks of life. The threaded discussions associated with each article host lively informal interchanges where people share ideas, opinions, personal experiences and relevant news stories, offer suggestions for improving the articles, and ask for and offer advice concerning personal choices. They're a great way to get inspired, find people with similar interests, and even start up collaborations. Some users take the extra step of authoring new articles that concisely present the critical issues, promising new ideas, and important arguments that emerged from these discussions. Editors, which are users with more experience with the system, help ensure the articles are organized in a way that makes it easy for people to find, and potentially add to, knowledge in any given area. The community as a whole, by offering comments and ratings, helps ensure that articles are up-to-date, and that the best articles are easy to identify.

Kara was impressed, in fact, by the depth and breadth of the contributions. The material covers a wide range of topics, ranging from ideas for making houses more energy-efficient, to identifying and comparing different kinds of climate-related governmental policies. The contributors represent scientists, educators, policy makers, engineers and so on from think tanks, corporations, university research centers, and NGOs from around the world. Many are able, by virtue of their access to specialized expertise, large databases, and sophisticated simulation models, to make authoritative contributions concerning the strengths and weaknesses of different climate-related technologies and policies. People of all background and political stripes are represented. Everybody gets to make their own pitch, and all the ideas or arguments on a given subject appear side-by-side so they can be compared directly with one other. Nobody can dominate the discussion because each idea and argument appears just once, in its own article. There seems to be a real community ethos built around the idea of careful well-founded critical thinking. You can tell: the highest-rated articles, and the most active and respected members of the community, all seem to reflect this aspiration.

The fact that players of this caliber are involved has meant that movers and shakers in our society, including educators, policy makers, charitable foundations, social entrepreneurs, and business executives, are increasingly looking to the Collaboratorium to get a sense of what they can do about climate change: they can assess, for example, which options have the most scientific and popular support, and who the key players are. The overall effect, Kara realizes, is to allow an unprecedented number of people to participate effectively in the process of finding solutions to a problem that critically affects all of us. This makes Kara excited about the idea of contributing to the Collaboratorium in some way. She just bought a Prius: perhaps should could start by entering some comments on her experiences with selecting and operating the car? She uses the search function to look for articles on hybrids, and begins to think about what she wants to say …

Looking back on it, Kara had an insight into what makes the Collaboratorium so valuable. The Web has a ton of material on climate change, but it’s so voluminous and fragmented and politicized that it would be a full-time job to figure out what is going on. She realized that what you really want in a situation like this is an unbiased expert: someone who can concisely and clearly lay out all the main issues and options for you, explain the pros and cons of the different approaches, and help you understand who supports which ideas and why. The Collaboratorium represents, in effect, a diverse community coming together to *collectively* create that kind of shared understanding, on an unprecedented web-enabled scale.

Our Approach

We have implemented a prototype web-based system, called the Collaboratorium, which aims to realize the vision described above. In the sections below, we describe how the Collaboratorium is designed, and why we believe this design has the potential to successfully enable large-scale high-quality deliberation for complex systemic challenges like climate change.

System Design

As we have seen, existing technologies display complimentary strengths and weaknesses in terms of their ability to support large-scale, high-quality deliberation. Sharing tools enable large-scale, but unsystematic and mixed quality, idea sharing. Argumentation tools enable systematic and high-quality deliberations, but heretofore only at small scales. And funneling tools enable the large-scale consensus identification that the other tools do not address. We believe that effective large-scale deliberation systems can be realized by *integrating* these three classes of technologies in novel ways. This is the approach we have taken.

The basic design of the system is simple. Users are asked to create a network of articles organized according to argumentation theory. While a range of argumentation approaches have been explored over the years (see section 2), we have chosen to adopt the IBIS formalism (Conklin 2005) because it is simple to use and has been applied successfully in hundreds of collective decision-making contexts. In this formalism, there are four kinds of articles: issues, ideas, pros, and cons. An *issue* represents a question that needs to be answered (e.g. “how can we increase the energy efficiency of personal transportation?”). An *idea* (e.g. “use hybrid gas-electric cars”) represents one possible answer for a given issue. *Pros* and *cons* capture arguments concerning the strengths or weaknesses, respectively, of an idea or other argument. Articles are interlinked to form compact and well-organized “argument maps” that capture the essence of a deliberation.

As with Wikipedia, each article has an associated discussion page which allows people to interact on topics related to that article. These discussion pages, in the Collaboratorium, take the form of threaded discussion forums (i.e. where contributors can add comments to as well as rate other’s comments). Such forums have proven highly successful in such large-scale systems as slashdot, digg, e-the-people, and so on. Discussions are informal and can cover anything related to the article topic, including suggestions on how to improve the article, hotlinks to relevant news items, opinion pieces, requests for advice, and stories about people’s success (or failure) at implementing the idea described in the article. The goal of the discussion forums is to allow people to interact freely using a medium that is already easy and familiar for millions of computer users.

Interleaved with this informal discussion process, users are also able to add new articles (issues, ideas, pros, and cons) to the argument map. In this case, a set of wikipedia-like conventions must be followed to ensure the articles are well-structured and of high quality. Articles should for example be clearly written, and reference reliable sources. Each article should represent a single issue, idea, pro, or con, should not replicate a point that has already been made elsewhere in the knowledge base. It should be attached to the appropriate part of the argument map. Changing an article in order to undermine someone else’s point of view is forbidden: if one disagrees with an idea or argument, the user can capture this by creating *new* articles that present their alternative ideas or counter-arguments. These simple conventions help create a disciplined forum for interacting about complex and controversial topics. Every individual can present their own point of view, using the strongest arguments they can muster, without fear of being sabotaged by anyone else. All the ideas and arguments on a given topic are located near to each other on the argument map so they can be compared directly. The community can register, through such mechanisms as voting or prediction markets, which ideas or lines of argument they find most compelling.

The system is designed so as minimize the barriers to early users, while also allowing more expert users to work efficiently. Users in the initial stages are typically drawn by simple curiosity. At this stage, it is important that the system appear interesting, useful, and accessible. This implies providing a simple intuitive UI, as well as enough content and user activity to give users a sense that the system is alive and well. Nothing cools interest as quickly as the appearance of being a 'dead' site. Users will get positive feedback for their initial forays, e.g. by sending email thanking them for contributing, and by highlighting new content.

Because good argument-mapping skills are not universal, editors with that expertise will be available to ensure that new articles are correctly structured. Experienced authors, with a track record of successful argument map creation, can for example be recruited to join the editor pool.

The interactions that emerge thus consists of two concurrent threads: (1) informal *discussions* centered around articles, and (2) a quality-controlled process of *harvesting* the best ideas from these discussions into the argument map.

This process is supported by capabilities that have proven invaluable for enabling productive interactions in large-scale sharing systems, including:

- *Rating*: users are able to rate the quality of articles (i.e. whether they are clearly-written and well-referenced). This helps the community distinguish "the wheat from the chaff", encourages higher-quality articles, and provides an important source of information concerning author credibility.
- *Voting:* users can indicate, either in binary (yes-no) or prediction market (stock price) form, which ideas they prefer. The aggregation of these votes allows the community to get a sense of which ideas are most widely and strongly supported.
- *Home pages*: every user has a customizable home page which lists which articles and comments they have contributed or endorsed, as well as the people they are connected with. This allows people to develop, if they wish, an on-line presence, facilitating reputation-building, networking and community-building, and makes it straightforward to find people who are experts on different subjects.
- *Version histories*: all articles include a history of all the versions they have gone through. This allows one to "roll back" an article quickly to a previous version.
- *Watchlists*: users are able to specify which articles or comments they are interested in, so they can be automatically notified when any changes are made thereto. When coupled with easy rollbacks, this helps make the knowledge-base "self-healing" – if an post is compromised in some way, this can be detected and repaired rapidly.
- *Search*: users are able to do keyword-based search for comments and articles, with optional filtering that limits matches to recently changed articles, articles in a given branch of the argument map, and so on. This helps people quickly find articles on topics that they are interested in reading, commenting on, rating, or adding to.

The Collaboratorium explores new territory in terms of collaborative system design, and as a result raises novel design challenges that will need to be resolved. These include:

Understanding complex argument maps: Since climate change is a complex topic, we can expect that the argument maps will grow to be substantial in size. How can we help non-expert users quickly get a sense of what information is available in the system? Argumentation systems have traditionally used graphical displays. Such displays are good for small argument maps, since entries will tend to have stable visual locations that make it easier to navigate quickly among them, but this advantage is lost when the number of entries increases and older entries frequently have to be moved, or collected into sub-maps, in order to make space for new entries. They also have a low information density, and thus rapidly become unwieldy as the argument map grows. We plan to explore the use of high-density textual displays to address these concerns, and to develop tools that help automatically summarize key features of the argument map (e.g. by identifying which areas are currently most controversial, or by converting the logical structure of an argument map into the kind of linear narratives most people find familiar and comfortable).

Access privileges: Who can edit which posts when? This question has been handled differently in different sharing systems. Wikis, for example, typically allow anyone to change anything, where the last writer "wins". In threaded discussion forums, every post has a single author; people can comment on but not modify submitted posts. The wiki model is helps ensure that diverse perspectives are incorporated and errors corrected. But it also has some weaknesses. Uninformed authors can overwrite expert contributions, which can discourage content experts from participating. Credit assignment for good articles is muddied because of the open authorship, making it harder to identify who is expert in which domains. And controversial topics can lead to edit wars as multiple authors compete to give their own views pre-eminence in a post. The forum model, by contrast, encourages expert commentary, but the community has much less opportunity to impact the quality of a post. This may be a less critical problem, however, for an argumentation system because, since each post is intended to capture just one perspective, rather than everyone's. We plan to explore a range of design alternatives along this continuum, wherein only the author, the author and his/her assigned proxies, or the whole community, can edit a post.

Ensuring a high quality argument structure: In an open system, we can expect that many participants will not initially be experts on argument mapping. They may fail to properly "unbundle" their contributions into their constituent issues, ideas, and arguments, or may link them to the wrong posts. Different people may also conceptually divide up a problem space differently, leading to the possibility of multiple competing argument maps covering the same subject. The sheer volume of postings may make this redundancy less than obvious. There are many possible approaches to addressing this problem. We can rely, for example, on users to correct poor argument mapping post-hoc. Alternatively, we can require that all new posts be "certified" by editors (experienced users with proven argument mapping skills) before they become "published" (i.e. generally visible in the argument map). Authors can recruit editors to help them author new articles. Alternatively, editors can take the initiative and periodically search for promising (e.g. highly-rated) comments in their areas of interest, and formalize them into the argument map. Perhaps argument mapping for a particularly rich comment could be crowd-sourced (i.e. divvied up among multiple editors). We will explore the possibility that making activity information explicit (e.g. by having more frequently viewed parts of the argument map appear in a larger font) will help the user community converge on a consensus argument map. Given a choice, users will presumably want to locate their posts in the argument branch that is most frequently viewed, because that maximizes their opportunities to be seen and endorsed. This should produce a self-reinforcing push towards consolidation in the argument maps. We are also exploring the idea of relying upon a small cadre of domain experts to create an initial argument map carefully designed to present the critical issues and options in an intuitively organized way. This "skeleton" will help ensure that users put new posts in a consistent place, and can then be fleshed out and, if necessary, modified by the full user community (see Appendix "Climate Change Issue Structure" for an initial version of this). Finally, we will explore the use of "argument schemes" (Walton 2005) to help ensure the argument maps cover important issues. An argument scheme represents a prototypical kind of argument, such as a syllogism, a reference to authority, or an analogy. Each scheme has different critical questions. Any reference to authority, for example, should consider whether that authority is an expert on the topic addressed and whether he/she has a potential conflict of interest. Users will be asked to classify which argument scheme they are using, and the system will prompt users to consider the critical questions relevant to that scheme (Iandoli and Klein 2007).

Mediating attention sharing: In small-scale face-to-face meetings, it is relatively straightforward to ensure that the group devotes the bulk of its energy to resolving the issues that are considered most important and to exploring the ideas that seem most promising. Facilitators often play a key role in this. In large-scale distributed systems with no one "in charge", however, there is no guarantee that this will happen. The community may either converge prematurely (i.e. spend too much time arguing about which ideas to select before a broad enough palette of ideas has been put on the table) or fail to converge at all (i.e. get stuck considering marginal issues or ideas). This problem is exacerbated by the fact that argument maps for complex problems define not just a few solution ideas, but rather an entire (and generally vast) *design space*. Imagine, for example, that we are designing a cap-and-trade system to reduce the emission of greenhouse gases. One issue may concern *how* the emissions certificates are allocated: are they, for example, sold on an open market, or are they allocated by the government? Another issue concerns how *many* certificates are allocated. These issues are *orthogonal*: one can imagine different solutions that combine different decisions re: certificate allocation and volume. The issues choices will often also be *interdependent:* the best choice for allocation strategy may be affected for example by whether certificates are numerous or scarce. This means that independently selecting the 'best' option for each issue is unlikely to result in the best overall solution. We therefore need to consider different *combinations* of ideas, to see which ones are compatible and optimize the outcomes. The resulting decision space may be too large to practically enumerate: an argument tree with only 10 orthogonal issues and 10 (non-mutually-exclusive) options per issue produces, for example, (2^10)^10 (over 10^30) possible solution options. This is not a 'problem' with argumentation systems, but rather a result of their ability to represent the inherent complexity of systemic problems. How can we help the user community identify the most critical issues and best ideas (or idea combinations) as rapidly as possible? We plan to address these challenges by making it easy for the user community to see where attention is currently being put, as well as to register and aggregate their opinions about what areas warrant further attention i.e. applying wisdom of the crowds to attention sharing. The argument map display, for example, can use font size to make highly active posts visually salient (more active posts are larger). This may help ensure that "fertile" parts of the argument map (i.e. ones where people are generating lots of new content) are more likely to get attention and thereby be "exploited" rapidly. We can provide a "hot list", where community members can post pointers to issues ideas or arguments they think warrant more attention, with the most popular posts appearing on top. Rating scores can be propagated through the pro/con chains attached to an idea, in order to give a numeric measure of how much the community likes that idea (Iandoli and Klein 2007). All these possibilities contribute to providing "social translucence" i.e. helping members of the user community get a more acute sense of what other community members are doing, and why (Erickson, Halverson et al. 2002). It may also prove useful to apply distributed nonlinear optimization techniques to help guide the community through the design space towards win-win solutions (Klein, Faratin et al. 2003).

Ensuring good collective decisions: Once the community has done its best to identify the most promising ideas for responding to a given issue, the final step is for them to decide which idea(s) they will select for adoption. This implies at least two key challenges: (1) how do we ensure that individual's choices are as well-founded as possible, and (2) how do we ensure that these individual judgments are "funneled" (e.g. by voting) to maximize the "collective intelligence" of the community's decision? The rich semantics captured by an argument map hold promise of addressing these challenge in new powerful ways. In particular, it is straightforward to run a "consistency check" wherein the ideas a user votes for are vetted against the arguments that the user had rated highly. If a user selects an idea that is not well-supported by the arguments he or she found most compelling, that suggests that the user is using bias rather than evidence and logic to make the selection, and may motivate the user to engage in further reflection, for example by re-considering the ratings, votes, arguments and endorsements he or she made. Another possibility is to introduce proxy voting (Rodriguez, Steinbock et al. 2007), wherein users transfer their vote on a given issue to other users whose judgment they trust in that area. The intent, of course, is to ensure that votes reflect the best judgment of the community, but this process may have other emergent properties, both positive and negative. For example, organizations may be incented to recruit Collaboratorium users in order to get more proxy power, thereby growing the community towards critical mass. On the other hand, users may too readily hand their decision-making over to a relatively small number of individuals, potentially undermining the very collective effects that the Collaboratorium attempts to achieve. This will need to be assessed. Other possibilities to be explored include weighted voting (where people can allocate points to ideas, instead of simply selecting one), as well as prediction markets.

Ensuring community robustness: Large-scale deliberation implies a degree of open-ness that makes the community potentially vulnerable to incompetence or even sabotage by its users. Some potential risks are, of course, moderated by the structure of the Collaboratorium itself. Edit wars, for example, are much less likely in a model where each post has controlled authorship, and where contrarian ideas can appear in their own posts. But other risks remain. Users may attempt, for example, to "stack the deck" by creating multiple posts that all express the same point, or by providing highly biased ratings in an attempt to manipulate the collective decision outcome. The community may become balkanized, wherein groups of users only attend to content and other users they already agree with. The editors who are charged with maintaining the argument map structure may be biased or incompetent or simply miss problems due to the sheer size of the argument map. Addressing such risks will require defining a compelling and effective shared ethos as well as a set of processes and supporting tools. We can, for example, adopt a variant of the Slashdot meta-moderation process to help create and maintain the quality of the editor population. Users can be given rating "budgets", so each can only rate a limited number of posts and thus will be incented to only rate posts they have strong feelings about. Search tools can be provided to flag potential problems like multiple highly similar co-located posts. Social network analysis tools can help uncover balkanization in the discussions. We can adapt time-tested peer-production guidelines to the special needs of large-scale argumentation systems. Wikipedia's famous "neutral point of view" guideline ("articles should not be edited to remove significant points of view") will need to be modified, for example, since each Collaboratorium post expresses just one point of view. A possible replacement is the rule: "articles should not be edited to weaken the point of view expressed therein".

Why This Will Work

Experience with sharing systems has shown that relatively simple technologies can enable remarkably powerful emergent phenomena at unprecedented scales (Tapscott and Williams 2006) (Sunstein 2006) (Surowiecki 2005) (Gladwell 2002). These phenomena help user communities generate high-quality content quickly and effectively:

- *Idea synergy*: the ability for users to share their creations in a common forum (e.g. youtube, flickr) can enable a *synergistic explosion of creativity*, since people often develop new ideas by forming novel combinations and extensions of ideas that have been put out by others.
- *The long tail*: sharing systems give individuals access to a much *greater diversity* of ideas than they would otherwise: "small voices" that would otherwise not be heard can now have significant impact.
- *Many eyes*: open-source efforts (e.g. Wikipedia and Linux) can produce remarkably *high-quality* results, even with unconstrained participation, by virtue of the fact that there are multiple independent verifications - many eyes continuously checking the shared content for errors and correcting them.
- Q*uality spiral*: many sharing systems (e.g. slashot, digg) use collective rating and ensure that highly-rated contributions are more visible than those that are poorly rated. This implicitly fosters a competition among contributors to produce highly-rated posts, and thus drives greater overall value in the contributions being made by the community at large.
- *Wisdom of the crowd*: Large groups of (independent, motivated and informed) contributors can collectively produce assessments (e.g. via voting or prediction markets) that are often far more accurate than those produced by individuals, even experts
- *Many hands make light work*: the sheer scale of voluntary participation enabled by social software systems has meant that even very large tasks can be taken on and achieved remarkably quickly and cheaply.
- *Small worlds*: social software systems have made it qualitatively easier for people to find others with related interests, irregardless of physical separation, enabling the creation of interest-focused communities that would have previously been too small and dispersed to come into being.
- *Economies of scale*: the large scale of social software systems has meant that it has become reasonable for users to invest substantial resources to create tools that result in substantial

efficiency benefits once implemented. Wikipedia editing robots (software that automatically scans articles to correct problems like spelling mistakes) are an example of this.

- *Ad hoc organizations*: social software reduces coordination costs drastically, so that users can routinely create highly-distributed project-specific work groups, perform the task at hand, and disband, without the need for long-standing formalized organizational affiliations. This greater agility means that more collaborative tasks can be performed more quickly and effectively, with lower fixed costs.
- *Matchmaking:* idea marketplaces (e.g. innocentive.com) allow quick and low-cost matchmaking between people with problems to solve and people with the skills potentially relevant to solving them, increasing the likelihood that the right match will be made and a superior solution found. This phenomenon is especially powerful because useful solutions often come from unexpected directions and disciplines.

All of these properties of sharing-type systems are clearly highly relevant to enabling large-scale high-quality deliberation.

Current sharing systems do face, however, the potential weaknesses listed in section 2 above. This is where argumentation technology steps in:

- *Low signal to noise ratio*: Argument maps provide a compact well-structured way to capture key issues ideas and arguments. It is easy to find the right place to add new stuff, and therefore easy to ensure that each unique issue idea and argument appears just once. This logical as opposed to chronological structure thus enables a qualitatively higher signal-to-noise ratio.
- *Unsystematic coverage*: With an argument map, it is easy to assess which issues have been explored in depth (have many associated sub-issues, ideas or arguments) and which have not, so it is easy for contributors to find gaps in the argument map and fill them in, bottom up.
- *Dysfunctional argumentation*: The pro/con-based argument scheme implicitly enforces logic/evidence-based reasoning. Collective rating ("wisdom of the crowd") can help distinguish good from bad argumentation and encourage a quality spiral in argumentation quality and critical thinking skills. Flame and edit wars can be avoided because the system enables people with divergent opinions to each create their arguments as strongly as possible in parallel without needing to compete for a single thread of attention or a single Wikipedia page. Noisy people have no way to monopolize the discussion, because the argument map is only updated when they make new points, as opposed to simply repeating old ones.
- *Balkanization:* Because all ideas and arguments relevant to an issue are co-located, inputs from different communities become visible to each other. There is no guarantee, of course, than any individual will read through the ideas and arguments, but at least he or she will easily be able to see what those differing opinions are.

Funneling systems complete the picture by enabling aggregation of individual contributions into a collective judgment. Every issue in the argument map represents a candidate for a vote or prediction market. Each idea represents a proposal (for a voting system), or a "stock" (for a prediction market).

The key uncertainty with our approach concerns whether argument mapping will work with large scale communities. Previous argumentation systems, as we have seen, have been applied almost exclusively in small-scale settings, with dedicated facilitators, because participants often find argument mapping burdensome. But if we are to scale up to thousands of users, we need to distribute the mapping activity across the user population. Will users take on this task? That remains an open question, for now, but there are good reasons to believe that users will find the incentives for argument mapping far more compelling in large-scale settings than small-scale ones.

One reason concerns communication styles. In small groups, relationship management is usually primary, so participants often prefer to communicate implicitly and indirectly, especially at first, in order to avoid unnecessary conflicts. In this context, a technology like argument mapping that forces people to make their arguments explicit can actually be a liability, rather than an asset. But this effect flips on its head at large scales. Implicit communication becomes unnecessary (and often confusing) when most people do not know each other personally, or can float potentially contentious proposals under aliases. So we can expect that resistance to argument mapping will decrease as the user community grows in size.

Other reasons concern the nature of incentives in sharing systems. It has been found (Hars and Ou 2002) (Lakhani and Wolf 2005) (Roberts, Hann et al. 2006) (Bonaccorsi and Rossi 2004) that users contribute to such systems predominantly for two reasons[2]:

- *Finding your tribe:* For many contributors, the predominant value of sharing systems is how they allow them to find, and be found, by a community of people they share interests and values with and therefore feel they can connect with. People contribute as a way of making themselves known. This is typically the “hook” that converts a casual browser into a regular contributor.
- *Becoming a hero:* It is well-known that contributions to sharing systems typically follow a power law: there is a relatively small number of heavy contributors (who contribute the bulk of the content as well as performing editing and management tasks), plus a larger crowd that makes more modest contributions. Heavy-duty contributors seem to be motivated largely by the desire to have a substantial positive impact. Enabling this seems to rely on two key factors: *shared norms and values*, so users believe in the goals of the system (a compelling mission statement, such as wikipedia’s “a free high-quality encyclopedia in everyone’s native language”, can be enormously helpful in this regard), as well as continual *opportunities to contribute* to meaningful tasks (where the need for help is visible, compelling, do-able in scale, appropriate to user skills, and users believe that successful contributions will have a positive impact).

[2] There are of course many other possible motivations for contributing to a sharing system, such as developing your real-life reputation, advertising goods or services, learning about a subject by contributing to it, or simply because you will personally benefit if the product you are working on (e.g. a software application) comes into existence. Most contributors to sharing systems, however, apparently do it largely for social and altruistic reasons.

These incentives have proven enormously powerful in peer-production efforts such as Wikipedia and Linix, and are highly relevant to motivating argument mapping. There is widespread disaffection with the low signal-to-noise ratio of current sharing tools. It seems clear that the number of distinct issues, ideas, and arguments in a discussion will grow, after a certain point, much more slowly than the number of participants. The larger the user community, therefore, the more potential redundancy there will be, the more value argument mapping offers in terms of improving the signal to noise ratio. We can thus expect that, as the system grows, users will increasingly recognize and respond to the opportunity to "become a hero" by contributing something (i.e. summarizing key points in the argument map) that is highly valued by the community. Authors will be motivated to make their postings in argument map form because it increases the likelihood that their contributions will be widely read, helping them "find their tribe". While contributing to an unstructured discussion is easy, the high volume and redundancy of such forums means that most posts will probably be overlooked by most readers. In an argument map, by contrast, if you have a unique point to make, it will become a part of a structure with a much higher signal-to-noise ratio, and thus a much greater chance of being read.

Another question concerns whether the user community will be able to provide sufficient *numbers* of people (what we call "editors") with good argument mapping skills. There are several reasons to believe this will be possible. There is already a substantial world-wide community of people with argument mapping skills. One organization alone (cognexus.org) has already trained and certified hundreds of people in the IBIS methodology that underlies the Collaboratorium. Other similar argument mapping tools, such as debatemapper.com, have their own user communities that could also contribute. Argument mapping is, in addition, a natural skill for lawyers, philosophers, mathematicians, library scientists, and others who create proofs or ontologies for a living. Such individuals can probably learn an IBIS-like formalism very quickly, and may be inspired by the opportunity to contribute their critical thinking skills to something as critical as the climate change debate, even if they do not have content expertise in that area. Another point is that, as the system scales, we can assume that the argument-mapping burden *per editor* will decrease. This is because the number of mapping-savvy users should scale linearly with the overall user population, but the number of novel ideas that need to be mapped should scale (as we mentioned above) *less* than linearly. If the Collaboratorium works like most peer-production systems, user contributions will follow a power law, so we can expect that a relatively small corps of "power users" will take on the bulk of the argument mapping tasks. Most people will just have to read or comment on, as opposed to create, argument maps.

We can summarize all this by reviewing how well the Collaboratorium design addresses the requirements for large-scale high-quality deliberation that we identified in section 2:

- All important issues are considered
- The broadest possible range of high-quality solution ideas are identified.
- The strongest arguments for and against each idea are captured.
- People can distinguish good from bad arguments.
- The solution ideas that individuals eventually select are chosen rationally.
- The aggregate results constitute a good representation of the "wisdom of the people".

The Collaboratorium helps ensure, through such phenomena as idea synergy, the long tail, and many eyes, that the user community creates a comprehensive collection of relevant issues, ideas, and arguments. The use of argument maps to structure this material helps ensure systematic coverage as well as clear and succinct argumentation, with competing ideas and arguments co-located and directly comparable. The quality spiral helps ensure that good arguments are more salient and thus more common. People can use the argument maps to verify whether the ideas they favor are indeed backed up by the arguments they find most compelling. Funneling techniques like voting and prediction markets provide, finally, time-tested ways to aggregate individual contributions into collective wisdom.

Conclusion

The key innovations of this work involve exploring how ideas that have proven successful in large-scale open source/peer production systems can be melded with argumentation theory to enable effective large-scale deliberation. To our knowledge, this kind of large-scale argumentation approach has not been previously explored. The Collaboratorium has been evaluated in a range of increasingly large-scale settings, the latest of which involving 200 users collaborating over a two week period on the topic of whether and how biofuels should be adopted in Italy. Out future work will focus on (1) understanding what it takes to make the system capable of enabling high-quality self-sustaining deliberations with thousands of users, (2) deploying the system so it has a positive impact on key systemic challenges such as climate change.

Acknowledgements

I would like to gratefully acknowledge the many valuable contributions, to the ideas described herein, made by Marco Cioffi, Luca Iandoli, Thomas Malone, Kara Penn and David Price.

References

Bonaccorsi, A. and C. Rossi (2004) "Altruistic individuals, selfish firms? The structure of motivation in Open Source software." Volume, DOI:

Chklovski, T., V. Ratnakar and Y. Gil (2005). "User interfaces with semi-formal representations: a study of designing argumentation structures." Proceedings of the 10th international conference on Intelligent user interfaces: 130-136.

Conklin, J. (2005). Dialogue Mapping: Building Shared Understanding of Wicked Problems, John Wiley and Sons, Ltd.

Erickson, T., C. Halverson, W. A. Kellogg, M. Laff and T. Wolf (2002). "Social Translucence: Designing Social Infrastructures that Make Collective Activity Visible." Communications of the ACM 45(4): 40-44.

Gladwell, M. (2002). The Tipping Point: How Little Things Can Make a Big Difference, Back Bay Books.

Hars, A. and S. Ou (2002). "Working for Free? Motivations for Participating in Open-Source Projects." International Journal of Electronic Commerce 6(3): 25 - 39.

Heng, M. S. H. and A. de Moor (2003). "From Habermas' s communicative theory to practice on the internet." Information Systems Journal 13(4): 331-352.

Iandoli, L. and M. Klein (2007). Channeling the wisdom of the crowds: how large scale collaborative platforms can effectively leverage collective intelligence. XIV Congress of

International Association for Fuzzy-Set Management and Economy: Computational Intelligence Applied to New Digital Economy and Business Ecosystems. Poiana Brasov – Romania.
Jonassen, D. and H. R. Jr (2005). "Mapping alternative discourse structures onto computer conferences." International Journal of Knowledge and Learning 1(1/2): 113-129.
Karacapilidis, N., E. Loukis and S. Dimopoulos (2004). "A Web-Based System for Supporting Structured Collaboration in the Public Sector." LECTURE NOTES IN COMPUTER SCIENCE: 218-225.
Kirschner, P. A., S. J. B. Shum and C. S. C. Eds (2005). "Visualizing Argumentation: Software tools for collaborative and educational sense-making." Information Visualization 4: 59-60.
Klein, M., P. Faratin, H. Sayama and Y. Bar-Yam (2003). "Protocols for Negotiating Complex Contracts." IEEE Intelligent Systems 18(6): 32 - 38.
Lakhani, K. R. and R. G. Wolf (2005). Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects. Perspectives on Free and Open Source Software. J. Feller, B. Fitzgerald, S. Hissam and K. R. Lakhani, MIT Press.
Lowrance, J. D., I. W. Harrison and A. C. Rodriguez (2001). Capturing Analytic Thought. First International Conference on Knowledge Capture: 84-91.
Moor, A. d. and M. Aakhus (2006). "Argumentation Support: From Technologies to Tools." Communications of the ACM 49(3): 93.
Roberts, J., I.-H. Hann and S. Slaughter (2006). "Understanding the Motivations, Participation and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects." Management science 52(7): 984-999.
Rodriguez, M. A., D. J. Steinbock, J. H. Watkins, C. Gershenson, J. Bollen, V. Grey and B. deGraf (2007). Smartocracy: Social Networks for Collective Decision Making. 40th Hawaii International Conference on System Sciences. Hawaii.
Shum, S. J. B., A. M. Selvin, M. Sierhuis, J. Conklin and C. B. Haley (2006). Hypermedia Support for Argumentation-Based Rationale: 15 Years on from gIBIS and QOC. Rationale Management in Software Engineering. A. H. Dutoit, R. McCall, I. Mistrik and B. Paech, Springer-Verlag.
Sunstein, C. R. (2006). Infotopia: How Many Minds Produce Knowledge, Oxford University Press.
Surowiecki, J. (2005). The Wisdom of Crowds, Anchor.
Tapscott, D. and A. D. Williams (2006). Wikinomics: How Mass Collaboration Changes Everything, Portfolio Hardcover.
Walton, D. N. (2005). Fundamentals of Critical Argumentation (Critical Reasoning and Argumentation). Cambridge (MA), Cambridge University Press.
Walton, D. N. and E. C. W. Krabbe (1995). Commitment in dialogue: Basic concepts of interpersonal reasoning. Albany, NY, State University of New York Press.