DCLA Meet CIDA: Collective Intelligence Deliberation Analytics

Simon Buckingham Shum[1]**, Anna De Liddo**[1] **and Mark Klein**[2]

[1] Knowledge Media Institute
The Open University
Walton Hall, Milton Keynes
MK7 6AA, UK

[2] Center for Collective Intelligence
Massachusetts Institute of Technology, Cambridge MA, USA
& Dynamic and Distributed Information Systems Group,
Universität Zürich, CH-8050 Zürich, Switzerland

s.buckingham.shum@gmail.com
anna.deliddo@open.ac.uk

m_klein@mit.edu

Abstract: This discussion paper builds a bridge between *Discourse-Centric Learning Analytics (DCLA)*, whose focus tends to be on student discourse in formal educational contexts, and research and practice in *Collective Intelligence Deliberation Analytics (CIDA)*, which seeks to scaffold quality deliberation in teams/collectives devising solutions to complex problems. CIDA research aims to equip networked communities with deliberation platforms capable of hosting large scale, reflective conversations, and actively feeding back to participants and moderators the 'vital signs' of the community and the state of its deliberations. CIDA tends to focus not on formal educational communities, although many would consider themselves learning communities in the broader sense, as they recognize the need to pool collective intelligence in order to understand, and co-evolve solutions to, complex dilemmas. We propose that the context and rationale behind CIDA efforts, and emerging CIDA implementations, contribute a research and technology stream to the DCLA community. The argument is twofold: (i) The context of CIDA work connects with the growing recognition in educational thinking that students from school age upwards should be given the opportunities to engage in authentic learning challenges, wrestling with problems and engaging in practices increasingly close to the complexity they will confront when they graduate. (ii) In the contexts of both DCLA and CIDA, different kinds of users need feedback on the state of the debate, and the quality of the conversation: the students and educators served by DCLA are mirrored by the citizens and facilitators served by CIDA. In principle, therefore, a fruitful dialogue could unfold between DCLA/CIDA researchers and practitioners, in order to better understand common and distinctive requirements.

1. Introduction

This discussion paper introduces work in an emerging field, which we will refer to as Collective Intelligence Deliberation Analytics (CIDA), as a contribution to the conversation at the Discourse-Centric Learning Analytics (DCLA) workshop on relevant literature and technology in the DCLA research and design space. Our argument proceeds as follows. All over the world, collectives are forming to tackle the challenge of devising practical, collectively owned solutions to some of society's most pressing problems. Such conversations are taking place daily at many scales, from small teams, organisations and networks of a few dozen, to several hundred, to even thousands of participants in open consultations, and networks for advocacy or social innovation. This is becoming a recognizable research field. For instance, the European Commission Horizon 2020 R&D programme now has a theme dedicated to the role of technology in building *Collective Awareness Platforms for Sustainability and Social Innovation*, which is currently supporting the work reported here (EC-Horizon2020), and a series of annual international workshops has been held in the field of

organizational collaborative computing, devoted to ideation and deliberation tools for Collective Intelligence (CI) (CICDA, 2012).

A contribution that the web makes to such collective intelligence (CI) are new kinds of processes and tools to facilitate *deliberation*—the generation of ideas, and evidence-based critical evaluation, of potential solutions to issues. In this context, the distinctive strand of our work focuses on addressing the acute limitations of current social platforms that merely invite the submission of comments and 'ideas', on which people then vote with simple 'thumbs up' or 'like' clicks. CI deliberation platforms require participants to build more reflectively on each others' contributions. Efforts are now under way to develop analytics services that provide ways to make sense of how much progress is being made. This then, is the context of CIDA research.

Meanwhile, educational institutions are being challenged on many fronts. A fast moving technological landscape is opening up new possibilities, as documented by numerous horizon-scanning learning technology reports. In higher education the quality of the student experience is under intense scrutiny, as educational systems reflect at national and international levels on their fitness for purpose and value for money. Coupled with calls from business for graduates who are not just academically excellent, but have transferable skills and competencies equipping them for the complexities of the workplace, this is serving as a driver for action research into new models focused on the wholistic design of learning, catalysing academics (Deakin Crick, 2009; Gardner, 1983; Perkins, 1993; Claxton, 2001), national networks and funding programmes. A particular focus of these initiatives is on 'deeper learning' that is more authentic in nature (Whole Education, 2014; Hewlett Foundation, 2014; Fullan and Langworthy, 2013)—to the extent that the educator may not know the 'right answer' but is learning with the students. Indeed, there may be no knowable right answer, such is the open-ended nature of truly "wicked problems" whose very definition sometimes defies consensus (Rittel, 1972). While accuracy of conceptual understanding remains as important as ever, in the absence of a knowable correct solution, it becomes increasingly important to evidence mastery of the appropriate *processes* through which one may tackle such open-ended problems. To the extent that knowledge is constructed through the use of language to share and context ideas, discourse becomes a window into the learner's mind, hence the importance of discourse analytics.

This paper therefore explores the DCLA-CIDA interplay: between deliberation platforms and associated analytics in formal educational contexts (DCLA), and informal, applied contexts (CIDA). We explore the proposition that the educational need to scaffold authentic learning on problems that matter, and moreover the need to monitor the quality of that process, converges very naturally with developments in CI deliberation platforms. Both contexts require that analytics make more visible what is going on in the platform for the benefit of participants (whether *students* or *citizens*) and those charged with facilitating them (whether *teachers* or *e-participation moderators*).

> At various points we highlight reflections on the differences between DCLA and CIDA use contexts, or what the DCLA or CIDA community may learn from each other.

2. The need for reflective deliberation tools

2.1 The limits of forums and commenting

The online discussion spaces we see on the Web today typically provide for the addition of flat listings of comments, listed by date (e.g. comments in Facebook; on web articles; on blog posts), or threaded in a strict tree which can be additionally viewed by 'subject' line (e.g. Google or Yahoo Groups;

Forum-Software.org). These are fundamentally *chronological* views of the unfolding conversation drawing attention to the most recent utterances, but offer no insight into the *logical* structure of the ideas — the coherence of the argument. At a glance, all one can tell is which threads have most contributions, or have been most active.

> **Learning context:** Online learning platforms are similarly dominated by threaded discussion forums, and in recent years, social web style platforms with even simpler flat commenting (e.g. via status updates, blogs or wikis). These do not have affordances that promote reflective contributions, or assist analysis of the state of the debate.

2.2 The limits of ideation tools

Another class of CI platform is the Idea Management System (IMS) for creative, collective ideation. These are designed to support grassroots innovation systems for the employees of an organization to deliberate about innovation or for its customers to suggest new products (e.g., Bailey and Horwitz; 2010).[1]

Klein (Klein, 2012) has critiqued such tools, noting that when an entire community is participating in the idea management and deliberation process, then a much larger number of ideas are generated, and the selection and judgment can become prohibitively lengthy and time consuming. Bjelland and Wood (2008) studied the process of idea generation at IBM and stressed the critical role played by managers and the large amount of work that they did. The role of facilitators was essential together with the software for identifying and nurturing good ideas as these were generated by the organization. Klein reports also that the work of facilitators can be prohibitively lengthy and time consuming. In the context of one Google project (www.project10tothe100.com), thousands of people from around the world submitted about 154,000 submissions. Google had to recruit 3,000 employees to filter and consolidate the large number of ideas received in a process that put the company nine months behind their original schedule. The work of Convertino (Baez & Convertino, 2012) exemplifies the next generation of IMS which seeks to tackle these bottlenecks.

> **Learning context:** The number of 'likes' that a contribution to a social/ideation platform receives can be motivating for students, giving them a sense of identity in a group. However, a deeper learning orientation must move beyond the popularity of ideas, and challenge learners about the rationale behind a vote: what is the quality of reasoning that leads a student to vote an idea up or down? Can they make that thinking visible in their discourse, and engage with contrasting views, in an appropriate manner?

2.3 Summary

The very low 'entry threshold' that all of these tools set for contributing (in order to maximise the number of users signing up, for either commercial purposes, or to maximize participation on democratic grounds) is in tension with the need to promote, and understand, a high *quality* conversation, which is critical to more advanced forms of collective endeavour. Our interest is in the more complex forms of collective action, socio-political dilemmas where people may disagree on the nature of the problem, what might count as a solution, where there is insufficient data, ambiguity about its trustworthiness, and uncertainty about the impact of actions. The answers to complex
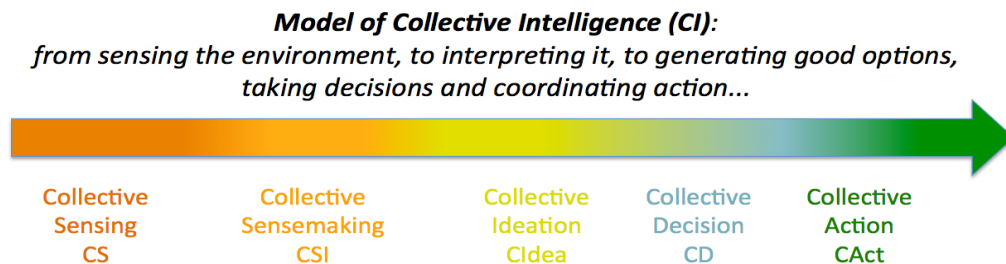
---

[1] e.g. http://www.spigit.com, http://ideascale.com

questions of this sort are rarely simple, but we propose that their complexity can be managed using techniques to show more clearly the nature of agreement and disagreement, poorly developed lines of reasoning, and where a contribution might make most impact rather than duplicate prior work. Resolving these requires high quality deliberation, which could benefit from more powerful scaffolding.

3. Pain points in online deliberation for social innovation

Collective intelligence can be viewed as comprising a spectrum of capabilities that ranges from collective *sensing* at one end (where a collective pools data on its environment), through *sensemaking* (interpreting data to identify patterns that warrant action), *ideation* (developing ideas about which actions to pursue), *decision-making* (selecting the best actions), and *action* (implementing these actions in a coordinated effective way) — Figure 1.

Model of Collective Intelligence (CI):
from sensing the environment, to interpreting it, to generating good options, taking decisions and coordinating action…

| Collective Sensing CS | Collective Sensemaking CSI | Collective Ideation CIdea | Collective Decision CD | Collective Action CAct |

Figure 1: A spectrum of CI activities

At the left side of the spectrum, sensor infrastructures are becoming ubiquitous, while towards the right end, relatively mature online voting and workflow coordination technologies support option prioritization and plan execution. We are currently much less able to collectively make sense of the vast amounts of data now available to us, and to come up with innovative and effective arguments for solutions, especially in domains where there are many competing perspectives on how to understand and solve a problem (Tversky and Kahneman, 1974; Sunstein, 2006; Schulz-Hardt et al., 2000; Cook and Smallman, 2007; Klein, 2012).

As part of the European Union *Catalyst* project, we have recently completed a requirements consultation with five organisations who specialize in hosting and moderating medium-large scale online conversations for social innovation collectives. Five workshops with on average 10 participants followed a methodology in which they were asked to prioritise the most pressing "pain points" that they experience in facilitating such community deliberations. We summarise next the highest priority issues, and the reader is invited to consider how closely these parallel the challenges of scaffolding quality student discourse in more formal learning contexts.

3.1 Hard to visualise the debate

The consultees confirmed that the effective visualisation of concepts, new ideas and deliberations is essential for shared understanding, but suffers both from a lack of efficient tools to create them and from a lack of ways to reuse them across platforms and debates. Yet, most users of these platforms (multilingual, multicultural communities, Generation Y, etc.) wish to have access to easy-to-

understand image/video-based content that they can grasp very rapidly and share easily with their peers via social media channels.

Many consultees reported that poor summarisation and visualisation are "the biggest problems as these both result in platforms which are unappealing to the user, and therefore suffer from a lack of participation". In particular, visualisation of the debate was considered important for both the community who participate to the debate and community manager. Some claimed that: "As a user, visualisation is my biggest problem. It is often difficult to get into the discussion at the beginning. As a manager of these platforms, showing people what is going on is the biggest pain point."

A good argument has particular attributes that are well understood by scholars of informal logic and other kinds of argumentation (Walton et al., 2008). Normally, of course, arguments are expressed in prose, and it is left to the reader to tease apart the elements in order to form a judgement about the line of reasoning. However, it is well established that the general public have poor argumentation and critical reasoning skills (Kuhn, 1991; Rider and Thomason, 2008). One of the technical goals of CIDA platforms is, consequently, to find more effective ways of rendering arguments, in order to make them more tractable for both humans and computational intelligence to parse and evaluate.

> **Learning context:** Somewhat disappointingly, university students are no exception to the finding that most citizens have poor reasoning skills. Undergraduate students find it challenging to write in the appropriately critical manner that is the hallmark of academic writing (Andrews, 2010; Lillis and Turner, 2001)

3.2 Poor summarisation

Summarisation is a prerequisite to informed participation in online debates. Participants struggle to get a good overview of what is unfolding in an online community debate. Only the most motivated participants will commit a lot of time to reading the debate in order to identify the key members, the most relevant discussions, etc. The majority of participants tend to respond unsystematically to stimulus messages, and do not digest earlier contributions before they make their own contribution to the debate, such is the cognitive overhead and limited time. This problem is crucial since it also influences other pain points such as idea duplication, shallow idea contribution and therefore poor participation.

> **Learning context:** The demands of reading prior contributions before posting one's own apply as equally to student forums as others. The difference may, however, be that student forums can be more tightly controlled either through imposing practices (e.g. "your postings must engage with ideas in at least two previous postings"; your postings should seek to forge meaningful connections between earlier ones"), or through imposing technical constraints on users permissions ("you will be randomly assigned to a team of 6 who can only access their team forum").

3.3 Poor commitment to action

Decision-support systems that underestimate the complexity of socio-technical problems suffer from problem definitions not being collectively owned (Conklin, 2006), hence our focus on high-quality deliberation process. However, even once candidate courses of action are clear, bringing motivated audiences to commit to action is difficult. Enthusiasts, those who have an interest in a subject but have yet to commit to taking action, are left behind.

This was a central issue for most of the participants. Participants cared about ways to prompt action in community members, to the extent that reaching a consensus was considered less important than being enabled to act.

> **Learning context:** The seriousness of this problem in educational settings may depend on how much autonomy students are given. In increasingly authentic learning contexts, one would expect greater student choice about the direction and focus of a project, and hence the growing importance of commitment to act to achieve that goal.

3.4 Sustaining participation

About 80% of the workshop participants considered lack of participation either important or very important. Of these, 40% considered it crucial. Even with optimal methodologies, it is typical that only a fraction of any group will actively participate in online deliberation, as reflected in the well known 'long tail' distribution of social media activity, whereby a very small percentage of participants are highly active. Moderators reported that motivating participants with widely differing levels of commitment, expertise and availability to contribute to an online debate is challenging and often unproductive. Participation is usually one-off, and maintaining presence as well as interest in the deliberation is a challenge.

The focus of interest in the workshops seemed to be more on how to maintain participation rather than on how to enlarge participation. As a general and recurring comment, also across workshop groups, many claimed: "it is better to have quality input from a small group than a lot of members but very little content". The real issue was considered to be "the lack of worthwhile and productive input. For example, 'liking' something on Facebook is a way of participating, but it is not necessarily that productive".

Several participants suggested that, in terms of priorities, rather then seeking simply to engage more and more participants in the conversation, it is more important that the actions discussed are then followed through and implemented, and that the platform should allow users to track the outcomes of those actions.

> **Learning context:** In an earlier paper (Buckingham Shum and Deakin Crick, 2012) we noted the disturbing levels of disengagement in the schools of many developed countries (Gilby et al., 2008; Willms et al., 2009; Yazzie-Mintz, 2009). These point to a widening disconnect between what motivates and engages many young people, and their experience of schooling. Such students, including the high ability but disengaged, want to know the point of bothering to re-engage, and are thus analogous to the hard-to-engage citizens that the consultees spoke about. This *sustaining participation* pain point thus points to the importance of (i) making learning authentic, and (ii) fostering forms of engagement that go beyond the lowest common denominator *like* act, or superficial comment.

3.5 Shallow contributions and unsystematic coverage

Open innovation systems tend to generate a large number of relatively shallow ideas. Open innovation sites do not in general encourage or support the collaborative refinement of ideas that could allow the development of more refined, deeply considered contributions. The majority of the contributions thus tend to be repetitions of a few obvious ideas. Moreover, there is no inherent mechanism for ensuring that the ideas submitted comprehensively cover the different facets of the problem at hand. The space of possible solutions is generally not specified up front, and there is no easy way for potential

contributors to see which problem facets remain under-covered. The repetition mentioned above is detrimental to consideration of how similar ideas actually differ (improving understanding) or focusing on new ideas. As a result, social innovation systems often produce very partial coverage of the solution space.

About 70% of consulted people considered this problem important or very important to enable effective online deliberation while the remaining 30% considered this of moderate or minor importance.

Learning context: This pain point is highly relevant to learning contexts requiring the rational analysis of a well-understood problem space. The task here is to understand the key dimensions that may be used to differentiate one important class of solution from another, and/or the hierarchies that may be used to clarify how different solutions compare.

It is important to note, however, that in a learning context, *shallow contributions* — in terms of social exchanges not specifically about the curriculum material — can play helpful roles by helping to foster a socially welcoming online environment, or building a community of practice: they build social capital, and create the conditions of trust in which learners are more likely to be open to challenge and new ideas. Such qualities have a role to play in authentic, socially rich learning contexts, but are less valued in assessments for which this is considered irrelevant. Shallow comments would therefore normally be considered weak contributions (unless the course is, for instance, about learning to host social learning spaces, resolve conflict, facilitate professional learning, etc.). However, if building trust and social capital is not the end of the story, they are preconditions to subsequent contributions becoming "deeper", for instance, evidencing a grasp of more complex concepts, modes of critical thinking and argumentation, or more challenging forms of discourse that might undermine one's preconceptions or threaten a worldview.

3.6 Poor idea evaluation

When there are thousands of ideas, as is common for innovation about complex problems, many emergent effects can deeply undercut the value of community ratings. Most people are likely to evaluate only a tiny fraction of the ideas, usually the ones at the top of the list. If, as is often the case, ideas are sorted by their average rating, one can expect that the system will quickly "lock" into a fairly static, and arbitrary, ranking, where the first few winners take all, even if they are inferior to many other ideas in the list. Even worse, stakeholders with vested interests can game the rating mechanisms in open innovation systems in order to manipulate which ideas rise to the top. A single idea post with focused voting, for example, may beat out a better idea that had its votes spread over many redundant instantiations. In an open ideation engagement, there is also often a disconnection between the voting of the contributors, and the idea evaluation criteria (often implicit) held by the person who voted.

This problem is exacerbated when, as is often the case, people evolve their understanding of what they want as they learn more about the space of possible solutions. In general, current open innovation systems provide little support for a varied crowd building upon each other's evaluative expertise, or mentoring one another. They cannot see why other contributors provided the ratings that they did, nor is there a good way for them to examine and correct each other's facts and reasoning. At best, current open innovation systems just provide comment streams to capture discussions about the worth of an idea, and these quickly become unwieldy as the number of comments on an idea increases.

Among the workshop participants, while this was not ranked as high as preceding pain points, approximately 60% ranked this "important" or "moderately important".

4. Collective intelligence deliberation platforms

4.1 Essence of the approach: semantic hypertext

The inaugural conference devoted to Collective Intelligence (CI), defined it as:

> "…behaviour that is both *collective* and *intelligent*. By collective, we mean groups of individual actors, including, for example, people, computational agents, and organizations. By intelligent, we mean that the collective behaviour of the group exhibits characteristics such as, for example, perception, learning, judgment, or problem solving." [www.ci2012.org]

We have previously drawn a distinction between CI derived from the aggregation of low level click traces left as a by-product of user activity, and CI emerging from intentional, user acts of synthesis and interpretation (De Liddo and Buckingham Shum, 2010). The latter is the focus of CIDA efforts, since wicked problems require significant intellectual effort in the construction of plausible narratives that make sense of data, which depending on the context may be scarce or abundant—but it is its trustworthiness and significance which must be decided.

There is an established research literature that provides key insights into the problem of making the structure and status of such deliberation visible. This focuses on the design of semiformal representations designed to aid both human and machine interpretation. Elsewhere, we have documented the roots to this field (Buckingham Shum, 2003). Explicit semantic networks provide a computational system with a more meaningful understanding of the relationships between ideas than natural language.

Following the established methodological value of Concept Mapping (Novak, 1998), the mapping of issues, ideas and arguments extends this to make explicit the presence of more than one perspective and the lines of reasoning associated with each. More formal approaches, derived from the convergence of AI and argumentation theory (Walton et al., 2008) model argument structures in finer detail, thus enabling automated evaluation. However, a longstanding research challenge is to add such computational power without sacrificing usability for non-experts.

A relatively simple 'lightweight' scheme for structuring deliberation has emerged as a popular choice for deliberation platforms. This uses a scheme originally devised in the 1970s called the *Issue-Based Information System* (IBIS) (Rittel, 1984 (1972); Rittel and Webber, 1973). The key elements of this form of mapping conversational moves are shown in Figure 1.
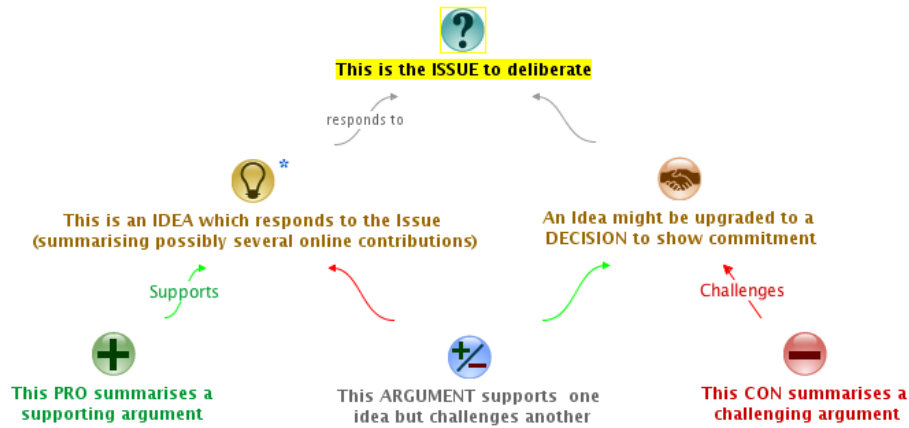
Figure 1: Core elements of a map using the IBIS scheme, underpinning many deliberation platforms, which render the structures using diverse visualizations

An extensive literature documents the adoption issues that some users face when they attempt this level of discourse structuring in synchronous face-to-face contexts (Buckingham Shum et al., 2006). Counter-balancing this are accounts of high impact mapping by 'cartographers' who are fluent with this, around which we are now developing a theoretical account and evidence base (Selvin et al., 2010). There is a growing body of evidence showing under what conditions tools which structure dialogue and argumentation into visualizations can support groups to build shared understanding, explore solutions to complex problems, record their rationale, and make better informed collective decisions (Okada et al., 2008). These include *e-participation/e-democracy* (Renton and Macintosh, 2007; Iandoli et al., 2009) *environmental policy* (Conklin, 2003), *participatory urban planning* (Culmsee and Awati, 2011), *distributed science in the field* (Sierhuis and Buckingham Shum, 2008), and *emergency response* (Tate et al., 2007). This line of work has led us to propose that we need to develop a distinctive class of CI system for *Contested CI* in which consensus cannot be assumed, but rather, the norm is to have perspectives in tension, the central task being to make those visible in order to understand them, with human and machine intelligence harnessed to this end (De Liddo et al., 2012).

The emergence of IBIS as a 'lingua franca' in CI deliberation platforms reflects its relative simplicity, and hence comprehensibility and usability, compared to richer, more expressive argument modeling approaches, which require correspondingly higher skill levels.

4.2 Examples of collective intelligence deliberation platforms

To make these technologies more tangible, let us give a few examples, noting that there is a growing range of tools[2], some of which are in use by citizens with no training. *DebateGraph* has been used to host public debates around political policy and other societal dilemmas, in association with government consultations, and many other public and not-for-profit sector organisations Figure 2. DebateGraph is being used in over 100 countries in many different fields, including education, health, governance, media, publishing, environment, conflict resolution, conferences, group facilitation, and public consultation and planning.

---

[2] e.g. *Online Deliberation: Emerging Tools*: ODET 2010 workshop: www.olnet.org/odet2010
   Arguing on the Web 2.0: ISSA 2014 workshop: http://www.sintelnet.eu/content/arguing-web-20-0
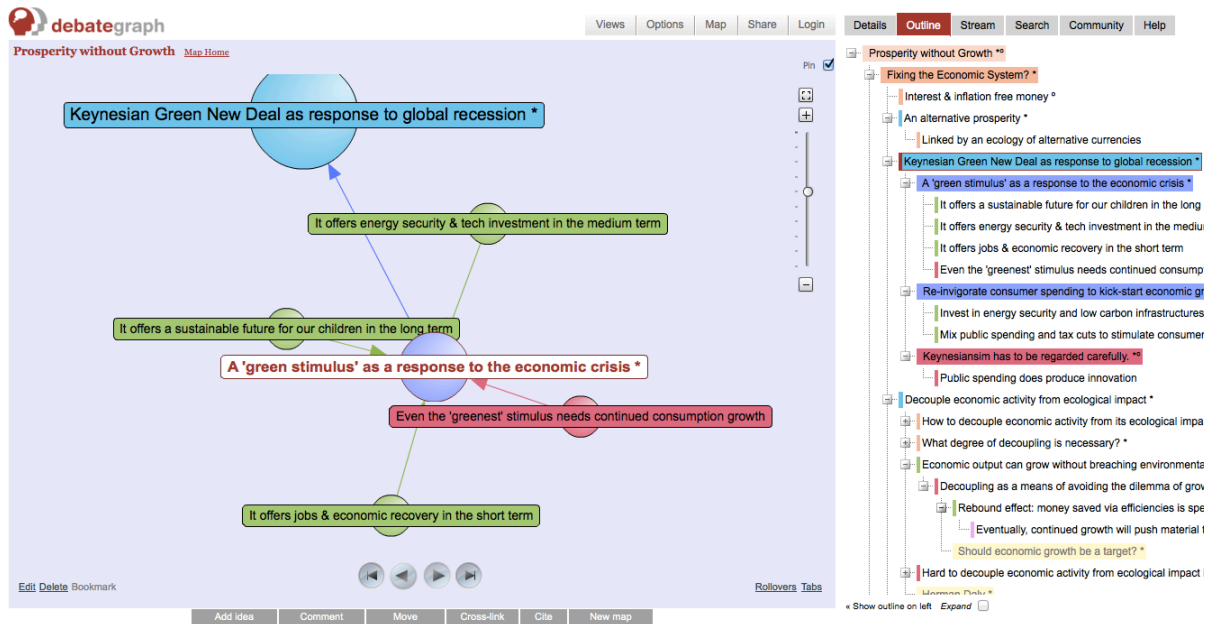
Figure 2: DebateGraph supporting a White House online consultation

These deployments confirm that untrained users are willing to engage with more structured deliberation platforms when they care enough about the issues at stake, but there has not yet been academic research into usage patterns.

Another IBIS network mapping application is Cohere (Figure 3), which adds web annotation functionality (to directly insert clips from websites into the map) and social network analysis. The goal is to produce aggregated, integrated views of both the social and discourse dynamics. For instance, while social network analysis networks typically show social ties based on direct interaction of some sort, undifferentiated except by the strength of the tie, this example shows a 'semantic social network' visualization of the kinds of social ties connecting participants, based on their agreement or disagreement around ideas (see colour-coded green/red/grey links) (De Liddo et al., 2011).
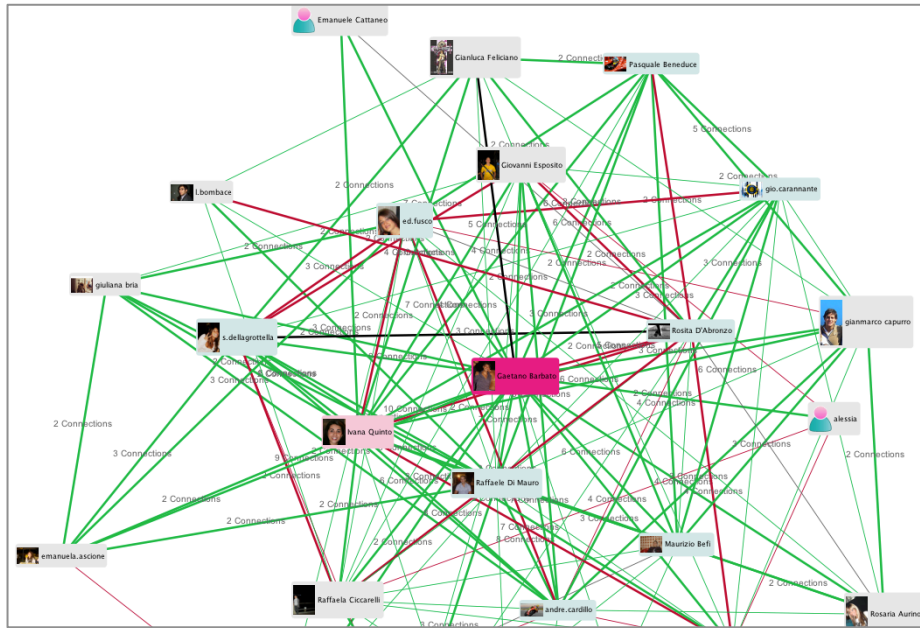
Figure 3: Cohere's Semantic Social Network view colour-codes the ties between people based on an algorithm weighting how they are connected via semantic argumentation moves

The OU's *Evidence Hub* (De Liddo and Buckingham Shum, 2013a) has been designed as CI site for pooling fragments of clippings from the web, and the posting of issues and ideas, but then scaffolds the process of moving from these into 'evidence-based stories' which make meaningful, argumentative connections between these elements. Since a harmonious 'big picture' rarely emerges in contested fields such as education, business or policy making, the Evidence Hub aims also to show where people disagree and why, and where there appears to be a lack of evidence to back any of the claims.

Figure 4 shows the story wizard which scaffolds the user through the process of submitting an evidence-based claim.

Figure 4: The Evidence Hub's Story Wizard scaffolds the submission of an IBIS map without users needing to understand the underlying semantics

An example of the IBIS-based structure generated by the story wizard is shown in Figure 5, rendered as discussion forum-style threads, with added semantic types from IBIS:

Figure 5: An Evidence Hub outline knowledge tree view of the IBIS structure. Graph visualization and a zoomed-in details view are also provided.

A number of empirical studies now document the costs and benefits of different versions' user interfaces (Iandoli et al., 2013; De Liddo and Buckingham Shum, 2013b).

MIT's *Deliberatorium* also uses the IBIS scheme, and has been evaluated in both educational contexts and as a medium for public political discourse (Iandoli et al., 2009; Klein et al., 2012). Figure 6 is taken from a city-wide deployment for citizen debate in an Italian election.

Figure 6: Deployment of Deliberatorium in citizen debate during a Naples city election

While the structure that IBIS introduces requires greater effort from participants, evaluation of this deployment in comparison to a conventional web forum:

> "suggests that argument maps, despite the additional demands they place on users in terms of structuring their contributions, do *not* suppress participation compared to the much simpler, and more familiar, medium of web forums. This is, we believe, an important result in terms of the viability of argument-centric large-scale deliberation, and is also potentially relevant to the viability of other social computing approaches (e.g. collaboratively-authored semantic networks) that use semi- formal representations."

After years of experimenting with the design and usage of graphical argument maps to promote critical thinking in public and educational discourse (van Gelder, 2003), Tim van Gelder is now concentrating on more conventional, simpler interfaces for civic debate. YourView[3] offers a two-column interface to see, for each proposed idea, the list of supporting and challenging arguments (Figure 7).

---

[3] YourView: https://yourview.org.au Blogs: http://timvangelder.com/category/yourview/

Figure 7: YourView's two-column view of for/against arguments

Although IBIS is relatively simple as a modeling scheme, additional visualizations can be generated from it. It is possible to start mapping peoples' positions compared to a norm, such as how much you agree or disagree with peers (Figure 8), or with the addition of domain-specific data, how citizens' votes are close to a specific political party view (Figure 9).

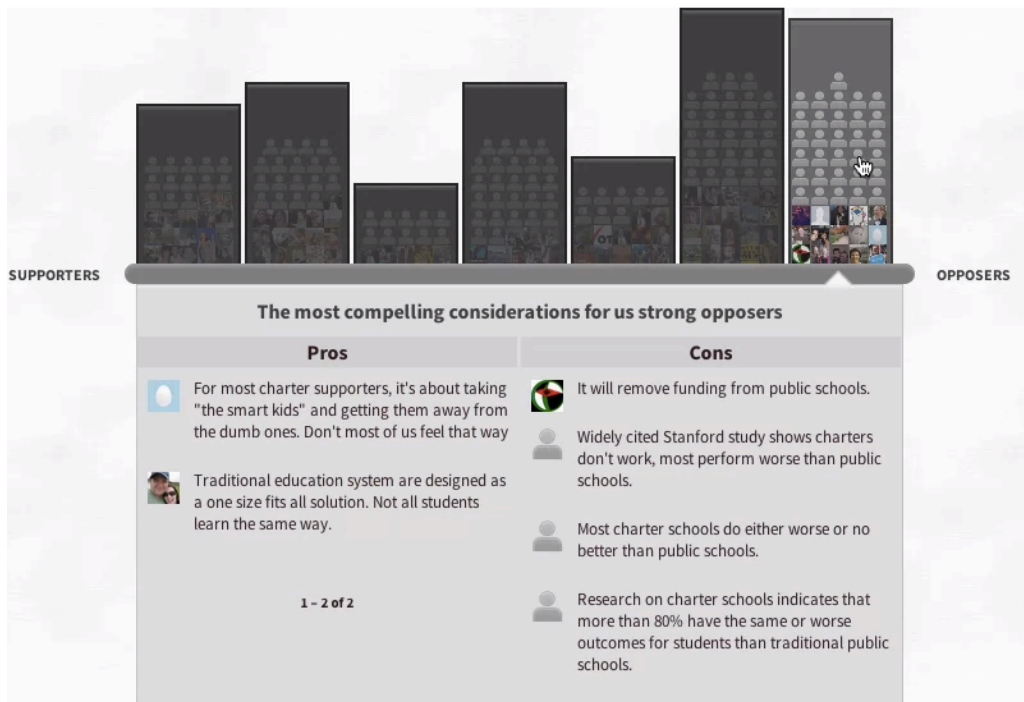**Figure 8: Consider.it group's thinking histogram groups people in terms of how much they support or challenge an idea, and shows the pros and cons each group appeals to.**
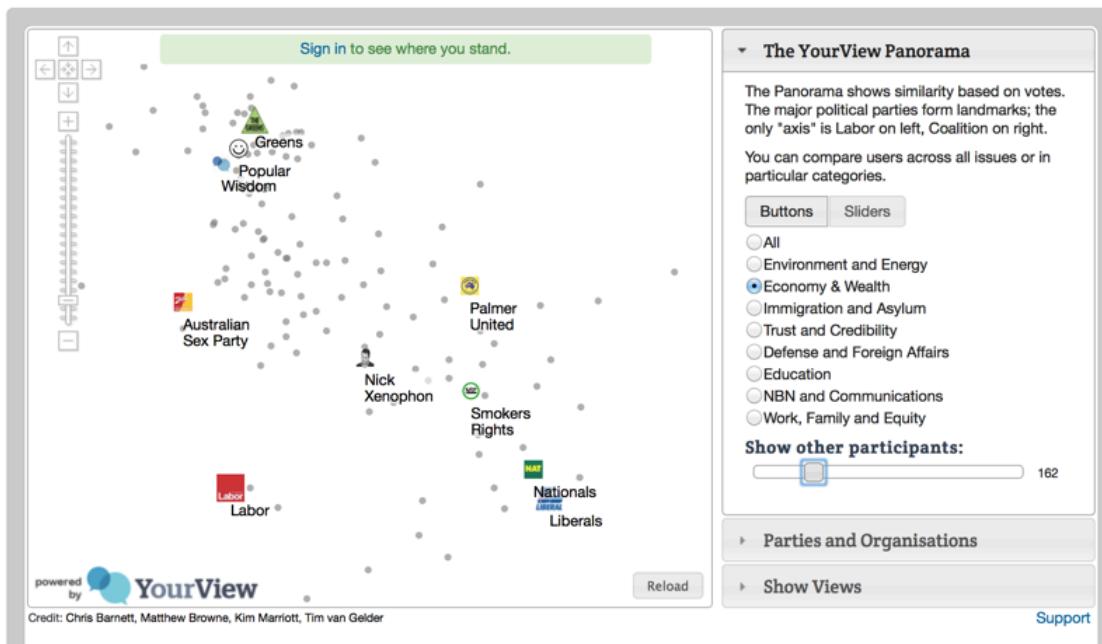

Figure 9: YourView's Election Panorama shows how peoples' votes locate in the landscape of political parties, other organisations, and YourView users. The Panorama also displays where all participants collectively stand.

In a nutshell, IBIS based deliberation platforms are collective intelligence platforms for what van Gelder terms "deliberative aggregation": they harness the deliberative power of communities and make them smarter by providing argumentation's *feedback*, *analytics* and *visualizations*, in the very attempt to improve *collective reflection* and *awareness* for the *public good*.[4]

> The DCLA community could take inspiration from the kinds of user experiences that these tools offer, since they are designed for members of the public with no formal training, outside the constraints and requirements associated with the formal educational contexts in which CSCL research is conducted. One might hypothesise that students will find such tools engaging and closer to the quality of user experience that they are accustomed to on the web at large, and in other social media specifically.
>
> The question when migrating these tools into educational contexts is what if anything needs to be added to create, and evidence, learning. That of course is determined by the kinds of learning one is aiming to deliver. Advances within *CSCL* and *AI in Education* argumentation research have much to offer to the DCLA-CIDA dialogue, as discussed later.

5. Deliberation analytics to quantify the health of a debate

Even moderately complex societal challenges can involve scores of problems to solve, hundreds of possible solutions, and thousands of arguments for and against these possible solutions. A critical challenge for making sense of such a deliberation is thus *attention allocation.* How can users know which particular topics are most in need of attention? Which areas are progressing well, and which may require some kind of intervention? Which parts part of a summary map is "mature" (i.e. comprehensively covers the key problems, solutions, and arguments) and thus ready to studied in detail?

We are seeking to develop *deliberation analytics*, by which we mean algorithms that yield analytics of quality, and map them to personalized attention mediation suggestions. If these algorithms work effectively, participants should be more aware of *where their efforts can do the most good,* to help maximize the collective intelligence of the system.

> Both the DCLA and CIDA communities should build on the advances in argumentation analytics developed at the intersection of computer-supported collaborative learning (CSCL) and artificial intelligence and education (AIED). Working in formal educational contexts, an established strand of research has sought to scaffold student argumentation and discussion skills through a variety of techniques which will be of great interest to the learning analytics community.

For example, a recent literature review (Scheuer et al., 2012) helpfully summarises the approaches that have been taken within the CSCL/AIED research communities to automated analysis of semantic argument models (

---

[4] http://timvangelder.com/category/deliberative-aggregator

Table 1).

Table 1. Overview of analysis approaches for argument models (Scheuer et al., 2012)

Analysis Approach	Description
Syntactic analysis	Rule-based approaches that find syntactic patterns in argument diagrams Systems: *Belvedere*, *LARGO*
Problem- specific analysis	Use of a problem-specific knowledge base to analyze student arguments or synthesize new arguments Systems: *Belvedere*, *LARGO*, *Rashi*, *CATO*
Simulation of reasoning and decision making processes	Qualitative and quantitative approaches to determine believability / acceptability of statements in argument models Systems: *Zeno*, *Hermes*, *ArguMed, Carneades*, *Convince Me* (Yuan et al., 2008)
Assessment of content quality	Collaborative filtering, a technique in which the views of a community of users are evaluated, to assess the quality of the contributions' textual content Systems: *LARGO*
Classification of the current modeling phase	Classification of the current phase a student is in according to a predefined process model Systems: *Belvedere*, *LARGO*

With specific reference to the need to help moderators prioritise which parts of multiple discussions require their attention, consider the ARGUNAUT tool below, which provides a moderator's interface which will be of interest to CIDA researchers (Figure 10), rendering the results of both shallower and deeper analysis of students' graphical argument maps (McLaren et al., 2010).
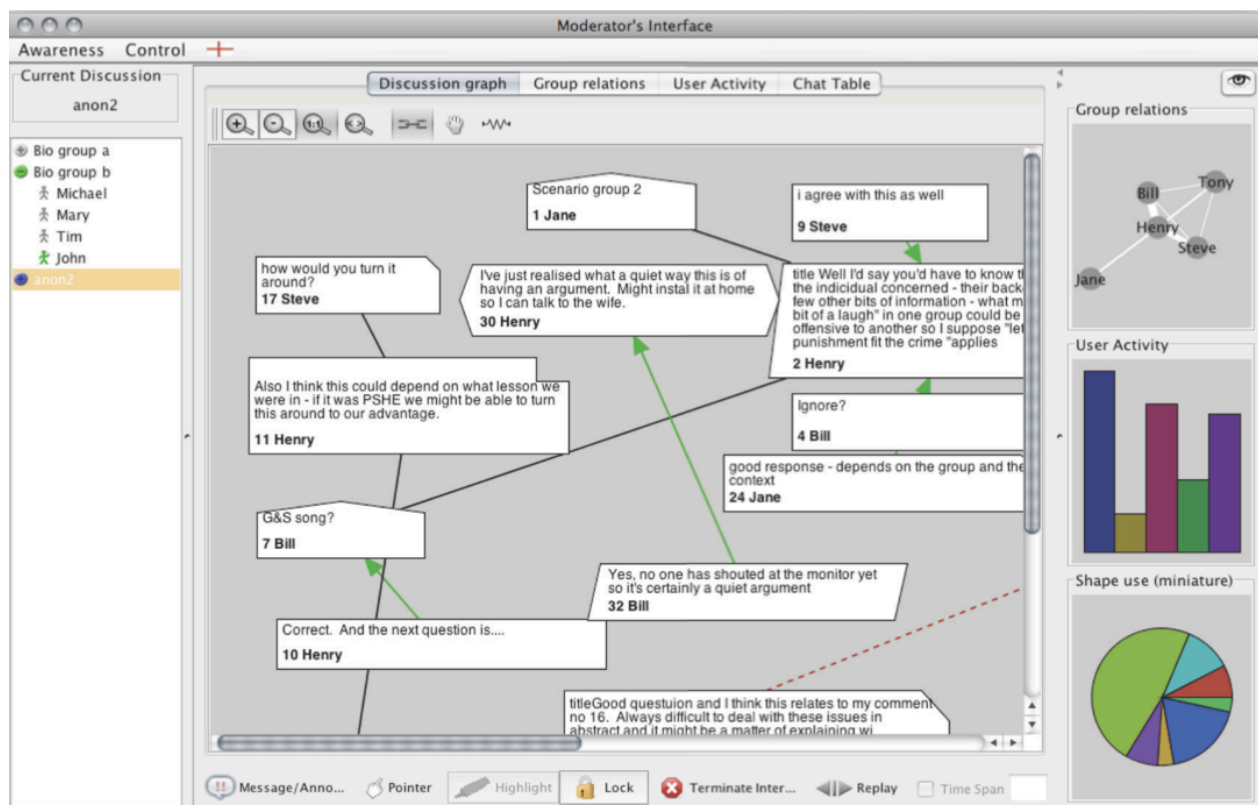
Figure 10: Prototype argument mapping moderator's interface using automated parsing of the argument graph to generate visual analytics (McLaren et al., 2010).

This is described as follows by MacLaren et al.:

"A teacher can monitor multiple ongoing discussions in parallel using a tool called the "Moderator's Interface" […] The teacher can toggle between the different e- discussions by selecting the different groups shown in the list on the left (i.e., Bio group a, Bio group b, etc.). Within each discussion, important aspects are visualized as *awareness displays*. These displays are presumed to be helpful to a teacher as he or she tries to find pedagogically meaningful aspects of the discussion (e.g., critical thinking, dialogism). In [the figure], on the right, three awareness displays are shown. The graph in the upper right shows the frequency with which students in the currently selected discussion have responded to one another's contributions, as well as which students are most in the center of the discussion through frequent responses to others' contributions. The middle right graph shows a comparison of the number of contributions made by each student, a rough indication of student engagement, while the graph in the lower right shows a comparison of the types of contributions made by students, a rough indication of *how* students are engaging with one another (e.g., Are the students asking one another questions? Making arguments?)."

In the *Process-Goal-Exception analysis* approach we describe next, several of the above techniques are considered, and we will also benefit from more deeply understanding their strengths and limits when scaled to the larger groups of citizens involved in CIDA.

Also of direct interest to CIDA, especially given the scaling challenge, MacLaren et al document a range of machine learning approaches they tested to train classifiers to recognize increasingly complex map elements, from single nodes to pairs, to larger structures. Consider this extract reporting their analytics for argumentative moves that would also be of relevance to CIDA systems Table 2.

Table 2: Extract from (McLaren et al., 2010) showing examples of argument map network structures which they trained automated classifiers to recognise

The Categories and Annotations of the Clusters

Category	*Explanation / Coding*	*Example*	*# of Ann.*
Argument + Evaluation	Relatively strong and explicit evidence of a discussant evaluating the argument of another discussant. These are relatively difficult clusters to identify; many "borderline" cases were found. Typically these are smaller clusters, usually just a pair, which may be part of larger clusters of different types (e.g., a *Chain of Opposition* cluster can contain a sub-chain in which one of the discussant evaluates the argument of another).	1st shape text (Argument): "It should be prohibited to experiment on animals because it has no special benefit. For example: one puts an unpleasant material in its eyes just to see whether it cause him damage. experiments should be prohibited also because the effect [on us] is not the same. for example: monkeys calm down if you ampute part of their brain, but humans will get handicapped as a result of this. Animals also have families and they are teared apart from their families, or that they born in the lab instead in nature." 2nd shape text (Claim), from a different person than the 1st shape: "your opinion is partly correct. indeed, there are some experiments that hurt animals and are very cruel, and are not necessary, but other experiments are usefull and important, like what I wrote in my argument. your opinions are too extreme." (*Link between shapes is "opposition"*)	36 (24 maps)
Chain of Opposition	These are linear sequences of contributions with (typically) two people arguing back and forth, each time raising counter-arguments to one another's claim or argument. The minimal pattern is of three shapes from two different users. Chains of opposition are often quite long (4-6 shapes) and some expression of disagreement, either	1st shape text (Argument), Student 1: "I agree with Riki in saying that the fetus - sick or healthy - is a human being and you can't kill a human being because he is different or disabled...." 2nd shape text (Argument), Student 2: "I think they shouldn't bring him to the world and he is not a human being with a soul in my opinion until 4 months"	20 (24 maps)

In our work, we are exploring the strengths and weaknesses of several approaches:

1. machine learning on natural language discourse, e.g. textchat (Ferguson et al., 2013)

2. rhetorical parsing of metadiscourse in natural language, e.g. discussion forums, argument maps, and research reports (Simsek et al., 2013; De Liddo et al., 2012)

3. rule-based agents monitoring IBIS structures in order to diagnose the health of a deliberation and make recommendations as to where attention might best be directed (Klein, 2003; 2012)

Since the first two approaches have been reported to the LAK community already we will elaborate the third approach, in particular because it is a new approach to DCLA, and introduces new ways of thinking about deliberation quality from fields such as organization science and group psychology.

6. Process-Goal-Exception analysis for deliberation analytics

Deliberation analytics can be identified using *Process-Goal-Exception* analysis, a technique developed by Klein (Klein, 2003). The key idea is that analytics can be viewed as the processes we put in place to identify, and respond, when a process deviates from its ideal functioning. This methodology allows us to identify process deviations and their associated responses in a systematic way that fosters complete coverage (Figure 11).
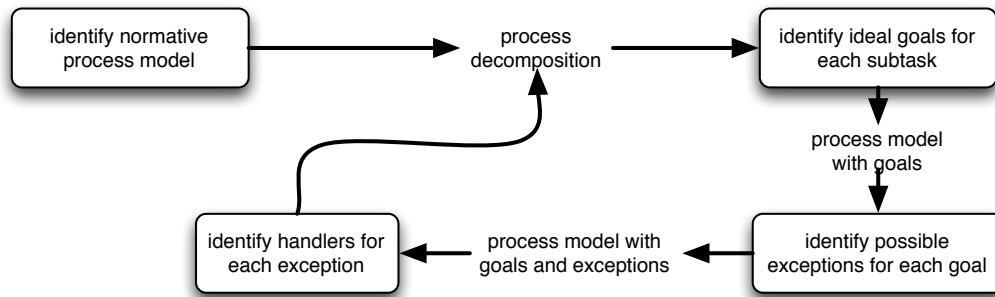
Figure 11: Process-Goal-Exception analysis, the methodology used for identifying analytics.

The results of process-goal-exception analysis are captured using the following structure:
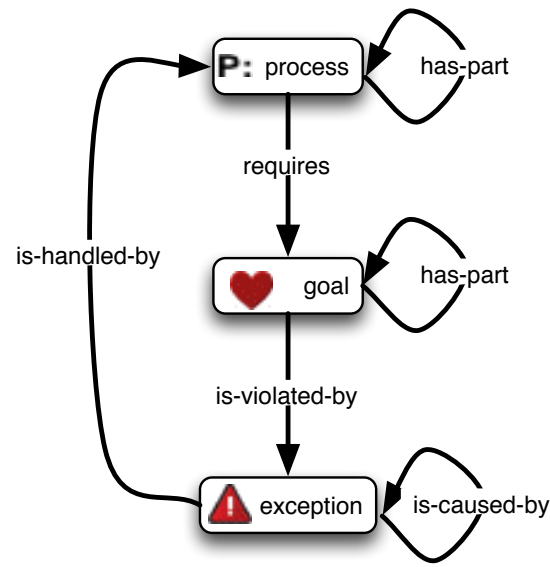
Figure 12. Entities and relationships that represent the results of process-goal-exception analysis.

Tasks in a process model are linked to their subtasks as well as to the goals they try to achieve. Goals are linked to their sub-goals, as well as to the exceptions that can violate them. Exceptions, finally, are linked to the other exceptions that may cause them, as well as to the (analytic and attention mediation) processes that can detect and resolve these exceptions.

6.1 Identify normative process model

The first step is to identify a model of how the target process should work. The core process we want to support is social innovation. Our model of this process consists of the following subtasks (Eemeren and Grootendorst, 2003; Walton and Krabbe, 1995):

1. Identify problems to solve
2. Identify possible solutions for these problems
3. Evaluate the candidate solutions
4. Select the best solution(s) from amongst the candidates
5. Enact the selected solution(s)
6. Learn from experience

This model is potentially iterative: enacting a selected solution (step 5) can, for example, lead the community to identify new problems to solve (step 1). Note also that social innovation engagements will necessarily include *all* these steps: it depends on the *purpose* of the engagement (Conklin, 2005), which can for example include:

> *Brainstorming*: create a list of solution options for a problem (step 2). Examples of this include strategic crowdsourcing in a company (before prioritization and decision by executive committee), or public consultation for a City. This can include using creativity techniques such as recombining known ideas

Argumentation: debate the relative merit of competing solution options (step 3). This can include the use of simulation and forecasting tools to assess the probable impact of the options under consideration.

Decision -making: select the preferred option from among a menu of alternatives (step 4).

Design enhancement: refine an existing solution design (i.e. start with step 5, and then loop back to step 1).

The social innovation process includes two key sub-processes. One is *harvesting*, wherein participants feed content, e.g. found in conventional social media, into the social innovation system. The harvesting sub-process consists of the following subtasks:

Harvesting Process			
Find interesting content	**Summarize as map**		
	unbundle	**tag**	**organize**

Harvesters find interesting content in social media platforms (such a Facebook, Twitter, mailing lists and blogs) where discussions about social innovations are taking place. This content is then parsed into "atoms" (i.e. individual issues, ideas, or arguments) tagged with their type (e.g. issue, idea, pro or con, evidence) and their topic area. These tagged atoms, in turn, are organized into summary maps.

The second key sub-process is *certification*, wherein moderators check the content contributed by authors in order to ensure it is organized so as to maximize it's ease-of-use for contributors and customers:

Certification Process							
Acquire post that needs attention	**Check post for correctness**				**Take action on post**		
	Bundling	**Title**	**Location**	**Substance**	**[de-] certify**	**discard**	**repair**

All new posts began with "pending" status, and become visible to the community at large only when certified. Moderators acquire posts that need attention (i.e. either pending posts, or certified posts that have been tagged as having problems) and then check the posts for correctness (i.e. whether they have substantive relevant content, are "unbundled" into individual issues ideas and arguments, have a clear title, and are placed in the correct part of the map). The moderator can then [de-]certify the post, discard it, and/or repair it.

6.2 Identify goals

The next step is to identify what each task in the process should ideally achieve: its *goals*. Our current model of the social innovation process includes the following goals, which we are modeling in a web application shown in Figure 13.
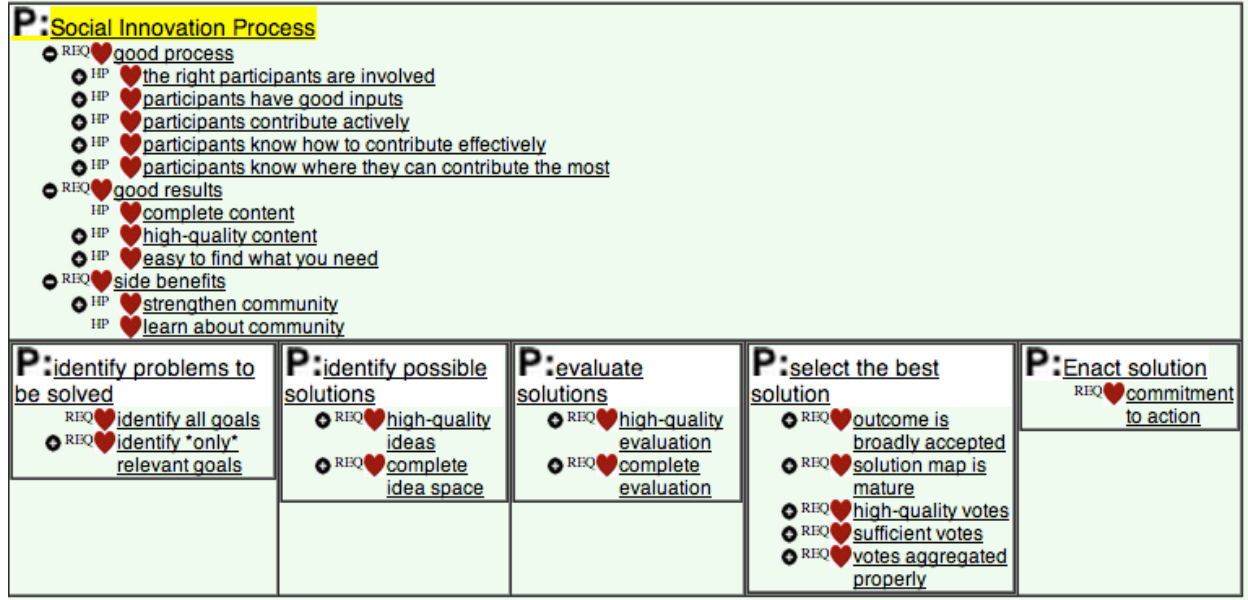
Figure 13: Top-level process model for social innovation. P:= process ♥ = goal.

This model proposes, for instance, that a *good process* for effective social innovation is one in which *the right people contribute actively and effectively to performing the most critical tasks*, to achieve *good results* defined as *complete, high-quality, well-organized content*, with the *side benefits* of strengthening and learning about the members of the user community.

6.3 Identify exceptions

For each goal, we then identify how it can be violated (the *exceptions*). The goal of having *the right participants involved*, for example, can have the following exceptions:
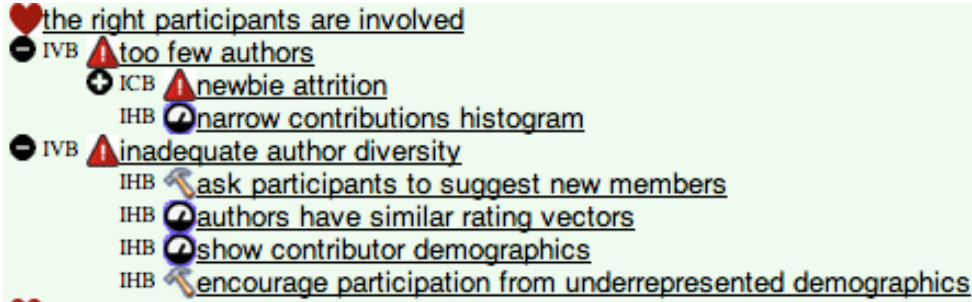
Figure 14: Exceptions and handlers for a goal in the social innovation process

⚠ **= exception** ⊘ **= analytic process** ⚒ **= handler process**

We can have *too few authors*, for example, or *inadequate diversity* in the author population.

6.4 Identify handlers

For each exception, finally, we identify handler processes that can (1) *detect* when the exception is taking place (i.e. via analytics), and (2) *resolve* that exception (i.e. via attention mediation

interventions). We can *detect too few authors,* for example, using an analytic that assess the width of the contribution activity histogram, and we *handle low author diversity* by encouraging participation from community members with underrepresented demographics (figure 5). Exception handler processes, like any other process, can themselves fail: the exception analysis process can be applied to handlers, just like any other process.

The attention-mediation interventions are personalized based on each participants' roles and past activity. The *customer* for a deliberation, for example, can be notified of topics that are mature and ready to be "harvested". A *topic manager* (responsible for ensuring a social ideation engagement achieves useful results) can be notified about which parts of the deliberation are dysfunctional (e.g. exhibit balkanization or groupthink). A *moderator* can be notified about users who consistently do (or do not) author well-structured and well-regarded posts, in order to inform training, moderator recruitment and/or rewards for top contributors. A *contributor* can be notified of content they can contribute to, such as pet ideas whose support has dropped, or posts where their ratings appear to exhibit an irrational bias. The *class* of the contributor (Preece and Shneiderman, 2009) (e.g. heavy contributors vs. peripheral users) should also likely impact which attention notification alerts they receive.

7. Defining CIDA: top-down and bottom-up

Our goal is to formalize a model of effective deliberation for social innovation, in order to assist the formalization of computational analytics and recommenders. The model being developed uses a combination of top-down and bottom-up analysis:

> *Top-down*: The model incorporates insights from the research literature for such fields as organizational science, cognitive and social psychology, political and communication science, computational social science, computer-supported cooperative work, complexity science, and economics.

> *Bottom-up*: The pain points requirements consultation reviewed earlier tested some of these against practitioners' field experience, and is now eliciting additional ideas via the web catalogue introduced below.

> We suggest that DCLA researchers might consider an analogous process in which robust theory and pedagogy can drive the definition of analytics top-down, while field practice might generate bottom-up requirements (e.g. craft moderator or mentor skill developed in the field, but lacking more formal academic validation).

One platform in which we are developing these concepts is MIT's Deliberatorium, introduced earlier. This provides analysts with a 'back end' web-based modeling environment, which allows a member of the team to search and browse the analytics 'catalogue' that is emerging.

Figure 15. A screenshot of the web-based system for process-goal-exception analysis.

Users can click on the ● and ● icons to incrementally hide and reveal components of the analysis, and click on the components themselves to see more details on each one. A search capability allows users to find components with a given type and keywords.

Our analysis efforts have resulted, at the time of writing, in a model with nearly 300 components, with a particular focus on the first three steps of the social innovation process ("identify problems", "identify possible solutions", and "evaluate solutions") since these were identified as the most critical elements by the research and community partners. Every component in the model is tagged with the system role (author, moderator, manager, customer) it is relevant to, as well as with whether it was considered, at the time of writing, to be a promising candidate for early implementation in our system.

This deliberation analytics analysis should be viewed a *living document*. We will update it throughout the project as we develop an increasingly complete understanding of how to achieve more effective large-scale social innovation processes. We have developed a web-based collective intelligence system that we will use to gather feedback about our existing analytics, as well as suggestions for new analytics, from the argumentation-mapping community (Figure 16).

> Potential communities for providing feedback on the analytics include the many users of argument-mapping systems reviewed earlier, but potentially, the DCLA community might begin to formalize notions of 'quality discourse' by submitting proposals.
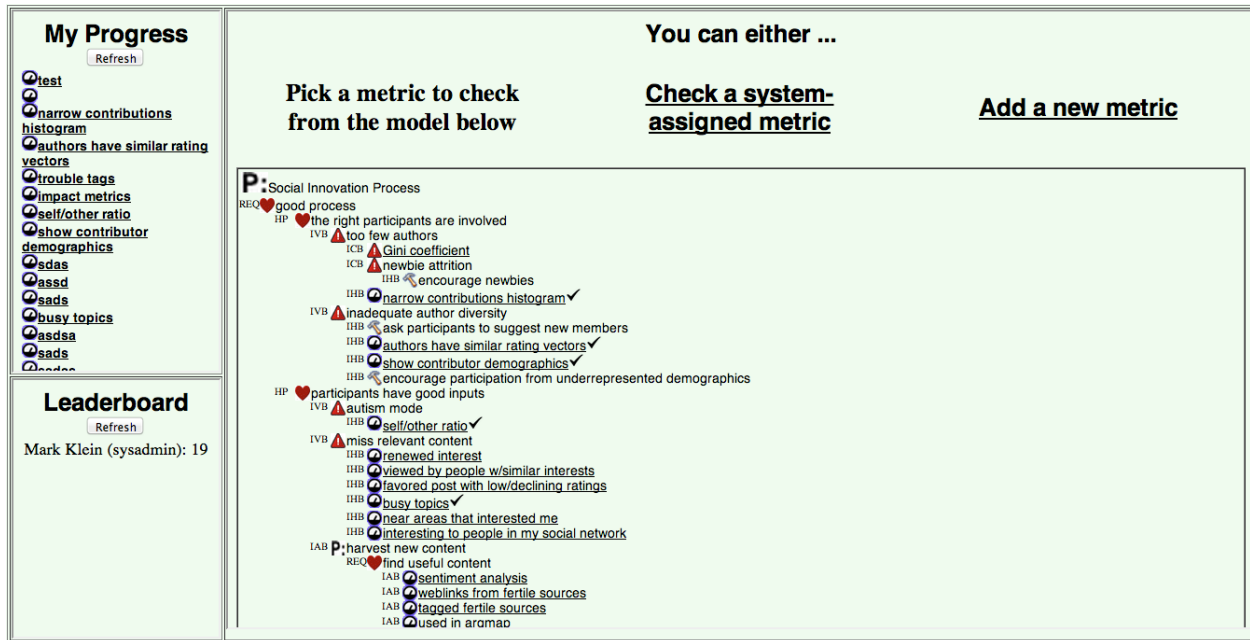
Figure 16. A screenshot of the web-based collective intelligence system for tapping the argument-mapping community's knowledge about deliberation analytics.

7.1.1 An Exception Analysis of Large-Scale Deliberation

We now present examples of CIDA based on this exceptions analysis approach, taken from (Klein, 2012). Our deliberation model formalizes a straightforward view of what makes up a good collaborative decision-making process. The main steps (**P:**) and goals (♥) in the process are as follows.
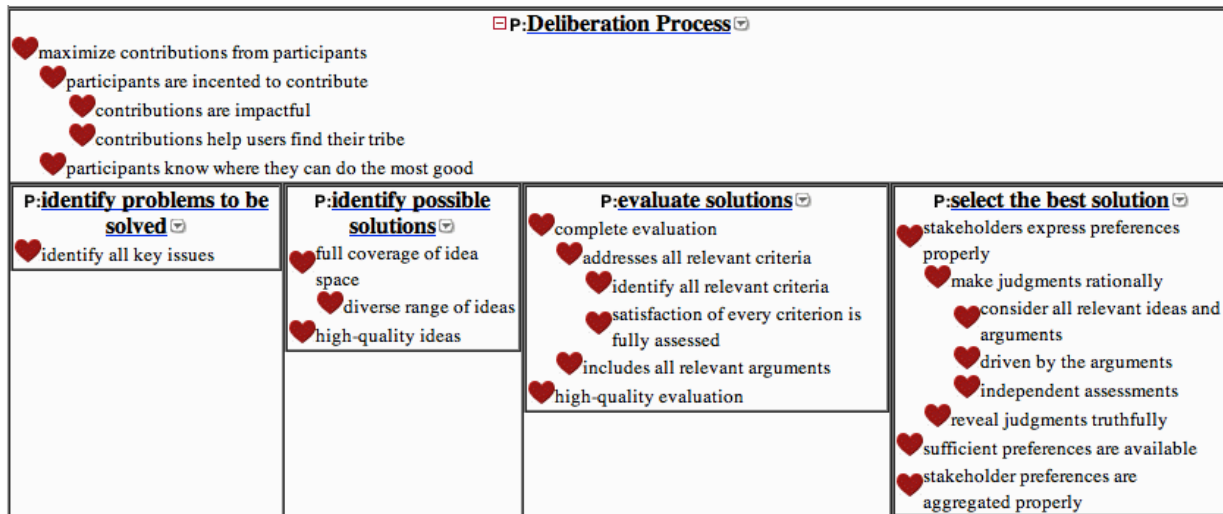
Figure 17. Normative process model for collaborative decision-making.

The complete analysis of this model (available upon request) is too large to review in its entirety in this paper. We will provide instead, in the following paragraphs, illustrative examples of the exceptions and analytics derived from our analysis.

Analytics for "Identify Possible Solutions": One of the key goals for any deliberation process is to achieve excellent coverage of the space of possible solution ideas, but this goal may not always be achieved, leading to such exceptions (⚠) as "dominated by narrow range of ideas.

P:*Deliberation Process*
   has-part **P**:*identify possible solutions*
      ⊟*requires* ❤full coverage of idea space
        ⊟*has-part* ❤diverse range of ideas
          ⊟*raises* ⚠dominated by a narrow range of ideas
            ☐*is-handled-by* **P**:detect: non-disjoint communities of idea authors
            ☐*is-handled-by* **P**:detect: vocabulary diversity across ideas
            ☐*is-handled-by* **P**:detect: narrow contributions histogram
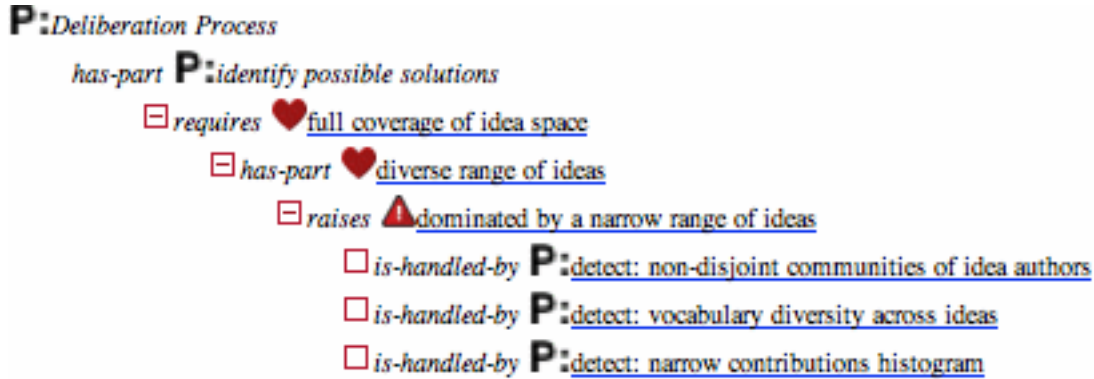
Figure 18. (partial) exception analysis for "identify possible solutions"

We have identified several possible analytics for detecting this exception. One is to look for issues where the ideas come from "non-disjoint" author sets i.e. where all the authors tend to agree about the pros and cons for different ideas and therefore probably share an intellectual frame. This is straightforward to assess in an argument map: we can for example perform vector orthogonalization (Householder, 1958) on participants' rating vectors, followed by a simple vector distance calculation, to assess how much the opinions for different users diverge. If we have only a moderate degree of opinion divergence for the authors of the ideas for a given issue, this suggests that the ideas captured so far may be relatively less diverse. Another way of detecting this exception involves measuring the use of shared vocabulary in the ideas for a given issue. If there is heavy use of shared terminology, this again suggests that the ideas are only moderately diverse. A third approach is simply to count the number of authors that contributed ideas for an issue: if a small number of authors contributed the bulk of the ideas, this suggests the diversity may be relatively low.

Another key goal for solution generation is to identify *high-quality* (i.e. promising) solutions.

P:*Deliberation Process*
   has-part **P**:*identify possible solutions*
      ⊟*requires* ❤high-quality ideas⚙
        ⊟*raises* ⚠issue has poor quality ideas⚙
          ☐*is-handled-by* **P**:detect: ideas have poor ratings⚙
          ☐*is-caused-by* ⚠idea sabotage⚙
            ☐*is-handled-by* **P**:detect: edit wars by people who disagree⚙
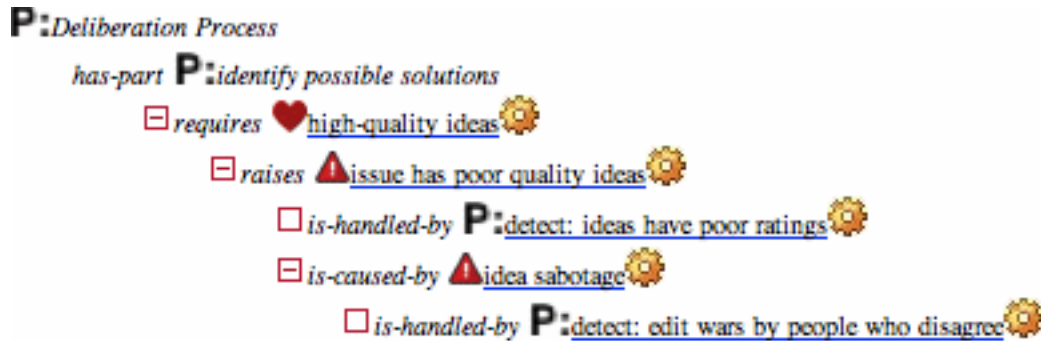
Figure 19. (partial) exception analysis for "identify possible solutions"

One straightforward analytic for identifying low-quality solutions is, of course, to look for idea posts with low aggregate rating scores. Note that ratings for posts in a argument map will generally be more meaningful than those in other forms of social media (e.g. wikis, web forums) because each rating is attached to a *single* point (issue, idea, or argument) rather than a *collection* of multiple points that may vary widely in quality.

A second analytic for assessing solution quality is based on *process,* rather than *outcome.* As we noted above, in our system argument map posts can be edited openly, like wiki pages, allowing the community to collaboratively improve each post. In controversial topic areas, however, open authoring can lead to destructive "edit wars", where people who disagree about a post repeatedly replace competitors' content with their own perspective. This "idea sabotage" can be detected by looking for alternating edits by users that appear to have divergent opinions (based on their rating behavior: see above) about the issue they are proposing solutions for.

Even these simple examples make it clear that it can be difficult to define analytics that *guarantee* that a given exception is taking place. Doing so would often require sophisticated natural language understanding technology well beyond the current state of the art. It *does* seem feasible, however, to use well-known data mining techniques (e.g. word frequency statistics, network analysis, vector orthogonalization and so on), taking advantage of the additional semantics provided by argument maps, to define analytics that assess whether or not an exception *might* be taking place. The deliberation participants can then decide for themselves whether the exceptions are indeed in play and need to be addressed.

Analytics for "Evaluate Solutions": When evaluating potential solutions for a problem, one key goal is to ensure that the evaluation of the ideas is *complete:* i.e. all key arguments for and against an idea have been identified.
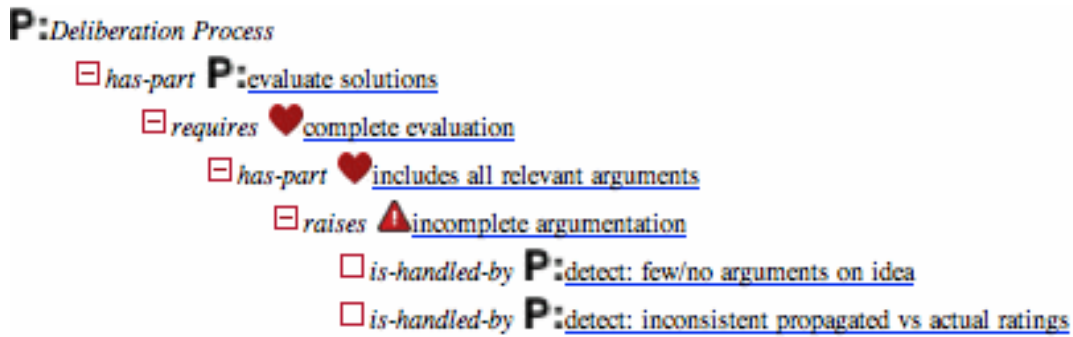
Figure 20. (partial) exception analysis for "evaluate solutions"

One straightforward analytic for the exception "incomplete argumentation" is to simply identify ideas which have few or no arguments attached to them. A second, more subtle, approach, is to assess whether a user has given ratings for ideas that are inconsistent with the ratings they gave the underlying arguments. We can use such techniques as Bayesian inference (Bolstad, 2010) to propagate a user's ratings for arguments up the argument map to predict how the user *should* have rated the ideas these arguments address. If there is a large divergence between a user's predicted and actual ratings for an idea, that suggests that the user has not yet entered arguments that are compelling to him or her.

Analytics for "Select Solutions": this aspect of the deliberation process (when the community is asked to converge upon one or more agreed-upon solutions for the key issues) has a range of challenging

exceptions. We will focus, here, on the exceptions associated with the sub-goal "stakeholders make judgments rationally".

P: *Deliberation Process*
  has-part **P:** *select the best solution*
    requires ♥ *stakeholders express preferences properly*
      ⊟ *has-part* ♥ make judgments rationally
        ⊟ *has-part* ♥ consider all relevant ideas and arguments
          ⊟ *raises* ⚠ premature decision
            ☐ *is-handled-by* **P:** detect: phases: issues to ideas to pros to cons to refinement to quiescence
          ⊟ *raises* ⚠ groupthink
            ☐ *is-handled-by* **P:** detect: attentional focusing
          ⊟ *raises* ⚠ balkanization
            ☐ *is-handled-by* **P:** detect: clustering user profiles
      ⊟ *has-part* ♥ independent assessments
        ⊟ *raises* ⚠ select based on popularity (non-independent ratings)
          ☐ *is-handled-by* **P:** detect: high correlation between early and final rating orders

Figure 21. (partial) exception analysis for "select the best solution"

One key exception involves participants converging on one or a few highly rated solution ideas for an issue before a full consideration of the relevant ideas and arguments i.e. before the argument map for that issue is "mature". Deliberation maturity can be assessed in several ways. Deliberations, especially in argument map contexts, tend to evolve from defining issues to proposing ideas to identifying trees of pro and con arguments. Early activity also tends to be characterized by creating new posts, and later activity by refining them, followed eventually by relative quiescence. An actively growing map with a few relatively shallow argument branches, as well as few edits per post, is thus probably relatively immature. We can also assess opinion churn for each issue (i.e. whether the highest-rated ideas for individuals, as well as the community as a whole, are still changing rapidly or not).

A second key exception is "groupthink", which can be defined as a group dedicating the bulk of its attention to refining a single solution idea, often the first one endorsed by an influential figure, rather than comparing several alternatives in depth. This is straightforward to assess in an argument map because we can measure when one idea under an issue is receiving the bulk of the community's attention (views, rates, edits and additions) while competing ideas and their underlying arguments remain largely untouched.

A third important exception is "balkanization", wherein a community divides itself into sub-groups where members of each group agree with one other but ignore contributions of groups with competing ideas. This exception can be assessed using a analytic that looks for clustering in the attention allocation of the participants in the discussion for a given issue.

The final exception we will consider is "non-independent ratings". It has been shown that when people are asked to rate competing ideas, if they can see the ratings made to date (e.g. they see the ideas in popularity-sorted order), then the first ideas that happen to get a rating advantage tend to become the eventual winners – they "lock in" to the winning position - even if they are worse than ideas that appeared later or started with lower ratings (Salganik et al., 2006). The underlying problem is that

people attend to, and rate, ideas based on their popularity rather than their inherent merits. This exception can be detected using a analytic that checks whether the popularity order for a set of competing ideas has remained relatively unchanged as the deliberation has progressed in maturity, especially if the idea popularity ranking diverges from what would be predicted by looking at the underlying argument ratings.

7.1.2 Generating Attention Mediation Suggestions

When a deliberation is actually taking place, in our approach, analysis algorithms generate personalized suggestions for users concerning which posts they might want to look at in order to contribute most effectively to the deliberation at hand. These suggestions are identified by (1) maintaining continuously updated values for the deliberation analytics (such as those we identified above), (2) identifying the exceptions that seem to be active based on these analytic values, and (3) notifying the user about the exceptions they can help resolve.

Calculating analytics: Analytics are calculated by running queries over the database created by users as they deliberate using the Deliberatorium. This database includes *entities* representing every user and post (issue, idea, argument, and comment), *relations* representing the connections between posts, and time-stamped *events* capturing all changes (creation, editing, moving) to these entities and relations. Since the database has a network structure, we use a graph pattern-matching language to express our queries. The current system uses PQL (Klein and Bernstein, 2004) but it could use other graph query languages such as SPARQL (Group, 2008) for the same purpose. One analytic, for example, involves looking for ideas or arguments where the propagated and actual ratings differ widely. This could be captured using the following PQL query.

Table 3: Process Query Language (PQL) graph query looking for ideas or arguments where the propagated and actual ratings differ widely

PQL query code	Explanation
(entity ?e isa (or idea pro con))	Binds idea, pro and con entities to variable ?e
(attribute average-rating of ?e is ?av)	Binds *actual* community rating to ?av
(attribute propagated-rating of ?e is ?pv)	Binds *propagated* rating to ?pv
(let ?diff be (abs (- ?av ?pv)))	Find the difference between the propagated and actual ratings
(when (> ?diff 2))	Ensures the difference is more than 2 (on a scale of 5) – smaller differences do not trigger this analytic
(let ?strength be ?diff)	Sets the analytic's "strength" to be a function of the size of the rating disconnect

Each analytic can be viewed, in effect, as a kind of "daemon" that continually scans the deliberation database for instances that trigger it. Every daemon has a "strength" value, representing how strongly it was triggered. In the above, for example, the daemon's strength is a function of the difference between the actual and propagated ratings.

Identifying exceptions: Exceptions are presumed to be occurring when their associated daemon(s) are firing strongly for a given part of the deliberation map. The more strongly the exception's daemons fire, the more likely the exception is considered to be. So, for example (see figure 7), if both the "inconsistent ratings" and "few arguments" daemons fire strongly for a given post in the map, the "incomplete argumentation" exception is considered to be active there.

It is difficult, as we noted above, to define analytics that guarantee a given exception is taking place, so there are likely to be "false positives" (i.e. cases where the daemons incorrectly flag an exception somewhere when it is not in fact taking place). Given that, users have to assess, for themselves, whether an exception is in fact "real" and requires attention. If such false positives are too numerous, however, it undercuts the whole point of the approach, which is after all to greatly reduce the user effort required to find the areas in the deliberation that can benefit from their particular knowledge and skills. This represents an open challenge for us at this point, but we believe that one promising strategy is to implement "crowd-sourced" exception validation. The idea is that, when a user views a suggested post, he or she can check off whether or not the proposed exceptions are actually taking place. As more and more users do so, the system can begin to highlight only those posts whose exceptions have been widely validated, reducing the impact of false positives. We can even imagine, eventually, applying machine learning techniques to this user validation data to help develop better exception detection analytics.

It is of course probably over-ambitious to imagine that software algorithms will be able to fully identify all departures from "ideal" deliberation without explicit human guidance. People, especially the "customers" for a deliberation, should be able to augment these suggestions. This is straightforward to do within our current structure: we need only create additional daemons that fire, more or less strongly, for parts of the deliberation map specified by human users.

Notifying users: Once the active exceptions are identified, users are provided with personalized suggestions corresponding to the exceptions they seem well-suited to help resolve. This matching process is conducted based on the user's roles, interests and skills.

Roles: The Deliberatorium includes several user roles, including *author*, *moderator* (responsible for assuring the map is well-structured), *topic manager* (responsible for assuring the deliberation is progressing effectively) and *customer* (the individuals seeking solutions to the problems being discussed). Different roles will be interested in different exceptions. Topic managers, for example, would probably be interested in the "idea sabotage" exception mentioned above because they have the power and responsibility to do something about it (e.g. by managing who is allowed to edit a given post).

Interests: Our model of user interests is based on an analysis of their deliberation activity, assuming that a user is interested in a post if they devoted a substantial amount of effort to viewing, rating, commenting on, or editing it (or its neighbors) in the past, and:

- the post is currently experiencing high/growing activity
- the post is receiving community scores that differ strongly from the rating the user gave the post
- the user expressed a strong opinion about the post (positive or negative) and it has a high/growing controversy score (i.e. the post itself has a wide rating variance and/or there are many highly rated pros and cons underneath it)

We also cluster users according to their map activity, so that we can identify whether a post has been interesting to someone who has shared interests and even opinions with you in the past, as with e-commerce recommender systems (Adomavicius and Tuzhilin, 2005).

Skills: A user's skills can be inferred by identifying the parts of the map in which they contributed to highly-rated posts.

Once the personalized suggestions are generated, they are presented to the users in the form of a deliberation map subset wherein the suggested posts are presented along with the *reasons* for their selection (i.e. listing the exceptions considered active for those posts, as well as why the user might be interested in them).
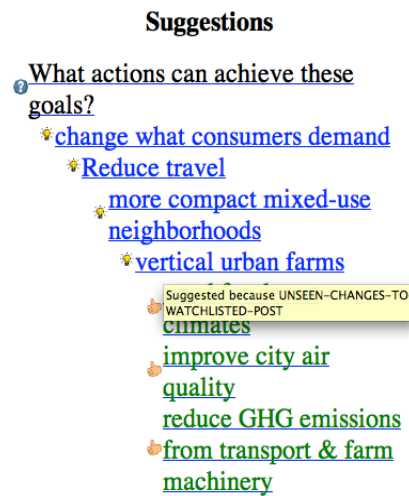

Figure 22. The personalized suggestions display.

More highly recommended posts are made visibly salient by appearing in a larger font. Users are then free to contribute to resolving the exceptions they feel most interested and capable of addressing.

8.   Concluding thoughts

This working paper has sought to bring together a number of threads of interest to the DCLA community. As has often been noted, learning analytics as a field is very much a confluence of existing tributaries, which flow from long standing 'mountains' in the research landscape. DCLA is enriched by taking into account the insights from *Hypertext, AIED, CSCL, EDM,* and now we propose, *CI,* so that we are not reinventing wheels. The challenge then is to integrate and scale these approaches to the larger datasets that learning analytics is interested in dealing with (although small and medium participant numbers are also highly relevant to many educational contexts), connect these with other educational datasets, and mainstream the approaches, so that the capabilities being demonstrated in small scale lab prototypes 'go to market'. *DCLA technologies* need to be connected to the mainstream learning environments, and *DCLA literacy* needs to be nurtured in learners and educators.

CIDA work brings design priorities that DCLA researchers should find relevant to the creation of authentic learning experiences: a focus on intuitive interfaces, untrained users, engaging visualizations, and complex problems that many citizens care about deeply. In turn, the analytics falling under the broad umbrella of DCLA (building, as this paper has sought to emphasise, on CSCL/AIED research) have made advances that CIDA research could build on. While CIDA is less interested in evidencing formal educational learning objectives, it is striking that argumentation researchers in education motivate their work precisely because critical thinking and the ability to engage constructively with widely differing viewpoints are such important qualities — not just for scientists or academics, but for all citizens to engage in participatory innovation of the sort demanded by the complexity of today's societal challenges, and in the broader picture, a healthy democracy. The DCLA-CIDA dialogue should therefore be a vigorous one, with potentially high impact within and beyond formal education.

9. Acknowledgements

10. References

Adomavicius G and Tuzhilin A. (2005) Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering* 17: 734-749.

Andrews R. (2010) *Argumentation in Higher Education: Improving practice through theory and research*: Routledge.

Buckingham Shum S. (2003) The Roots of Computer-Supported Argument Visualization. In: Kirschner PA, Buckingham Shum S and Carr C (eds) *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making.* London: Springer-Verlag, 3-24.

Buckingham Shum S and Deakin Crick R. (2012) Learning Dispositions and Transferable Competencies: Pedagogy, Modelling and Learning Analytics. *Proc. 2nd International Conference on Learning Analytics & Knowledge.* Vancouver, 29 Apr-2 May: ACM: New York, 92-101. DOI: http://dx.doi.org/10.1145/2330601.2330629 Eprint: http://oro.open.ac.uk/32823

Buckingham Shum S, Selvin A, Sierhuis M, Conklin J, Haley C and Nuseibeh B. (2006) Hypermedia support for argumentation-based rationale: 15 years on from gIBIS and QOC. In: Dutoit A, McCall R, Mistrik I and Paech B (eds) *Rationale Management in Software Engineering.* Berlin: Springer-Verlag, 111–132.

CICDA. (2012) *Collective Intelligence as Community Discourse and Action. A CSCW2012 Workshop: ACM Conference on Computer-Supported Cooperative Work, Feb 11-15, 2012, Seattle, WA.* Available at: http://events.kmi.open.ac.uk/cscw-ci2012

Claxton G. (2001) *Wise Up: Learning to Live the Learning Life,* Stafford: Network Educational Press.

Conklin J. (2003) Dialog Mapping: Reflections on an Industrial Strength Case Study. In: Kirschner PA, Buckingham Shum S and Carr C (eds) *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making.* London: Springer-Verlag, 117-136.

Conklin J. (2006) *Dialogue Mapping: Building Shared Understanding of Wicked Problems,* Chichester: Wiley.

Cook MB and Smallman HS. (2007) Visual Evidence Landscapes: Reducing Bias in Collaborative Intelligence Analysis. *Proc. 51st Human Factors & Ergonomics Society Annual Meeting.* Baltimore, Maryland, October 1-5, 2007.

Culmsee P and Awati K. (2011) *The Heretic's Guide to Best Practices: The Reality of Managing Wicked Problems in Organisations,* Bloomington IN: iUniverse Inc.

De Liddo A and Buckingham Shum S. (2010) Cohere: A Prototype for Contested Collective Intelligence. *Workshop on Collective Intelligence in Organizations, ACM CSCW 2010: Computer-Supported Cooperative Work.* Savannah, GA: Eprint: http://oro.open.ac.uk/19554.

De Liddo A and Buckingham Shum S. (2013a) The Evidence Hub: Harnessing the Collective Intelligence of Communities to Build Evidence-Based Knowledge. *Large Scale Ideation and Deliberation Workshop, 6th International Conference on Communities & Technologies.* Munich, Germany, 29 June-2 July. Open Access Eprint: http://oro.open.ac.uk/38002

De Liddo A and Buckingham Shum S. (2013b) Improving Online Deliberation with Argument Network Visualization. *Digital Cities 8 Workshop, 6th International Conference on Communities & Technologies.* Munich, Germany, 29 June-2 July. Open Access Eprint: http://oro.open.ac.uk/38003

De Liddo A, Buckingham Shum S, Quinto I, Bachler M and Cannavacciuolo L. (2011) Discourse-Centric Learning Analytics. *Proc. 1st International Conference on Learning Analytics & Knowledge.* Banff, CA: ACM: New York, 22-33. DOI: http://dx.doi.org/10.1145/2090116.2090120 Open Access Eprint: http://oro.open.ac.uk/25829

De Liddo A, Sándor Á and Buckingham Shum S. (2012) Contested Collective Intelligence: Rationale, Technologies, and a Human-Machine Annotation Study. *Computer Supported Cooperative Work* 21: 417-448. Open Access Eprint: http://oro.open.ac.uk/31052

Deakin Crick R. (2009) Pedagogical challenges for personalisation: integrating the personal with the public through context   driven enquiry. *Curriculum Journal* 20: 185-306.

EC-Horizon2020. *CAPS 2020: Collective Awareness Platforms for Sustainability and Social Innovation. European Commission ICT Programme.* Available at: http://caps2020.eu

Eemeren FHv and Grootendorst R. (2003) *A Systematic Theory of Argumentation: The Pragma-dialectical Approach*: Cambridge University Press.

Ferguson R, Wei Z, He Y and Buckingham Shum S. (2013) An Evaluation of Learning Analytics to Identify Exploratory Dialogue in Online Discussions. *Proc. 3rd International Conference on Learning Analytics & Knowledge.* Leuven, BE: ACM, 85-93 Open Access Eprint: http://oro.open.ac.uk/36664

Fullan M and Langworthy M. (2013) Towards a New End: New Pedagogies for Deep Learning. New Pedagogies for Deep Learning: A Global Partnership. http://www.newpedagogies.info/wp-content/uploads/2014/01/New_Pedagogies_for_Deep_Learning_Whitepaper.pdf

Gardner H. (1983) *Frames of Mind: The Theory of Multiple Intelligences,* New York: Basic Books.

Gilby N, Hamlyn B, Hanson T, Romanou E, Mackey T, Clark J, Trikka N and Harrison M. (2008) National Survey of Parents and Children: Family Life, Aspirations and Engagement with Learning in 2008. London: UK Dept. Children, Schools & Families. https://http://www.education.gov.uk/publications/RSG/Parentscarersandfamilies/Page7/DCSF-RR059

Hewlett Foundation. (2014) *Deeper Learning Initiative: William & Flora Hewlett Foundation.* Available at: http://www.hewlett.org/programs/education-program/deeper-learning

Iandoli L, Klein M and Zollo G. (2009) 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* 1: 69-91.

Iandoli L, Quinto I, De Liddo A and Buckingham Shum S. (2013) Socially augmented argumentation tools: Rationale, design and evaluation of a debate dashboard. *International Journal of HUman-Computer Studies* 72: 289-319. http://dx.doi.org/10.1016/j.ijhcs.2013.08.006

Klein M. (2003) A Knowledge-Based Methodology for Designing Reliable Multi-Agent Systems. In: Giorgini P, Mueller JP and Odell J (eds) *Agent-Oriented Software Engineering IV*. Springer-Verlag, 85 - 95.

Klein M. (2012) Enabling Large-Scale Deliberation Using Attention-Mediation Metrics. *Computer Supported Cooperative Work* 21: 449-473. http://dx.doi.org/10.1007/s10606-012-9156-4

Klein M, Spada P and Calabretta R. (2012) Enabling Deliberations in a Political Party Using Large-Scale Argumentation: A Preliminary Report. *Workshop on Large-scale Idea Management and Deliberation, COOP2012 Conference (May 29 2012, Marseille)*. http://lsdeliberation.wikispaces.com/papers

Kuhn D. (1991) *The Skills of Argument,* Cambridge: Cambridge University Press.

Lillis T and Turner J. (2001) Student writing in higher education: contemporary confusion, traditional concerns. *Teaching in Higher Education* 6: 57-68.

McLaren BM, Scheuer O and Mikšátko J. (2010) Supporting Collaborative Learning and e-Discussions Using Artificial Intelligence Techniques. *International Journal of Artificial Intelligence in Education* 20: 1-46.

Novak JD. (1998) *Learning, creating, and using knowledge: concept maps as facilitative tools in schools and corporations,* Mahwah, NJ: Lawrence Erlbaum Associates.

Okada A, Buckingham Shum S and Sherborne T. (2008) Knowledge Cartography: Software Tools and Mapping Techniques. London, UK: Springer.

Perkins D, Jay, E. & Tishman S. (1993) Beyond Abilities: A Dispositional Theory of Thinking,. *Merrill-Palmer Quarterly* 39: 1-21.

Renton A and Macintosh A. (2007) Computer supported argument maps as a policy memory. *The Information Society* 23: 125–133. Open Access Eprint: http://eprints.whiterose.ac.uk/3744

Rider Y and Thomason. (2008) Cognitive and Pedagogical Benefits of Argument Mapping: L.A.M.P. Guides the Way to Better Thinking. In: Okada A, Buckingham Shum S and Sherborne T (eds) *Knowledge Cartography: Software Tools and Mapping Techniques.* London: Springer.

Rittel H. (1972) Second Generation Design Methods. *Design Methods Group 5th Anniversary Report: DMG Occasional Paper* 1.

Rittel H. (1984 (1972)) Second Generation Design Methods (reprint 1972). In: Cross N (ed) *Developments in Design Methodology.* Chichester: J. Wiley & Sons 317-327.

Rittel HWJ and Webber MM. (1973) Dilemmas in a General Theory of Planning. *Policy Sciences* 4: 155-169.

Scheuer O, McLaren BM, Loll F and Pinkwart N. (2012) Automated Analysis and Feedback Techniques to Support Argumentation: A Survey. In: McLaren BM and Pinkwart N (eds) *Educational Technologies for Teaching Argumentation Skills.* Bentham Science Publishers, 71–124.

Schulz-Hardt S, Frey D, Lüthgens C and Moscovici S. (2000) Biased information search in group decision making. *Journal of Personality and Social Psychology* 78: 655-669.

Selvin A, Buckingham Shum S and Aakhus M. (2010) The Practice Level in Participatory Design Rationale: Studying Practitioner Moves and Choices. *Human Technology: An Interdisciplinary Journal of Humans in ICT Environments* 6: 71-105. Eprint: http://www.humantechnology.jyu.fi

Sierhuis M and Buckingham Shum S. (2008) Human-agent knowledge cartography for e-science: NASA field trials at the Mars Desert Research Station. In: Okada A, Buckingham Shum S and Sherborne T (eds) *Knowledge Cartography: Software Tools and Mapping Techniques.* London: Springer (Advanced Information and Knowledge Processing Series), 287–306.

Simsek D, Buckingham Shum S, Sándor Á, De Liddo A and Ferguson R. (2013) XIP Dashboard: Visual Analytics from Automated Rhetorical Parsing of Scientific Metadiscourse. *1st International Workshop on Discourse-Centric Learning Analytics, at 3rd International Conference on Learning Analytics & Knowledge.* Leuven, BE (Apr. 8-12, 2013). Open Access Eprint: http://oro.open.ac.uk/37391

Sunstein CR. (2006) *Infotopia: How Many Minds Produce Knowledge*: Oxford University Press.

Tate A, Buckingham Shum S, Dalton J, Mancini C and Selvin A. (2007) Co-OPR: Design and Evaluation of Collaborative Sensemaking and Planning Tools for Personnel Recovery. Technical Report KMI-06-07, Knowledge Media Institute, The Open University UK. Open Access Eprint: http://kmi.open.ac.uk/publications/pdf/KMI-TR-06-07.pdf

Tversky A and Kahneman D. (1974) Judgments under uncertainty. Heuristics and biases. Science. *Science Education* 185: 1124-1131.

van Gelder T. (2003) Enhancing Deliberation Through Computer-Supported Argument Visualization. In: Kirschner PA, Buckingham Shum S and Carr C (eds) *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making.* London: Springer-Verlag.

Walton D and Krabbe ECW. (1995) *Commitment in dialogue: Basic concepts of interpersonal reasoning,* Albany, NY: State University of New York Press.

Walton D, Reed C and Macagno F. (2008) *Argumentation Schemes,* Cambridge: Cambridge University Press.

Whole Education. (2014) Available at: http://WholeEducation.org

Willms JD, Friesen S and Milton P. (2009) What did you do in school today? Transforming classrooms through social, academic, and intellectual engagement. Toronto: First National Report, Canadian Education Association. http://www.ccl-cca.ca/pdfs/otherreports/WDYDIST_National_Report_EN.pdf

Yazzie-Mintz E. (2009) Charting the Path from Engagement to Achievement: A Report on the 2009 High School Survey of Student Engagement. Center for Evaluation & Education Policy, Indiana University. http://ceep.indiana.edu/hssse

Yuan T, Moore D and Grierson A. (2008) A Human-Computer Dialogue System for Educational Debate: A Computational Dialectics Approach. *Intl. Journal of Artificial Intelligence in Education* 18: 3-26.