The Mit Deliberatorium - Enabling Large-scale Deliberation about Complex Systemic Problems.

Conference Paper · May 2011

DOI: 10.1109/CTS.2011.5928678 · Source: DBLP

1 author:

Mark Klein
Massachusetts Institute of Technology
426 PUBLICATIONS **8,221** CITATIONS

Using Metrics to Enable Large-Scale Deliberation

Mark Klein

Center for Collective Intelligence
Massachusetts Institute of Technology
77 Massachusetts Avenue, NE25-754
Cambridge MA USA 02139
m_klein@mit.edu

Abstract. Existing social media (such as such as email, blogs, wikis, chat rooms, and web forums) provide unprecedented opportunities for interacting on a massive scale, but have yet to realize their potential for helping people *deliberate* effectively around complex and contentious topics, typically generating poorly-organized, unsystematic and highly redundant contributions of widely varying quality. Large-scale argumentation systems represent a promising approach for addressing these challenges, by virtue of providing a simple systematic structure that radically reduces redundancy and encourages clarity. They do, however, raise an important challenge. How can we ensure that the attention of the deliberation participants is drawn to where it can best serve the goals of the deliberation? This paper describes how novel forms of metrics can help address this critical problem.

Keywords: deliberation; metrics; argumentation

1 The Challenge

Humanity now finds itself faced with a range of highly complex problems – such as climate change, the spread of disease, international security, scientific collaborations, product development, and so on - that call upon us to bring together large numbers of experts and stakeholders to deliberate collectively on a global scale. Collocated meetings can however be impractically expensive, severely limit the concurrency and thus breadth of interaction, and are prone to serious dysfunctions such as polarization and hidden profiles [1]. Social media such as email, blogs, wikis, chat rooms, and web forums provide unprecedented opportunities for interacting on a massive scale, but have yet to realize their potential for helping people *deliberate* effectively, typically generating poorly-organized, unsystematic and highly redundant contributions of widely varying quality. Large-scale argumentation systems represent a promising approach for addressing these challenges, by virtue of providing a simple systematic structure that radically reduces redundancy and encourages clarity. They do, however, raise an important challenge. How can we ensure that the attention of the deliberation participants is drawn, especially in large complex argument maps, to where it can best serve the goals of the deliberation? How can users, for example, find

the issues they can best contribute to, assess whether some intervention is needed, or identify the results that are mature and ready to "harvest"? Can we enable, for large-scale distributed discussions, the ready understanding that participants typically have about the progress and needs of small-scale, collocated discussions?

This paper will address these important questions, discussing (1) the strengths and limitations of current deliberation technologies, (2) how large-scale argumentation can help address these limitations, and (3) how we can use novel deliberation metrics to enhance the effectiveness of deliberations mediated by argumentation systems.

2 Review of existing Deliberation Technologies

Let us define deliberation as a process where communities (1) identify possible solutions for a problem, and (2) select the solution(s) from this space that best meet their diverse needs [2] [3]. How well do existing technologies meet this challenge?

A wide range of social computing technologies have emerged in the past few decades, including email, chat, web forums, wikis like wikipedia, media sharing sites like youtube and flickr, open source software development efforts such as Linux, solution competitions such as Innocentive.com, idea-sharing systems such as ideastorm.com, peer-filtering sites such as Slashdot, group decision support (GDSS) systems [4] [5] [6] [7] [8], and scientific collaboratories [9]. Experience with such systems has shown that they foster, by virtue of reducing the cost of participation, voluntary contributions at a vast scale, which in turn can lead to remarkably powerful emergent phenomena [10] [1] [11] [12] that include:

- *Idea synergy*: the ability for users to share their creations in a common forum 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*: social computing systems enable access to a much *greater diversity* of ideas than they would otherwise: "small voices" (the tail of the frequency distribution) that would otherwise not be heard can now have significant impact.
- *Many eyes*: social computing efforts can produce remarkably *high-quality* results by virtue of the fact that there are multiple independent verifications - many eyes continuously checking the shared content for errors and correcting them.
- *Wisdom of the crowds*: large groups of (appropriately independent, motivated and informed) contributors can collectively make better judgments than those produced by the individuals that make them up, often exceeding the performance of experts, because their collective judgment cancels out the biases and gaps of the individual members.

To understand the strengths and limitations of these technologies, it is helpful to divide them up based on how they structure content. One category is **time-centric** tools, i.e. tools like email, chat rooms, and web forums where content is organized based on *when* a post was contributed. Such systems enable large communities to

weigh in on topics of interest, but they face serious shortcomings from the perspective of enabling collective deliberation [1]:

- *Scattered content:* The content in time-centric tools is typically widely scattered, so it's hard to find all the contributions on a topic of interest. This also fosters *unsystematic coverage*, since users are often unable to quickly identify which areas are well-covered, and which need more attention.
- *Low signal-to-noise ratio*. The content captured by time-centric tools is notorious for being voluminous and highly repetitive. This is a self-reinforcing phenomenon: since it can be difficult to find out whether a point has already been made in a large existing corpus, it's more likely that minor variants will be posted again and again by different people. Some authors may do so simply hoping to win arguments by sheer repetition. This low signal-to-noise ratio makes it difficult to uncover the novel contributions that inspire people to generate creative new ideas of their own.
- *Balkanization:* Users of time-centric 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 blogs and forums – so they tend to see only a subset of the issues, ideas, and arguments potentially relevant to a problem. This tends to lead people to take on more extreme, but not more broadly informed, versions of the opinions they already had.
- *Dysfunctional argumentation*: Time-centric 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.

Enormous effort is typically required to "harvest" the corpuses created by time-centric tools to identify the most important issues, ideas, and arguments. Intel, to give a typical example, ran a web forum on organizational health that elicited a total of 1000 posts from 300 participants. A post-discussion analysis team invested over 160 person-hours to create a useful summary of these contributions (at 10 minutes a post, probably longer than it took to write many of the posts in the first place). The team found that there was lots of redundancy, little genuine debate, and few actionable ideas, so that in the end many of the ideas they reported came from the analysis team members themselves, rather than the forum[1].

It could be argued that many of these concerns are less prominent in **topic-centric** tools such as *wikis* and *idea-sharing* systems. In wikis, for example, all the content on a given topic is captured in a single article. But wikis are deeply challenged by deliberations on complex and controversial topics [13] [14]. They capture, by their nature, the "least-common-denominator" *consensus* between many authors (any non-consensus element presumably being edited out by those that do not agree with it), and the controversial core of deliberations are typically moved to massive talk pages for the article, which are essentially time-centric venues prone to all the limitations we noted above. *Idea-sharing* tools – such as Dell's Ideastorm.com, the Obama

[1] Based on personal communication with Catherine Spence, Information Technology Enterprise Architect, Computing Director/Manager at Intel.

administrations' Open for Questions web site, and Google's project10tothe100.com - are organized around *questions*: one or more questions are posted and the community is asked to contribute, rate, and comment on proposed solutions. Such sites can elicit huge levels of activity – the Obama site for example elicited 70,000 ideas and 4 million votes in three weeks – but they are prone to several serious shortcomings. One is *redundancy:* in all of these sites, many of the ideas represent minor variations of each other. When there are thousands of posts submitted, manually pruning this list to consolidate equivalent posts is a massive undertaking. In Google's case, for example, the company had to recruit 3,000 employees to filter and consolidate the 150,000 ideas they received in a process that put them 9 months behind their original schedule. Another issue is *non-collaborativeness*. Idea-sharing sites tend to elicit many fairly simple ideas. The ideas generated by the google project, for example, (e.g. make government more transparent, help social entrepreneurs, support public transport, create user-generated news services) were in large part not novel and light on detail. Surely that massive amount of effort could have been used to compose a smaller number of more deeply-considered ideas, but idea-sharing sites provide little or no support (or incentive) for this, because people can not collaboratively refine submitted ideas.

3 Large-Scale Argumentation

Large-scale argumentation represents a promising approach to addressing the weaknesses with current deliberation technologies. We describe this approach below.

Argumentation tools [15] [16] [17] take an **argument-centric** approach based on allowing groups to systematically capture their deliberations as tree structures made up of *issues* (questions to be answered), *ideas* (possible answers for a question), and *arguments* (statements that support or detract from an idea or argument) that define a space of possible solutions to a given problem:

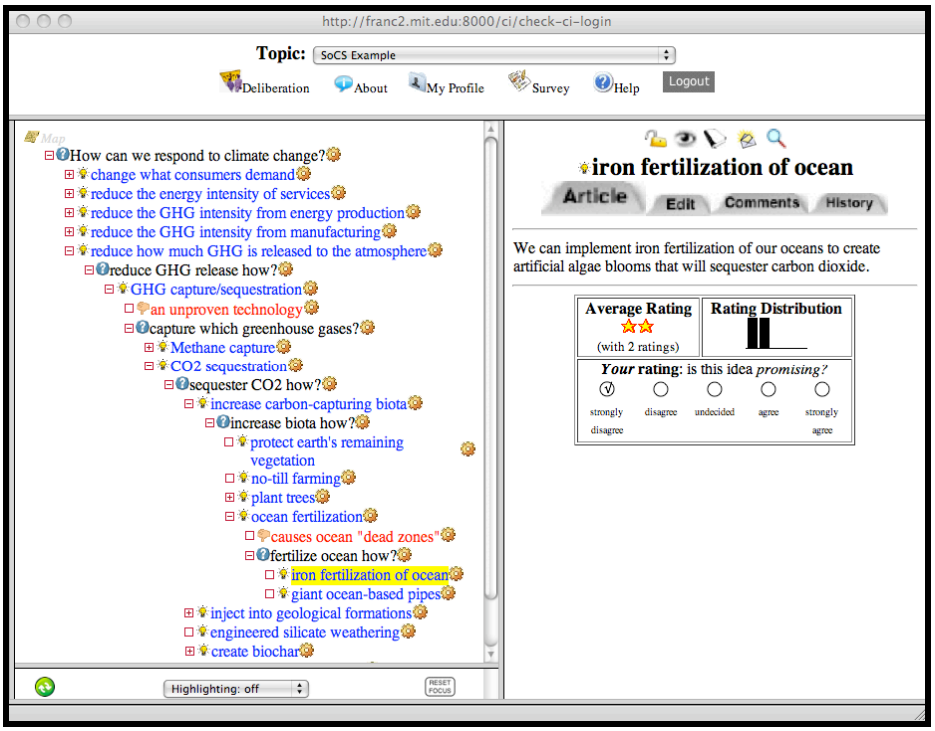

Fig. 1. A screenshot from the Deliberatorium, a large-scale argumentation system.

Such tools have many advantages. Every unique point appears just once, radically increasing the signal-to-noise ratio, and all posts must appear under the posts they logically refer to, so all content on a given question is co-located in the tree, making it easy to find what has and has not been said on any topic, fostering more systematic and complete coverage, and counteracting balkanization by putting all competing ideas and arguments right next to each other. Careful critical thinking is encouraged, because users are required to express the evidence and logic in favor of the options they prefer [18], and the community can rate each element of their arguments piece-by-piece. Users, finally, can collaboratively refine proposed solutions. One user can, for example, propose an idea, a second raise an issue concerning how some aspect of that idea can be implemented, and a third propose possible resolutions for that issue. The value of an argument map can extend far beyond the deliberation it was initially generated for, because it represents an entire *design space* of possible solutions that can be readily harvested, refined and re-combined by other communities facing similar problems.

Most argumentation systems have been used by individuals or in small-scale settings, relying in the latter case on a facilitator to capture the free-form interactions of a collocated group in the form of an commonly-viewable argument map [19]. Argumentation systems have also been used, to a much lesser extent, to enable

distributed deliberations over the Internet [20] [21] [22] [23] [24] [25]. These maps tend to be poorly structured, however, because many users are not skilled argument mappers, and the scale of participation has been small[2], typically involving only a handful of authors on any given task.

The author and his colleagues have investigated, over the past several years, how an argument-centric approach can be extended to operate effectively at the same large scales as other social computing systems. Our approach is simple. Users are asked to create, concurrently, a network of posts organized into an argument map. We use the IBIS argumentation formalism [26] because it is simple and has been applied successfully in hundreds of collective decision-making contexts. A set of community conventions (similar to those that underlie other social computing systems like Wikipedia and Slashdot) help ensure that the argument map is well-organized. Each post should represent a single issue, idea, pro, or con, should not replicate a point that has been made elsewhere in the argument map, and should be attached to the post it logically refers to. A central tenet is the "live and let live" rule: if one disagrees with an idea or argument, the user should not change that post to undermine it, but should rather create *new* posts that present their alternative ideas or counter-arguments. Every individual can thus present their own point of view, using the strongest arguments they can muster, without fear of sabotage by anyone else. This process is supported by capabilities that have proven invaluable in other social computing systems, including *rating* (to help the community encourage and identify important issues, ideas and arguments), *watchlists* (which automatically notify users of changes to posts they have registered interest in), *version histories* (to allow users to roll-back an post to a previous version if it has been "damaged" by an edit), and *home pages* (which allows users to develop an online presence). The system also provides multiple forms of *social translucence* [27] (i.e. visual cues concerning who is doing what in the system), thereby fostering a sense of belonging as well as enabling self-organized attention mediation by the community. See [28] for further discussion of the issues underlying the design of large-scale argumentation capability. The system itself is accessible at http://franc2.mit.edu/ci/.

Because good argument-mapping skills are not universal, *moderators* help ensure that new posts are correctly structured. Their job is part education, and part quality control. Posts, when initially created, are given a "pending" status and can only be viewed by other authors. If a post doesn't adequately follow the argument map conventions, moderators will either fix it or leave comments explaining what needs to be done. Once a moderator has verified that a post follows the conventions, the post is "certified" and becomes available to be viewed, edited, commented on, or rated by the general user population. The certification process helps ensures well-structured maps, and provides incentives for users to learn the argument formalism. Moderators serve as honest brokers in all this: their role is not to evaluate the merits of a post, but

[2] 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.

simply to ensure that the content is structured in a way that maximizes its utility to the community at large.

We have implemented an initial version of these ideas, in the form of a web-based tool called the Deliberatorium [29] [30, and evaluated it to date with over 700 users deliberating on a wide range of topics. The largest evaluation was performed at the University of Naples with 220 masters students in the information engineering program, who were asked to use the system to deliberate, over a period of three weeks, about the use of bio-fuels in Italy [Klein, 2008 #4691]. We observed a very high level of user participation: all told, the students posted over 3000 issues ideas and arguments, in addition to 1900 comments. This is, to our knowledge, both the largest single argument map ever created, as well as (by far) the largest number of authors for a single argument map. Roughly 1800 posts were eventually certified, and about 70% of all posts could be certified without changes, demonstrating that, even after a relatively short usage period, most authors were able to create properly-structured posts. The certification ratio, in addition, increased over the duration of the experiment. The breadth and depth of coverage was, in the judgment of content experts, quite good: this community of non-experts was able to create a remarkably comprehensive map of the current debate on bio-fuels, complete with references, exploring everything from technology and policy issues to environmental, economic and socio-political impacts. We estimated, based on this experience, that there needs to be about 1 moderator for every 20 active authors, to ensure that posts are checked and certified in a timely fashion without undue burden on each moderator. This figure is well within the bounds of the percentage of "power users" that typically emerge in social computing user communities.

Other evaluations (including a deliberation with 120 students at the University of Zurich, with 73 users at Intel, and with 40 users at the US Federal Bureau of Land Management), have explored the efficacy of our large-scale argumentation tool for a range of topics and incentive structures. These evaluations support the idea that large-scale argumentation can be applied effectively to complex challenges. Substantial user communities with no initial familiarity with argumentation formalisms have been able, in a range of contexts, to rapidly create substantive, useful, compact, and well-organized maps on complex topics, while requiring levels of moderator effort much lower than those needed to harvest, post-hoc, discussions hosted by such conventional social computing tools as web forums.

Our mathematical analyses show that the per-moderator burden, as well as the cost-benefit ratio for authors, should decrease substantially as the user community grows, suggesting that the incentives will be especially compelling for larger-scale problems.

While these results are promising, our work has led us to conclude that, to fully realize argumentation technology's potential for supporting large-scale deliberations, we need to address the critical challenge of *attention allocation*. For the kinds of topics that most require large-scale deliberation, even a moderately large user community can quickly generate large and rapidly growing argument maps. How can we help users identify the portions of the map that can best benefit from their contributions, in maps that covers hundreds of topics? How can the stakeholders for such deliberations assess whether the deliberations are progressing well, whether some intervention is needed to help the deliberations work more effectively, and when

the results are mature and ready to "harvest"? Can we foster, for large-scale deliberations, the understanding that participants in small-scale discussions typically have about where the discussion has gone, what remains to be addressed, and where they can best contribute. Without this kind of big picture, we run the risk of severely under-utilizing the collective intelligence potentially provided by large-scale social media.

4 Metrics

We can meet this challenge, we believe, by developing a set of algorithms that can be used to provide users with a personalized and continuously-updated set of suggestions, based on deliberation metrics, concerning which parts of the argument map they should view, add to or rate, and why:

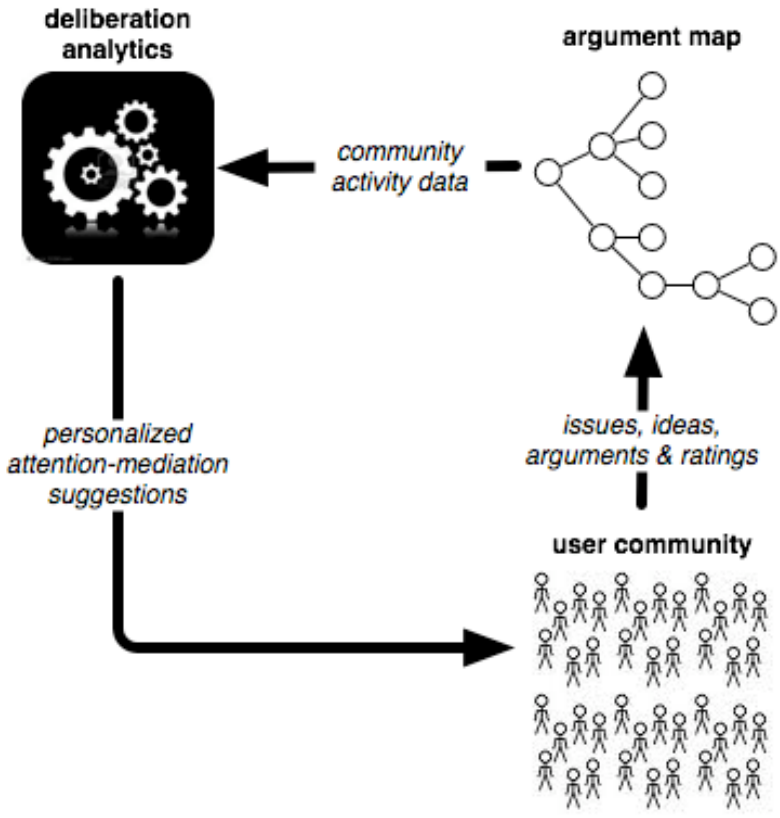

Fig. 2. Using metrics to enable attention mediation in large-scale deliberations.

Each user is free to accept or ignore suggestions as they like, but they know that the suggestions are based in an overview of the deliberation as a whole and are intended to help them apply their unique skills and perspectives to promising regions in the map. If the suggestions are reasonably well-done, the emergent effect is that the collective intelligence of the user community is maximized because each users contributes where they can do the most good.

How can such suggestions be generated? This can be done, we believe, building on a process we call process-commitment-exception analysis [31]. First we define a normative model that specifies what a good large-scale deliberation looks like, including its main steps, commitments, and failure modes (exceptions). Each commitments and exception is then mapped to one or more metrics intended to assess (by analyzing user activity data and the emerging argument map structure) to what extent the commitments is being achieved or, conversely, to what extent the exception is taking place. These metrics values are then mapped, based on a model of the users' roles and interests, into customized suggestions. We discuss these steps in more detail in the paragraphs below.

4.1 A Normative Model of Large-Scale Deliberation

Our normative deliberation model formalizes a straightforward view of what makes up a rational decision-making process. According to this model, deliberation consists of four key steps:

1. Identify the goals the deliberation is trying to achieve
2. Propose possible ways to achieve these goals
3. Evaluate the proposed ideas with respect to the deliberations goals
4. Select the best idea(s) from amongst the proposed solutions

The commitments (❤) and exceptions (⚠) in this model include:

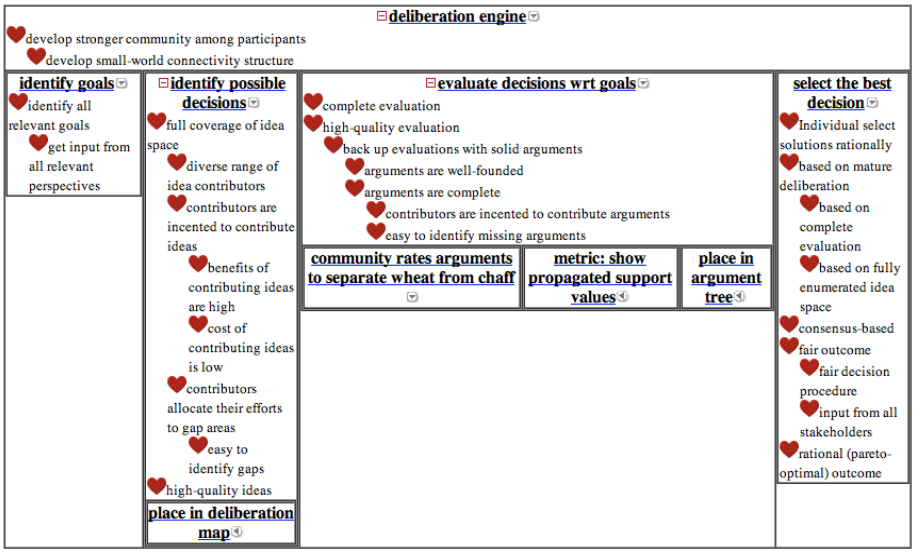

Fig. 3. A (partial) normative model enumerating deliberation commitments and exceptions.

The commitments for the "identify goals" step, for example, include identifying all relevant goals for the deliberation, which in turn is enabled by getting input from all stakeholders for the decision being made.

This generic deliberation model is then elaborated to include sub-steps that specify how these main steps are implemented in the context of a large-scale argumentation system. For example, the commitments "easy to identify gaps" of the step "identify possible decisions" is implemented, in an argumentation system, by capturing possible decisions as idea posts and placing them in the correct part of the argument map so that all the ideas for an issue are grouped together, making it easy to see what has and has not been proposed for that issue. Each of these additional steps may imply additional commitments and exceptions.

4.2 Identifying Metrics

The next step is to identify metrics that can use the information generated during a large-scale argument-centric deliberation to assess whether the deliberation commitments are being achieved, and the potential exceptions are occurring or not. We have identified over 100 possible metrics to date and describe, below, a few illustrative examples, highlighting those that take advantage of the additional semantics provided by an argument map:

- *Balkanization:* balkanization is the phenomenon wherein a community divides itself into sub-groups where members of each group agree with one other but tend to reflexively ignore the inputs of other groups that they do no agree with. This can be viewed as a deliberation dysfunction because it violates the goal

"individuals fully consider the options and tradeoffs" of the "select the best decision" step of our normative deliberation model. The structure of the argument map makes it clear which ideas represent alternatives for a given issue, as well as which arguments support and detract from these ideas, making it straightforward to assess when groups are ignoring the ideas, and supporting arguments, for competing ideas.

- *Groupthink*: groupthink can be defined as occurring when a community prematurely devotes an excessive proportion of its attentional resources to a small subset of the relevant issues, ideas and arguments. This is straightforward to assess in an argument map because we can readily measure when, for example, one idea under an issue is receiving the bulk of the community's attention while competing ideas and their underlying variants and arguments have remain largely untouched.
- *Irrational bias*: we define irrational bias as occurring when a user gives ratings for ideas or arguments that are inconsistent with the ratings they give the underlying arguments. We can use simple techniques to propagate a user's ratings for arguments up to produce a predicted rating for the higher level arguments/ideas, and then compare that with the actual ratings they give these posts.
- *Mature topics*: a mature topic is one where a fairly exhaustive inventory has been made of the relevant ideas and arguments. This can be estimated in a number of ways, including tree topology (more mature topics tend to gave both broader and deeper structures), activity history (argument-centric deliberations tends to transition, over time, from identifying issues to proposing ideas to presenting arguments to ratings posts to quiescence), and so on.
- *Controversial posts*: we can identify controversial posts because one can look not only for posts with many highly divergent ratings, but also for posts that have polarized rating distributions for the underlying arguments. The fact that each post represents a single logical point (issue, idea, or argument) rather than (as is often the case with other social media) a collection of points, means that the ratings give a more accurate picture of the community's assessment of each point.

A large-scale argumentation system requires that users parse their contributions into topically-organized structures of typed, individually-ratable issues, ideas, and argument. This structure provides, as we can see, rich fodder for such powerful techniques as social network analysis, belief propagation, singular vector decomposition, and so on. This in turn makes it possible to define real-time deliberation metrics that, for conventional social media, would require an impractical level of hand-coding for most settings.

4.3 Generating Suggestions

The final step of our approach involves generating suggestions for users concerning which posts they might want to look at in order to contribute most effectively to the deliberation at hand. This is done by identifying, based on a user model, which

metrics a user "should" be interested in, and then drawing their attention to parts of the argument map where these metrics have extreme values. The user's interests can be inferred based on their role, as well as their past activity and that of other members of the community. A topic manager (someone responsible for ensuring a deliberation achieves useful results) might, for example, be interested in identifying parts of the deliberation that are mature and ready to be "harvested" or, conversely, that are dysfunctional (e.g. exhibiting balkanization or groupthink) and need some kind of intervention. An author might be interested in being notified of controversies that have arisen in an area they previously contributed to, of pet ideas whose support has dropped and might be revived by the addition of additional supportive arguments, or of posts where there ratings appear to exhibit an irrational bias. In our current implementation, users are presented these suggestions in the form of an argument map subset wherein the suggested posts are highlighted and the reasons for the highlighting appear when they roll over the post:

Fig. 4. The personalized suggestions display.

The emergent effect of these automatically-generated suggestions, we believe, will be to help ensure that each part of the deliberation receives attention, and is fully developed by, the participants with the most interest and knowledge on the topic.

5 Conclusions

The key contribution of this work is to explore how automated algorithms can generate real-time metrics that help users allocate their deliberation efforts, in an argument map context, to where they can do the most good. This approach, if executed well, synergistically harnesses the creativity and judgment of human

communities along with the ability of computational systems to rapidly summarize and visualize large data sets.

While there has been substantial effort devoted to manually-coded, *post-hoc* metrics on the efficacy of on-line deliberations [32] [33] [34] [35] [36] [37], existing deliberation technologies have made only rudimentary use of automated *real-time* metrics to foster better emergent outcomes during the deliberations themselves. The core reason for this lack is that, in existing deliberation tools, the content takes the form of unstructured natural language text, limiting the possible deliberation metrics to the analysis of word frequency statistics, which is a poor proxy for the kind of semantic understanding that would be necessary to adequately assess deliberation quality. One of the important advantages of using argument maps to mediate deliberation is that they allow us, by virtue of their additional semantics, to automatically derive metrics that would require resource-intensive manual coding for more conventional social media such as web forums. We are aware of one other effort to develop real-time deliberation metrics for large-scale argument mapping, but this work [38] is based on measuring how well the deliberations adhere (e.g. in terms of audibility, simultaneity of messages, and mobility of the participants) to a normative model of small-scale, physically collocated *conversations* [39]. Our work is unique, we believe, in how it attempts to assess (and improve) how effectively large groups are *deliberating* (i.e. exploring and converging on problem solutions) rather than just how well individuals are *conversing*.

Our work to date has been largely conceptual, focusing on identifying what kinds of metrics could foster better emergent properties in large-scale argumentation-based deliberations. Our future work will focus on the empirical, analytic, and computational (simulation-based) assessment of the emergent impact of these metrics.

Acknowledgements. The author would like to gratefully acknowledge the many useful conversations he has had on the topic of deliberation metrics with Prof Ali Gurkan (Ecole Centrale Paris), Prof. Luca Iandoli (University of Naples), and Prof. Haji Reijers (Eindhoven University of Technology).

References

1. Sunstein, C.R., Infotopia: How Many Minds Produce Knowledge. 2006: Oxford University Press.
2. Walton, D.N. and E.C.W. Krabbe, Commitment in dialogue: Basic concepts of interpersonal reasoning. 1995, Albany, NY: State University of New York Press.
3. Eemeren, F.H.v. and R. Grootendorst, A Systematic Theory of Argumentation: The Pragma-dialectical Approach. 2003: Cambridge University Press.
4. Pervan, G.P. and D.J. Atkinson, GDSS research: An overview and historical analysis. Group Decision and Negotiation, 1995. 4(6): p. 475-483.
5. Luppicini, R., Review of computer mediated communication research for education. Instructional Science, 2007. 35(2): p. 141-185.
6. Gopal, A. and P. Prasad, Understanding GDSS in Symbolic Context: Shifting the Focus from Technology to Interaction. MIS Quarterly, 2000. 24(3): p. 509-546.
7. Macaulay, L.A. and A. Alabdulkarim, Facilitation of e-Meetings: State-of-the-Art Review, in e-Technology, e-Commerce and e-Service (EEE'05). 2005.

8. Powell, A., G. Piccoli, and B. Ives, Virtual teams: a review of current literature and directions for future research. ACM SIGMIS Database, 2004. 35(1): p. 6 - 36.
9. Finholt, T.A., Collaboratories. Annual Review of Information Science and Technology, 2002. 36(1): p. 73-107.
10. Tapscott, D. and A.D. Williams, Wikinomics: How Mass Collaboration Changes Everything. 2006: Portfolio Hardcover.
11. Surowiecki, J., The Wisdom of Crowds. 2005: Anchor.
12. Gladwell, M., The Tipping Point: How Little Things Can Make a Big Difference. 2002: Back Bay Books.
13. Kittur, A., et al., He says, she says: conflict and coordination in Wikipedia, in SIGCHI Conference on Human Factors in Computing Systems. 2007, ACM: San Jose, California, USA.
14. Viegas, F.B., M. Wattenberg, and K. Dave, Studying cooperation and conflict between authors with history flow visualizations, in SIGCHI conference on Human factors in computing systems. 2004: Vienna, Austria.
15. Kirschner, P.A., S.J.B. Shum, and C.S.C. Eds, Visualizing Argumentation: Software tools for collaborative and educational sense-making. Information Visualization, 2005. 4: p. 59-60.
16. Moor, A.d. and M. Aakhus, Argumentation Support: From Technologies to Tools. Communications of the ACM, 2006. 49(3): p. 93.
17. Walton, D.N., Fundamentals of Critical Argumentation (Critical Reasoning and Argumentation). 2005, Cambridge (MA): Cambridge University Press.
18. Carr, C.S., Using computer supported argument visualization to teach legal argumentation, in Visualizing argumentation: software tools for collaborative and educational sense-making, P.A. Kirschner, S.J.B. Shum, and C.S. Carr, Editors. 2003, Springer-Verlag. p. 75-96.
19. Shum, S.J.B., et al., Hypermedia Support for Argumentation-Based Rationale: 15 Years on from gIBIS and QOC, in Rationale Management in Software Engineering, A.H. Dutoit, et al., Editors. 2006, Springer-Verlag.
20. Jonassen, D. and H.R. Jr, Mapping alternative discourse structures onto computer conferences. International Journal of Knowledge and Learning, 2005. 1(1/2): p. 113-129.
21. Chklovski, T., V. Ratnakar, and Y. Gil, User interfaces with semi-formal representations: a study of designing argumentation structures. Proceedings of the 10th international conference on Intelligent user interfaces, 2005: p. 130-136.
22. Lowrance, J.D., I.W. Harrison, and A.C. Rodriguez, Capturing Analytic Thought, in First International Conference on Knowledge Capture. 2001. p. 84-91.
23. Karacapilidis, N., E. Loukis, and S. Dimopoulos, A Web-Based System for Supporting Structured Collaboration in the Public Sector. LECTURE NOTES IN COMPUTER SCIENCE, 2004: p. 218-225.
24. Heng, M.S.H. and A. de Moor, From Habermas' s communicative theory to practice on the internet. Information Systems Journal, 2003. 13(4): p. 331-352.
25. Rahwan, I., Mass argumentation and the semantic web. Journal of Web Semantics, 2008. 6(1): p. 29-37.
26. Conklin, J., Dialogue Mapping: Building Shared Understanding of Wicked Problems. 2005: John Wiley and Sons, Ltd.
27. Erickson, T., et al., Social Translucence: Designing Social Infrastructures that Make Collective Activity Visible. Communications of the ACM, 2002. 45(4): p. 40-44.
28. Klein, M., The MIT Collaboratorium: Enabling Effective Large-Scale Deliberation for Complex Problems. 2007, MIT Sloan School of Management.
29. Klein, M. and L. Iandoli, Supporting Collaborative Deliberation Using a Large-Scale Argumentation System: The MIT Collaboratorium, in Directions and Implications of

Advanced Computing; Conference on Online Deliberation (DIAC-2008/OD2008). 2008: University of California, Berkeley.
30. Iandoli, L., M. Klein, and G. Zollo, Enabling on-line deliberation and collective decision-making through large-scale argumentation: a new approach to the design of an Internet-based mass collaboration platform. International Journal of Decision Support System Technology, 2009. 1(1): p. 69-91.
31. Klein, M., A Knowledge-Based Methodology for Designing Reliable Multi-Agent Systems, in Agent-Oriented Software Engineering IV, P. Giorgini, J.P. Mueller, and J. Odell, Editors. 2003, Springer-Verlag. p. 85 - 95.
32. Steenbergen, M.R., et al., Measuring political deliberation: a discourse quality index. Comparative European Studies, 2003. 1(1): p. 21-48.
33. Stromer-Galley, J., Measuring deliberation's content: a coding scheme. Journal of Public Deliberation, 2007. 3(1).
34. Trénel, M., Measuring the quality of online deliberation: Coding scheme 2.4. 2004, Social Science Research Center: Berlin, Germany.
35. Cappella, J.N., V. Price, and L. Nir, Argument Repertoire as a Reliable and Valid Measure of Opinion Quality: Electronic Dialogue During Campaign 2000. Political Communication, 2002. 19(1): p. 73 - 93.
36. Spatariu, A., K. Hartley, and L.D. Bendixen, Defining and Measuring Quality in Online Discussions. The Journal of Interactive Online Learning, 2004. 2(4).
37. Nisbet, D., Measuring the Quantity and Quality of Online Discussion Group Interaction Journal of eLiteracy, 2004. 1: p. 122-139.
38. Shum, S.B., et al., A Debate Dashboard to Support the Adoption of Online Knowledge Mapping Tools. VINE Journal of information and Knowledge Management Systems in press.
39. Clark, H.H. and S.E. Brennan, Grounding in communication, in Perspectives on socially shared cognition, L.B. Resnick, J.M. Levine, and S.D. Teasley, Editors. 1991, American Psychological Association,: Washington, DC, US. p. 127-149.