

# A tale of struggles: an evaluation framework for transitioning from individually usable to community-useful online deliberation tools

Lucas Anastasiou  
lucas.anastasiou@open.ac.uk  
The Open University  
Milton Keynes, United Kingdom

Barbara Brayshay  
barbara.brayshay@rhul.ac.uk  
Royal Holloway University of London  
London, United Kingdom

Aldo de Moor  
ademoor@communitysense.nl  
CommunitySense  
Tilburg, The Netherlands

Anna De Liddo  
anna.deliddo@open.ac.uk  
The Open University  
Milton Keynes, United Kingdom

## ABSTRACT

In this paper, we discuss the importance of providing socio-technical support for technology-mediated public interest debates and outline the principles that need to be considered to ensure healthy and fruitful deliberation in online discussion processes. We highlight the challenge of transitioning from individually usable only to community-useful online deliberation tools and we propose a 4-layer evaluation framework for online deliberation technologies that take into consideration usability, discussion quality, debate quality, and societal context, under the prism of participants' sense-making. We present a new online deliberation tool (BCause), enhanced with computational aids for sensemaking support that conforms with our evaluation framework guidelines. We also present the hurdles encountered in two use case applications of BCause and reflect on the real-world challenges of deploying a novel deliberation tool within real communities.

## CCS CONCEPTS

• **Human-centered computing** → Collaborative and social computing design and evaluation methods; • **Information systems** → Summarization; Web interfaces.

## KEYWORDS

online deliberation, online discussion, computer supported tools, sensemaking

### ACM Reference Format:

Lucas Anastasiou, Aldo de Moor, Barbara Brayshay, and Anna De Liddo. 2023. A tale of struggles: an evaluation framework for transitioning from individually usable to community-useful online deliberation tools. In *The 11th International Conference on Communities and Technologies (C&T) (C&T '23)*, May 29–June 2, 2023, Lahti, Finland. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3593743.3593771>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*C&T '23*, May 29–June 2, 2023, Lahti, Finland

© 2023 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0758-2/23/05.

<https://doi.org/10.1145/3593743.3593771>

## 1 INTRODUCTION

Providing socio-technical support for public interest debates requires thorough consideration of various principles to ensure healthy and fruitful deliberation. *Deliberation* is the careful discussion before decision, and it can be defined as the thorough dialogical assessment of the reasons for and against a measure before a decision is made. As such debate processes are complex by their very nature, they require a rich mix of elements to become effective and serve the promise of deliberative democracy. Civic engagement projects face serious cooperative discussion challenges related to not just the scale of the participation but also their political context [34]. Building useful political deliberation environments requires overcoming many issues to enable effective public sphere interactions [48]. Well-designed socio-technical support systems and processes are therefore needed to enable effective large-scale argumentation for discourse communities [1].

In modern society, the public sphere has become a platform for open debate, discussion, and the exchange of ideas. Moreover with the wide adoption of digital technologies, such as social media, access to public discourse has been democratized. Our contemporary society has an abundance of debates that range from policy-making (e.g. how to stop climate change, the public healthcare system, abortion laws, nation trade agreements), to regulation (e.g. gun control, net neutrality, environmental regulations, immigration), up to philosophical dilemmas or questions (e.g. is AI an existential threat to humanity, should race be a factor for admissions, free speech vs censorship).

Debate-support technologies and tools are widely accessible - in the case of social media used by a substantial portion of the global population. Development teams that design, and support such tools take extra care to refine and polish usability functionalities. Typically, the focus in the HCI literature is on the design and use of interfaces and interactive systems, but it has only been recently that larger community interests are taken into account in the design of such technological infrastructure [26]. New paradigms are starting to infuse the design of such systems that are more sensitive to their larger societal context of use. Examples include community citizen science [26]; community informatics [11] and collective intelligence for the common good [47].

To organize the often rather fragmented socio-technical design suggestions coming from these fields, in this paper we propose a 4-layer socio-technical evaluation framework for online deliberation platforms that takes into consideration (from technical features to societal impact): i. usability, ii. discussion quality, iii. debate quality and iv. societal context. We examine the four layers through the prism of participants' sensemaking. We then describe an experiment with a new online discussion platform (BCause) that balances evenly between the formality of argumentative discussion and the naturalistic flavor of discussion while enhanced with computational aids for sensemaking support. To show that just building such a sophisticated platform is not enough to catalyze an effective public interest debate, we look at the application of BCause in two distinct use cases. Through the lens of the evaluation framework, we present a narrative of challenges and difficulties encountered. We hope this conceptual framework and the tale of our struggles inspires others to think deeply about how we can make our platforms not just more individually usable but also more community-useful in terms of public interest impacts.

## 2 RELATED WORK

Group deliberation happens online by means of online discussion platforms [18]. For example, studies have shown that people are able to come to well-reasoned and considered decisions through online discussions, even when they are not physically together [20, 23, 42]. Additionally, online discussion platforms can allow for larger and more diverse groups of people to participate in the deliberation process, which can help to ensure that a wider range of perspectives is taken into account [59]. Overall, the use of online discussion platforms for group deliberation is a promising approach that is increasingly being used in a variety of settings.

However, this form of dialogic communication also presents several generic issues that may impact the quality of the debate. Some of these issues include, but are not limited to:

- It is difficult to ensure that all participants have an equal opportunity to contribute to the discussion (e.g. Wikipedia edits [50]). This can be especially challenging in large group discussions, where some voices may be drowned out by others [51]
- Sometimes heated or unproductive discussions take place, with participants getting sidetracked or engaging in personal attacks. This can make it difficult for the group to reach a consensus or make a well-reasoned decision [38].
- Online discussions can also be subject to manipulation or bias, as it can be difficult to verify the identity of participants or ensure that they are acting in good faith [7, 25].

There are many root causes that influence the quality of societal debates - and their support, an exhaustive list we cannot produce here. However, three fundamental factors complicate discussions in such debates beyond just direct discussant engagement: different, often conflicting stakeholder interests (Polarization, Toxicity), debate content (Shallowness) and the complexity of meaning negotiation (Sensemaking) and artefact production (Collaboration).

- *Polarization*: participants become more entrenched in their positions and less willing to consider other perspectives. This can happen for a variety of reasons, such as the tendency of

people to seek out information that confirms their existing beliefs, or the fact that online discussions can sometimes become heated or adversarial [9, 49]. This leads to more division rather than coming to a consensus. This is usually abated by establishing ground rules (or a protocol of interaction), heavy moderation with users with elevated rights and explicit roles to ensure that discussion remains civil and productive, or encouraging participants to consider different perspectives [53].

- *Toxicity*: Much online behavior is unproductive, harmful, or counter to the goals of the discussion. This includes personal attacks, trolling, or other behavior that is intended to disrupt the conversation or make it difficult for the group to reach a consensus [57]. Toxicity is a major problem in online discussions, as it can make it difficult for participants to have a productive conversation and can even drive some people away from the discussion entirely.
- *Shallow content*: In some cases, online discussions may be quite deep and consist of well-reasoned, thought-provoking content [21, 24]. In the majority of the cases though, the content may be more shallow and consist of superficial or unoriginal ideas [36]. Ultimately, the depth of an online discussion will depend on the quality of the participants and the effort they put into contributing to the conversation [2].
- *Sensemaking* in online discussion can prove problematic [22]. As large discussions can be chaotic or disorganized, it is difficult for participants to follow the conversation or understand what is being discussed [3]. Additionally, not only the discussion itself but also the vast amount of participants, makes it challenging for individuals to keep track of all the different ideas being discussed and their provenance (who tells what).
- *Collaboration*: While the promise of online discussion is a highly collaborative environment where participants are working together effectively to generate new ideas, share information, and make decisions, in reality collaboration is usually less efficient [30], with participants struggling to effectively communicate and work together [17]. The quality of collaboration in online discussions is affected (among other factors) by the level of trust among participants, the clarity of the discussion goals, the diversity of perspectives represented, and the presence of effective moderators or facilitators.

Incorporating structure into discussions (a rudimentary implementation of argumentation theoretical tool) can be employed to tackle barriers that separate lay stakeholders from policy debates [12, 43]. Structuring discussions around arguments, i.e. claims with premises, or ideas and their supporting evidence or any sort of combination of argumentation components, where participants present their ideas or positions along with corresponding for and against arguments (e.g. IBIS [35]) can help to address some of the problems that can arise in online discussions. Although structuring discussion in arguments is not a silver bullet, argumentative discussion indeed allows for participants to better comprehend the essence of the debate [55, 58]. By explicitly stating their positions and the reasons for them, participants can help to make their ideas clearer and more understandable to others as different viewpoints

are being presented “by design”. Additionally, by considering and responding to counterarguments, participants can help to ensure that the discussion is grounded in evidence and reasoning, rather than just personal opinions or beliefs. This can help to prevent discussions from becoming unproductive or toxic, and can help participants to come to well-reasoned and considered decisions. However argumentative discussion is not mainstream because of the high cost of (complex) use and its effect on engagement -which is the currency used in online platforms [44]. Whereas argumentative discussion systems help participants to make their ideas more comprehensible to others, it is questionable how cognitively easier is for participants to follow the discussion. For example, structuring discussions around arguments could make the discussion more formal and academic in nature, which may not be suitable for all topics or groups. This approach may be particularly well-suited for discussions involving complex or technical issues, where participants need to carefully consider and evaluate different ideas and evidence. However, it may not be as effective for more casual or informal discussions, where a more relaxed and open-ended approach may be more appropriate. Additionally, some participants may find this approach intimidating or off-putting, as it can require them to carefully articulate their ideas and defend them against counterarguments, leading to reduced engagement. Summing up, an argumentative structure may not be as effective at fostering deep listening and empathy, it does however, facilitate more nuanced complex discussions by introducing discussion structure that helps to alleviate the shallowness barrier.

## 2.1 A landscape of debate-support tools

Over the past 30 years (since the advent of the World Wide Web), information and communication technologies (ICT) have been integrated in public consultation initiatives. We review here those software technologies that support online deliberation. Many technologies were not designed with this purpose, e.g. blogs, forums, message boards, however, this technological review aims to capture purposed or not tools that have been used extensively in public participation activities.

For clarity reasons, we define the concept of deliberation as a means of public consultation, which is the process in which participants engage in a reasoned opinion expression about an issue in an attempt to identify solutions about a stated problem and evaluate these suggested solutions [45]. Bachtiger and Parkinson extend this definition to its democratic aspect[5]: “Democratic deliberation is about using that mode in an inclusive and equal manner, oriented towards an effective, collective decision point and on into implementation”. Deliberation on a given issue of a community spans and progresses through a number of phases [54]. Initial phases correspond to *ideation* and *consolidation* where ideas are proposed, discussed, edited and evaluated. Later phases correspond to a *reconciliation* phase where proposals are aggregated and iteratively reevaluated and finally the *selection* phase - where a winning proposal is selected for implementation.

Referring to debate-support technologies, tools enormously vary by intent and purpose. It may include (among others):

- Online meeting platforms (Skype, Google Meet, Zoom, etc)

- Discussion forums (Reddit, Slack, Discord, simple phpBB forum): participants engage in asynchronous discussions
- Social media platforms (Twitter, Facebook) can be used to facilitate public interest debates by allowing participants to share their thoughts and engage in discussions online (massive scale - broad scope)
- Collaborative document-editing tools: (e.g. Google Docs) participants collaborate on documents in real-time
- Debate-specific platforms (Kialo, consider.it, debate.hub , bCause)

Existing solutions for public consultation and online deliberation can be arranged in the following three categories (as proposed by [32]), according to the anchoring concept of the participant contribution:

- *Time-Centric Systems*: Content is organised on a temporal basis (*when* it was contributed). Typical examples are email and chat rooms where usually posts appear in timely order, most recent first (or opposite). In general, time-centric systems thrive when it comes to the scale of participants but lack efficiency for public consultation purposes due to scattering of information (as evidenced in [4]).
- *Question-Centric Systems*: Contributions aim to answer a central question, the most representative example of such system are *Question-Answering* systems, e.g. stackoverflow.com. They usually focus on one domain and thrive in answering questions that are easily verified for correctness, e.g. *what is* type of questions. However, they have weak mechanisms to show the rationale or narrative of the responder. Often answers contain duplicate arguments (pieces of information that have been mentioned in other answers) and do not promote collaboration on the level of each answer but rather are usually flooded by many shallow and overlapping comments (shown at [37]).
- *Issue-Centric Systems*: Participants interact by not only providing their ideas, comments and therefore arguments but also explicitly linking those, creating deliberation argumentation maps [31]. Such augmentation in the deliberation process enables more systematic and structured discussion leading to healthier participation [14] and harnessing collective opinion and intelligence [8]. Also, the provision of evidence in arguments -evidence based reasoning- is directly linked to better decision making [19]. Also, they help to build shared understanding of the discourse [13] which is essential when tackling wicked problems. Key weaknesses of such systems revolve around the inherent complexity of the user interfaces and argumentation technologies for conversation, however, e.g. vagueness of concepts ideally requiring definition in advance of the argumentation scheme in place and significant training efforts in appropriately using the argumentation diagramming tools.

## 3 FROM USABILITY VIA USEFULNESS TO PUBLIC INTEREST

While traditionally the emphasis in the design of deliberation platforms or social media has been on usability (feature sets, user interfaces, and accessibility needs for enabling the highest levels of

engagement), little attention has been given to the overall goals of being *useful*, especially when moving from individual user to community and even societal needs. Increasingly, though, dedicated and advanced deliberation platforms like *consider.it* and *kialo.com* do pay more attention to the larger purpose of their discussions, for example by finding visual ways to scale arguments. Still, the design choices in moving these systems from not just being useful but also usable are often left implicit and still seem to focus on technical rather than larger socio-technical considerations. We argue that even if a public interest debate platform implements all the rich set of technical features that satisfy the usual HCI requirements, but fails to promote higher order public interest values such as participation, equality, diversity, common good etc. (the aforementioned higher-order “societal principles”) then is not useful. As a case in point: the Truth Social platform<sup>1</sup> conveys a perfectly usable platform following all the latest UI and UX principles. Though advertised as by its promoters a solution to “Big Tech censorship”, its “usefulness” from a public interest, common good and democracy point of view is highly debatable, as for example there is strong evidence that is a host to extremist voices.

So far we have introduced characteristics and examples of state of the art online debate support tools. The focus of our paper is not on these tools themselves, but rather on the collective and societal contexts of use in which they are to be applied. Although often strong claims are made about the impact these tools could have on public interest debates, most of the literature just focuses on the technical usability aspects and individual usability evaluations. However, how to shift the focus of these evaluations to actual societal impacts and the consequences these results may have in terms of socio-technical (re)designs of the debate tools and the ways in which they are (actually) most effectively used?

In the remainder of this paper we first outline a conceptual framework that helps us in transitioning from individual usability to public interest tool evaluation. We then use the framework to organize our observations around the usefulness of online debate tools in real world public interest settings. To this purpose, we examine the use of the state of the art Bcause debate tool in two cases around the same hot topic: climate change & heritage.

## 4 EVALUATION FRAMEWORK

In this section, we present a conceptual framework for the evaluation of public interest debate support tools. The examples presented in the following description of our framework parts are taken from the general literature. They are not an exhaustive and prescriptive list of aspects to (only) be taken into account, the reader may add their own classifications and examples. They are meant to help orientate a qualitative, generative and exploratory, yet coherent analysis of socio-technical aspects that can help make platform-supported societal debates more useful and impactful. In the construction of the framework levels and dimensions we build on the related social context model for discussion process analysis [15]. Our approach was also influenced by the approach employed by [27], which highlights the importance of looking at both “participation” and “reification” when designing and evaluating the mediation capabilities of online discussion technologies. Indeed, building on

Wenger’s theory on communities of practice [56], this work points to the importance of considering “Participation” as a more holistic process of ongoing collective action (combining participants’ interaction, ideas, information, goals, and values sharing) and “Reification”, as a process by which an abstract concept is objectified into something that can be reflected, discussed and negotiated upon from the group in an explicit way. Using this theoretical underpinning and also inspired by related conversation structuring frameworks like the Business Model Canvas, which helps organizations conduct structured, tangible, and strategic conversations around new businesses or existing ones [40], the research team distilled a novel evaluation framework, which captures crucial dimensions to assess public interest debate support tools. This consists of four levels, the rationale for which is described and motivated below.

### 4.1 Usability (individual contributions)

Beginning at the lowest level, the framework examines how well the platform features support individual discussion tasks. If we employ a chemistry analogy, this will constitute the debate **atoms**. The usability framework is derived from well-adapted HCI frameworks such as SUS (System Usability Scale) or Microsoft Desirability Toolkit:

**4.1.1 SUS (System Usability Scale).** [10] is a questionnaire-based tool used to measure the usability of a software product or system. It was developed by John Brooke in 1986 as a quick and dirty solution, yet it has been widely used in the industry and academia since then. The SUS consists of a 10-item questionnaire that asks participants to rate the usability of a system on a five-point scale. The SUS is a reliable and valid tool that can provide valuable feedback on the usability of a system, and it can be used to compare the usability of different systems.

**4.1.2 Microsoft Desirability Toolkit.** is a set of tools and techniques used to measure the desirability of a product or system. It was developed by Microsoft and consists of various methods, such as emotional response surveys, preference tests, and attribute ratings. The toolkit is designed to help designers and researchers understand how users perceive and experience a product, and it can provide insights into how to improve the product’s desirability (e.g. [6]). The Microsoft Desirability Toolkit is often used in combination with other usability testing methods to provide a more comprehensive understanding of a product’s usability and desirability.

Collectively, both toolkits can measure typical HCI quality aspects such as the ease of use, attractiveness (aesthetics), complexity, functionality and desirability of the product tested.

### 4.2 Discussion quality (individual sensemaking)

Moving on to a higher level, the framework examines how users make contributions to a particular discussion. In the previous chemistry analogy, this level constitutes the debate **molecules**. This employs close examination of platform features that aid in participants’ individual sensemaking, such as:

- **Visual aids:** visual analytic approaches provide condensed information which may improve human sensemaking performance in certain tasks (e.g. [39]).

<sup>1</sup><https://truthsocial.com/>

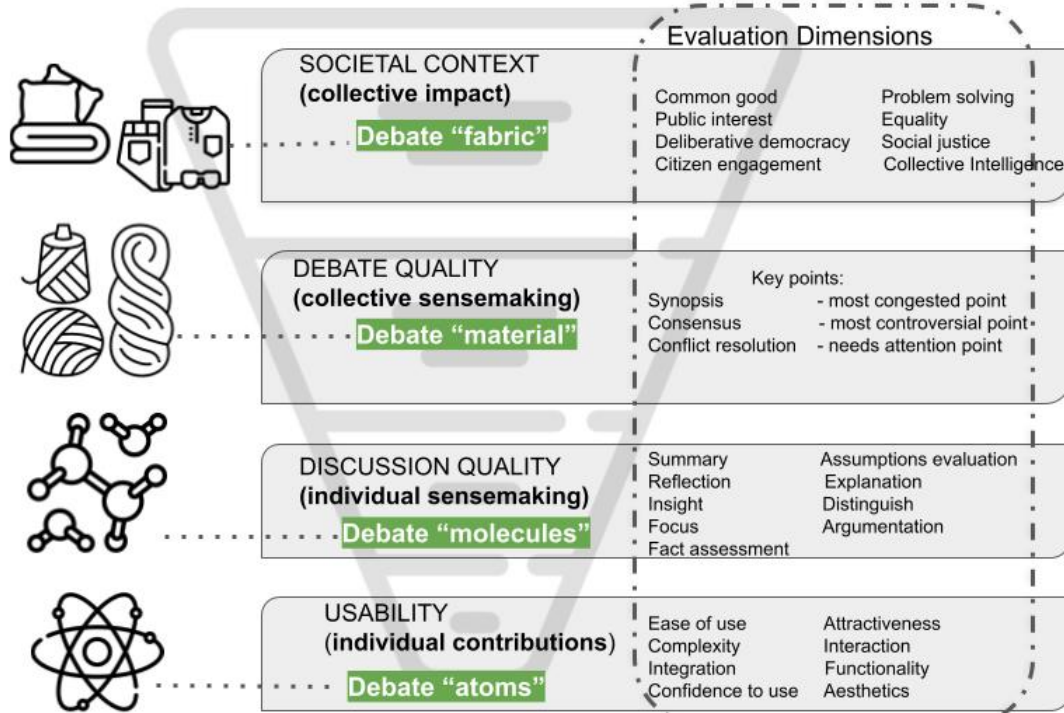


Figure 1: Evaluation framework 4 layers - debate "atoms", "molecules", "material" and "fabric"

- *Summarizing*: the affordance of quick reports provides a "quick glance" to users [3]; offering a manageable amount of information while guiding them to the points of interest. This has a positive effect in the cognitive cost of sensemaking [46].
- *Feedback* loops are crucial to improving users' sensemaking because they allow individuals to receive information about the consequences of their actions and make informed decisions based on this information.
- *Reflection* mechanisms allow users to pause, assess the current state of their actions, and proceed to make informed decisions. It helps individuals to develop a more nuanced understanding of a complex issue.
- *Structural organisation*: a system should accommodate a balance between complex structure of information and usability (ease of use), as incorporating complex argumentative formats though powerful can prove dangerous to the engagement of the platform [41].

### 4.3 Debate quality (collective sensemaking)

Further examining aggregate content quality, we move to a higher level of sensemaking by examining the overall debate quality and collective sensemaking. In our matter analogy, this could be considered the debate **materials**: useful components that are necessary conditions for creating products that serve actual needs. It evaluates whether the arguments presented arrive through a careful reasoning process to a conclusion that serves the - intended - goals

of the debate. For that, the framework focuses on the evaluation of the following platform features:

- *Evidence-based arguments*: are essential as they provide a systematic and objective approach to evaluating claims and assertions. Furthermore, they help to build consensus-building and cooperation in debates [33].
- *Synopsis* provides a comprehensive overview of the discussion (while summary typically provides only a condensed -short- form of a long text). It provides a common understanding of the key elements of a debate or what is the overall theme of the discussion
- *Key-points* By providing the key points, participants can distill their arguments compared to the other contributed and help to their reasoning and reflection. It should make sure that it includes all the different perspectives (re)presented in the discussion.

### 4.4 Societal context (collective impact)

As the top layer of the evaluation framework, we define the external societal context of the debate, i.e. what are the political goals of the debate and how well do the debate process and results fit those public interests. Continuing the matter-analogy: this could be seen as the "debate fabric" level, understanding how well the strands of are woven together into a cloth that actually meets the needs of the debate community and the societal stakeholders they represent. Essentially the framework evaluates the collective problem-solving abilities of a group or community to address common challenges

and achieve common goals in a similar fashion that Collective Intelligence for the Common Good (CI4CG) [47] is doing. For that it evaluates the following dimensions:

- *Problem-solving effectiveness* the platform should effectively engage stakeholders in a structured discussion and collaboration process towards the resolution of a specific issue. In the end, some actionable common ground should be identified along with potential areas of compromise
- *Diversity of perspectives* provides a thorough exploration of an issue and fosters a rich and nuanced discussion that prevents echo chambers and promotes inclusivity.
- *Collaboration and teamwork* are critical elements of high-quality debates. Working together in a coordinated and constructive manner enhances the quality of arguments and develops the foundation for long-term sustainable solutions.
- *Transparency and accountability* is promoted by open and honest discussions. Beyond legal reasons, holding participants responsible for their actions and decisions is essential for building public trust and the success of the deliberation process.
- *Inclusivity and equity*: is critical for successful debate, as they increase the legitimacy of the discussion, promote social cohesion and understanding, and help to promote social justice.

Note that we have only sketched the four layers of our evaluation framework and given some typical examples of socio-technical constructs and ways to evaluate them in each layer. By no means is this intended to be comprehensive. Rather, we intend it to act as a conceptual lens, an inspirational and organizing framework to broaden and focus public interest debate tool evaluations. In this way, it can spark ideas on the types of quality dimensions and aspects to take into account and questions to ask to make such tool support more effective. We now show how it helped us make "meta-sense" of the messy reality of real world computer-mediated public interest debates.

## 5 ONLINE DEBATING TOOLS IN THE WILD: USING BCAUSE IN THE CLIMATE CHANGE & HERITAGE CASES

We introduce the Bcause online deliberation tool and then briefly describe two real-world use cases on the same topic (climate change & heritage) we applied it to.

### 5.1 Bcause platform

The BCause "Reasoning for Change" platform is a structured and decentralised online discussion system for distributed decision making. The platform is being developed at The Open University (UK) with the goal of providing structured online discussions for groups to make decisions that are consulted, reflected and critically assessed by all discussion participants. It aims to overcome three fundamental limitations of discussion systems when applied to decision-making contexts: i. the lack of overall quality of discussion, particularly in terms of data structure and evidence-based reasoning; ii. the lack of functions that support sensemaking and situational awareness to enable people to participate meaningfully

in discussion; and iii. the lack of data ethics, in terms of data centralization, which implies that organizations must "barter" their data rights with outside companies in order to gain access to discussion technologies.

With BCause, we have adopted an approach that combines three main innovations: i. low-cost argument structuring: with an accessible user interface for users to contribute and analyze arguments in an online discussion process; ii. the distinctive use of automated discourse analysis and advanced visualizations with visual analysis and automated report/summary features to support sensemaking by discussion participants; iii. decentralization: with a data infrastructure that enables secure decentralization of discussion data and user identity, and gives users autonomy of choice and full control over the ownership of personal data. The core functionality of BCause is organising the discussion in positions and pro/con arguments (Fig. 2).

In addition to this core discussion functionality, some auxiliary computational debate support technologies are being developed for BCause. We outline their essence here but their technical details are beyond the scope of this paper and will be published in upcoming work.

*Dynamic summary of online discussions*: Taking advantage of previous results on the comparison of automatic summaries [3], we developed a summary of long texts based on the requirements of a large language model. This type of summary is more suitable for the specific scenario of an extremely long online discussion, as it accurately captures the essence of the discussion in a natural, human-understandable form. Several prompts and hyper-parameters were evaluated by human annotators for accuracy, factuality, and appropriateness. The summary is displayed in the left sidebar next to the debate and is also present in the debate preview, so that newcomers can get an idea of the status and progress of an online debate before they start contributing. It offers a synoptical overview of the state of the discussion, presenting the overall theme and main arguments of the discussion in a human-natural narrative style. Moreover, in the synopsis, key points of the debate are algorithmically identified and shown. Specifically, we automatically extract the most contested point, the most opposed point, and the point that requires the most attention from the participants based on the contingency of the alignment with analogous arguments or limited involvement. They are also shown in the left column (see Figure 3) and aim to nudge (navigate) the user to points of interest to contribute effectively to the discussion.

*Improving the quality of online debate by recommending arguments (evidence) extracted from the scientific literature*: Previous research has shown that online discussions often rely on poorly researched content and evidence of unreliable quality [29]. Therefore, we developed a recommendation system that aims to improve the overall quality of debate by retrieving data only from the scientific literature. We created a scientific argumentation recommendation system integrated with the Bcause platform that, given (i) a user's position on a given topic, (ii) his or her previous interactions with the platform, and (iii) the inferred position on the debated topic as input, proposes relevant statements or evidence as text fragments extracted from related scientific literature. This RecSys system can be evaluated for the key dimensions of relevance, argumentation,

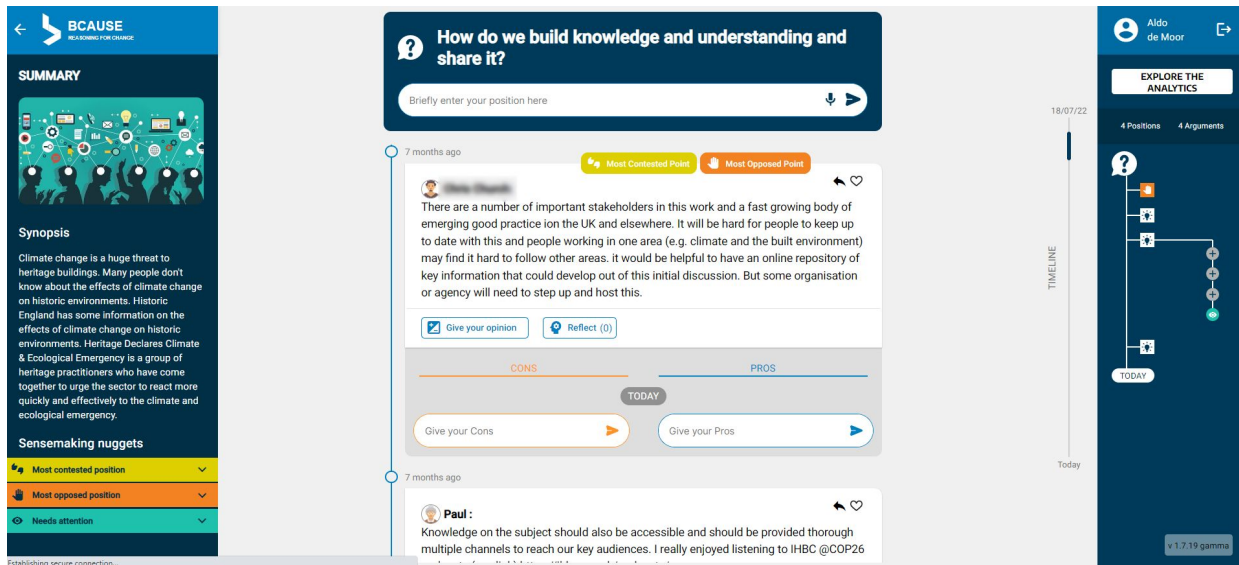


Figure 2: The core argumentation structure used in BCause; organising discussion in positions and pro/con arguments

and polarity. The system allows user to attach arguments sourced from scientific literature to a given position or argument.

## 5.2 The Use Cases

Considerable emphasis was placed on identifying a genuine need-use scenario. For example, in the first use case, we engaged Oxford Civic Society from the inception in the case planning and let them invite relevant stakeholders from their trusted network. Moreover, a substantial portion of the initial in-person workshop was devoted to the participants jointly defining the core debate questions, thus promoting community ownership and engagement. This section presents the application of the above evaluation framework to the use novel debate platform BCause.app<sup>2</sup> in two related use cases:

*Oxford Heritage Built Environment and Climate Change-debate:* This use case was initiated through contact with the Oxford Civic Society (OCS) who were interested in the issues surrounding heritage-built environments, climate change mitigation and net-zero targets. This is a key issue in a city with a rich legacy of conservation areas and heritage listed buildings. The potential to engage a wide range of stakeholders in a discussion of these issues in an online platform interested Oxford Civic Society as a way of bringing together diverse voices from across the heritage, building and environmental sectors. An elaborate process was designed to try and ensure the best possible fit of the BCause platform to the “debate fabric” needed by the OXFord stakeholders (see Appendix 1 for the use case process design). The consultation process started by bringing together an initial group of pathfinding stakeholders in a physical 2-hour workshop. First, the platform was introduced and the use case framework was co-designed, defining the discussion challenge, the discussion topics and potential process goals and outputs, Twelve participants attended and a set of four key topics to frame the discussion was agreed upon by the stakeholders. After this, an

alternating online and physical process was to take place of two more physical workshops (the second one on consolidating the debate and the third workshop on jointly defining outcomes and outputs) and in between two three-week BCause-mediated online debates. Although the online debate was off to a good start, the process came to an unforeseen halt. In all 4 discussions, with 7 positions and 9 arguments in total, and with 7 unique participants took place. In the next section, we examine some of the suggested reasons for the debate halting, as well as potential socio-technical remedies.

*The online Edge “Heritage & Net Zero” Debate:* An online workshop was undertaken with The Edge Debate, a multi-disciplinary built-environment think tank. It had already planned an online debate on “Heritage & Net Zero: A wicked problem?” gathering 135 participants, all professionals related to the Edge network. The Edge debate organizers had been contacted by one of the Oxford case organizers to form a second use case on this theme, which they considered a useful addition to their process. In the workshop, the short (5-minute) presentations by experts and stakeholders were each followed by a lively question-and-answer session. Instead of having an asynchronous process in which stakeholders themselves were to use the BCause platform, it was now used live by qualified mind mappers to interpret and record the online debate exchanges. This resulted in the creation of an extensive debate map consisting of 36 positions and 68 arguments.

## 6 A TALE OF STRUGGLES: USE CASE INSIGHTS

To illustrate the challenges involved in utilising a novel online deliberation platform in a real world community context, we present a narrative of struggles and obstacles encountered along the way. We use the evaluation framework outlined earlier as a resource to inspire and organize our evaluation thoughts. We focus our tale on

<sup>2</sup><https://bcause.app/home>

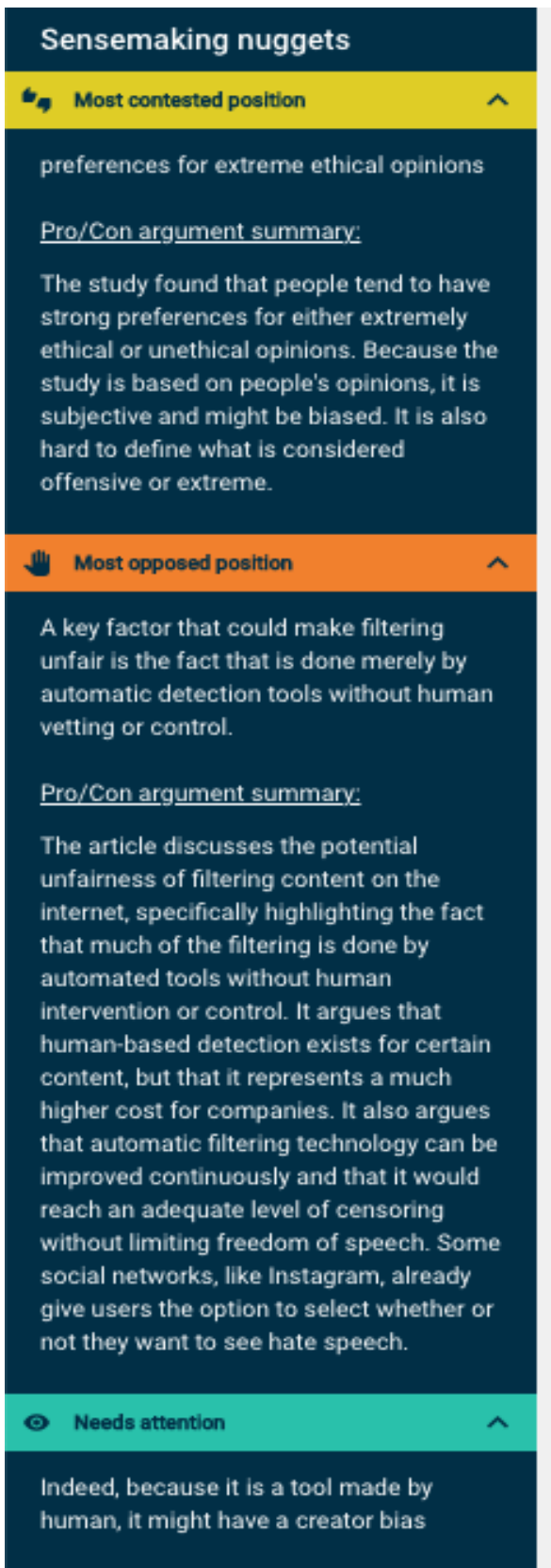


Figure 3: Sensemaking “nuggets” detailed view

the first use case, as that was the one most socio-technical efforts were put into. We use the Edge debate use case mostly to contrast and further refine some of the observations related to the first use case. We started our use cases with having a platform that was tested and found to be usable in many prior (lab) experiments. The use case itself was on a theme (climate change) that is very relevant and around which there is a huge general public engagement. In the Oxford case, we had a willing process owner who is a spider in a large local stakeholder web. We invested much effort and many resources in preparing the consultation process. And still, the debate did not take off.

In this section, we present some observations loosely organized according to the four layers of our evaluation framework. It is an exploratory analysis with only very tentative findings. We focus in particular on the higher layers of the framework, since this is where the “real world” context comes in.

The starting point of the observations was a questionnaire design focusing on the first three layers of the evaluation framework. It was used for generating questions in an in-depth interview by one of the team members with two of the most engaged users. Additionally, team members took notes of socio-technical issues surfaced in informal conversations with users throughout the process. Furthermore, observations of team members about usability and usefulness of the tool in the two use cases throughout the process were collected and jointly analyzed.

Although our evaluation process was rather limited, for an initial, qualitative exploration of framework layers and topics this sufficed.

## 6.1 Usability (individual contributions)

The initial engagement with Bcause was carried out without major technical usability issues. Users - having received training at the workshop - reported it to be a generally intuitive environment when they started working with the platform.

Despite the platform being generally usable in principle, we observed a cold start phenomenon which can be attributed to the small size of network involved. This remains a typical issue to newly founded platform projects. This phase was characterised by hesitations by the users to initiate a discussion. Another suggested reason was that it took quite a while to set up and configure the online debate after the pathfinder meeting, resulting in momentum lost.

To overcome this slow start, we had to employ a strategy of *nudging* participants to kickstart the discussion. Nudging involved reminder emails and notifications of new discussion activity that required the users’ attention.

## 6.2 Discussion quality (individual sensemaking)

Once the initial hesitation had been overcome, some promising discussions with several turn-takings indeed took place. However, again due to the small community size and the asynchronous nature of the discussion, there was a significant period of time without new entries, leading to stalled discussions that required manual intervention to resume them. This user feedback illustrates it:

“I had no problem using the platform – it was very intuitive and easy to use. However after making my



initial contribution very few people joined in the conversation – so it stalled and so I gave up in the end.”

One related missing (usability) feature that was reported repeatedly by users and might have overcome at least some of these interruptions were *notifications*. E-mails or other prompts with clickable links directly to an argument added in response to a user’s position or previous argument might have prompted more engaged and critical debate.

### 6.3 Debate quality (collective sensemaking)

Successful collective sensemaking took place in the initial pathfinder workshop, where the discussion challenge, the discussion topics and the potential process goals and outputs were jointly defined. Such collective sensemaking on sensitive political topics is usually much harder to do online. The importance of this physical meeting - and the stakeholder commitment that was generated is exemplified by this participant quote:

“My experience of the Pathfinder meeting was very positive, there was a good cross section of stakeholders present and people were really engaged. The key questions for the platform that emerged from the conversation were really relevant and it was good practice that they came from the stakeholders “bottom up” rather than “top down” approach. Really useful to start off with a F2F meeting with everyone in the room before moving the conversation online. People were very engaged and there was a good range of stakeholders.”

In the Oxford Case in particular, due to the limitations of the COVID pandemic and stakeholder commitments, it was not possible to hold another physical discussion as the original plan intended – instead participants provided feedback via a Zoom call in the second online workshop only. Based on the success of the first pathfinder workshop, and if we had had timely platform configuration and notification functionalities in place, the participants of the feedback meeting concluded that we could have satisfied some important necessary conditions to get a sustained, collaborative public interest debate underway.

In the second use case, we noticed that using Bcause to *live map* a live debate really requires well trained mind mappers and expert users to distill the essence of the natural discussion and simultaneously enter it in the online platform - where a subsequent deliberation by debate participants could then take place. Even with multiple expert mappers, however, it was difficult to keep up with the pace of the debate. Of course, discussions could be mapped post-event, but then a lot of the powerful symbolism of visualizing the debate as it happens - including the collective sensemaking that can happen by debate participants as they are triggered by particular arguments being added, is lost. It might be interesting to explore if aiding the expert mappers with speech recognition / advanced AI chatbots like ChatGPT could be of assistance here: with the experts providing prompts and indicating where to add arguments and the AI helping to then turning blurbs into argument-size meaningful chunks.

The interplay with the existing Bcause “sensemaking nuggets” such as automatic synopsis could enrich this process even more. Of

course, many other variations of such rich mixes of advanced sense-making technologies with human users in various debating roles could be envisioned. Such configurations could further improve collective sensemaking debate quality, including the socio-technical features they would require at the lower layers of the evaluation framework.

### 6.4 Societal context (collective impact)

A societal context factor that likely further increased the need for physical collective sensemaking meetings may have been the political nature of the debate. When stakeholders represent local organizations with vested and not always overlapping interests, having regular physical meetings may be even more of the essence to generate sufficient trust to engage sufficiently online afterwards and work toward contributions - individually and collectively - that do have a collective societal impact.

Related to this was the fact that the original process owner in the first use case due to personal circumstances was not able to lead the debate, engage stakeholders present more and attract new stakeholders so that a snowballing process of interest and investment could have been catalyzed. *Ownership by a community spider in the web* who can act as a facilitator and catalyst therefore seems to be another necessary condition for online public interest debates to take off.

A suggested interesting technical feature, very much originating from the societal context layer were *automated information reports* that can be used for accountability/transparency reasons to their stakeholders or constituents.

## 7 DISCUSSION AND CONCLUSIONS

The real-life testing of Bcause has confirmed the usability of individual features of the platform (with the lack of discussion notifications mentioned as a still missing main feature). However, it has also highlighted issues of maintaining engagement, especially in less structured and highly political interorganisational contexts such as in climate change & heritage use cases. Using the evaluation framework we tried to make “meta-sense” of what worked and what didn’t work in our use cases. Although the weight of the analysis was on the first one, we sketched the contours of the second use case as an illustration of the wide range and variety of socio-technical considerations to be taken into account in real world platform deployment.

In a mathematical analogy, we argue that while usability is a necessary attribute for usefulness, it is not sufficient on its own. Extending in the same way, the usefulness (of a platform) is necessary but not sufficient for achieving public interest. This kind of thinking is in line with earlier work on “The Pragmatic Web” which argued that Semantic Web technologies only become useful when applied in well understood communities of use [52].

In earlier work on Bcause, we have demonstrated the use of computational aids to improve the individual and collective sense-making of online debate participants. That include an *automatic summariser* that crafted a synopsis of the discussion enabling quick understanding of the state of the discussion and the *provision of external scientific arguments* that improve the accountability and trust of the system. We foresee that in the future, the advent of

highly-capable AI technology (such as chatGPT) will revolutionise user interaction, reflection and sensemaking in online debate platforms. We expect these tools to significantly reduce the cognitive load associated with participating in online discussion and overall enhance the efficiency and effectiveness of deliberation platforms, as long as AI-safety mechanisms are in place. The sky being the limit in the promises being made by the developers of these tools, however, we hope we have made the case that the human factor is as important as advanced technological features. Along the lines of Douglas Engelbart [16], we advocate that AI powered systems should be developed to *augment* human intelligence, rather than to replace it.

To increase user motivation, in the first use case we did identify a "community champion", the Oxford Civic Society, in the first use case. One could argue that to derisk the use case in terms of limited engagement, more champions should be identified. In practice, however, engaging champions is a very costly process in terms of time and energy spent. A hallmark of real world cases is exactly that resources are limited and one often has to make do with what is available. Another more political reason for not engaging multiple champions might be that the focus and quality of the debate could be lowered with different champions potentially have conflicting core interests. At any rate, we believe community champions to be a core necessary condition for making online-mediated societal debates work in practice. Obviously, who to engage in this role and how it should be played are still wide open research questions,

Although the uptake and output of the platform was somewhat disappointing, we think that it is still - or maybe especially - worth sharing our findings. Iteration and improvement are important attributes of design. Failure is part of that process and learning from failure is an important way to uncover key concepts and promote reflection [28]. Releasing sophisticated online deliberation systems in the unruly real world of stakeholder consultations on wicked problems entails a degree of socio-technical complexity that requires a lot of failure and learning to progress.

We hope our findings inspire others to present their own stories of failures and setbacks and what learnings those may have triggered in them. It is our strong belief that the hurdles presented must and can be overcome if we are to make our collective intelligence systems have a noticeable impact on the common good. The first step, though, is to become aware of and share those hurdles, and come up with initial socio-technical (re)designs of our systems and the contexts in which we use them.

## ACKNOWLEDGMENTS

Our sincere gratitude to Ian Green, Chair of Oxford Civic Society, Chris Church, Oxford Friends of the Earth and Oxford Civic Society, and Dr Kerry J Mashford OBE, The Edge Debate, for making the use cases possible and joining the struggle. This research was funded by the U.S. Office of Naval Research under award number N00014-19-1-2366

## REFERENCES

- [1] Mark Aakhus and Miriam Greenfeld Benovitz. 2008. Argument reconstruction and socio-technical facilitation of large scale argumentation. In *Proceedings of the 3rd International Conference on the Pragmatic Web: Innovating the Interactive Society*. 77–81.
- [2] Yair Amichai-Hamburger, Tali Gazit, Judit Bar-Ilan, Oren Perez, Noa Aharoni, Jenny Bronstein, and Talia Sarah Dyne. 2016. Psychological factors behind the lack of participation in online discussions. *Computers in Human Behavior* 55 (2016), 268–277.
- [3] Lucas Anastasiou and Anna De Liddo. 2021. Making Sense of Online Discussions: Can Automated Reports help?. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–7.
- [4] Pablo Aragón, Vicenç Gómez, and Andreas Kaltenbrunner. 2017. Detecting platform effects in online discussions. *Policy and Internet* 9, 4 (2017), 420–443.
- [5] André Bächtiger and John Parkinson. 2019. *Mapping and measuring deliberation: towards a new deliberative quality*. Oxford University Press.
- [6] Joey Benedek and Trish Miner. 2002. Measuring Desirability: New methods for evaluating desirability in a usability lab setting. *Proceedings of Usability Professionals Association* 2003, 8-12 (2002), 57.
- [7] Alessandro Bessi and Emilio Ferrara. 2016. Social bots distort the 2016 US Presidential election online discussion. *First monday* 21, 11-7 (2016).
- [8] Eric Bonabeau. 2009. Decisions 2.0: The power of collective intelligence. *MIT Sloan management review* 50, 2 (2009), 45.
- [9] Kenneth Boyd. 2020. Group epistemology and structural factors in online group polarization. *Episteme* (2020), 1–16.
- [10] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
- [11] John M Carroll, Patrick C Shih, and Jessica Kropczynski. 2015. Community informatics as innovation in sociotechnical infrastructures. *The Journal of Community Informatics* 11, 2 (2015).
- [12] Stephen Chaudoin, Jacob Shapiro, and Dustin Tingley. 2017. Revolutionizing teaching and research with a structured debate platform. *Journal of Political Science* 58 (2017), 1064–1082.
- [13] Jeff Conklin and Michael L Begeman. 1988. gIBIS: A hypertext tool for exploratory policy discussion. *ACM Transactions on Information Systems (TOIS)* 6, 4 (1988), 303–331.
- [14] Anna De Liddo, Ágnes Sándor, and Simon Buckingham Shum. 2012. Contested collective intelligence: Rationale, technologies, and a human-machine annotation study. *Computer Supported Cooperative Work (CSCW)* 21, 4-5 (2012), 417–448.
- [15] Aldo De Moor and Rolf Kleef. 2005. A social context model for discussion process analysis. In *Information systems for sustainable development*. IGI Global, 128–155.
- [16] Douglas C Engelbart. 1962. Augmenting human intellect: A conceptual framework. *Menlo Park, CA* 21 (1962).
- [17] Evren Eryilmaz, Nimer Alrushedat, Sumonta Kasemvilas, Justin Mary, and J Pol. 2009. The effect of anchoring online discussion on collaboration and cognitive load. *AMCIS 2009 Proceedings* (2009), 576.
- [18] Katharina Esau, Dennis Friess, and Christiane Eilders. 2017. Design matters! An empirical analysis of online deliberation on different news platforms. *Policy & Internet* 9, 3 (2017), 321–342.
- [19] J St BT Evans, David E Over, and Keith I Manktelow. 1993. Reasoning, decision making and rationality. *Cognition* 49, 1-2 (1993), 165–187.
- [20] Shelly Farnham, Harry R Chesley, Debbie E McGhee, Reena Kawal, and Jennifer Landau. 2000. Structured online interactions: improving the decision-making of small discussion groups. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. 299–308.
- [21] Patricia K Gilbert and Nada Dabbagh. 2005. How to structure online discussions for meaningful discourse: A case study. *British Journal of Educational Technology* 36, 1 (2005), 5–18.
- [22] Todd Graham and Tamara Witschge. 2003. In search of online deliberation: Towards a new method for examining the quality of online discussions. (2003).
- [23] Michel JJ Handgraaf, Kerry F Milch, Kirstin C Appelt, Philip Schuette, Nicole A Yoskowitz, and Elke U Weber. 2012. Web-conferencing as a viable method for group decision research. *Judgment and Decision making* 7, 5 (2012), 659–668.
- [24] Noriko Hara, Curtis Jay Bonk, and Charoula Angeli. 2000. Content analysis of online discussion in an applied educational psychology course. *Instructional science* 28 (2000), 115–152.
- [25] Susan C Herring. 2004. Computer-mediated discourse analysis: An approach to researching online behavior. *Designing for virtual communities in the service of learning* 338 (2004), 376.
- [26] Yen-Chia Hsu and Illah Nourbakhsh. 2020. When human-computer interaction meets community citizen science. *Commun. ACM* 63, 2 (2020), 31–34.
- [27] Luca Iandoli, Ivana Quinto, Anna De Liddo, and Simon Buckingham Shum. 2014. Socially augmented argumentation tools: Rationale, design and evaluation of a debate dashboard. *International Journal of Human-Computer Studies* 72, 3 (2014), 298–319.
- [28] Andrew Jackson, Allison Godwin, Scott Bartholomew, and Nathan Mentzer. 2022. Learning from failure: A systematized review. *International Journal of Technology and Design Education* 32, 3 (2022), 1853–1873.
- [29] Magnus E Jonsson and Joachim Åström. 2014. The challenges for online deliberation research: A literature review. *International Journal of E-Politics (IJEP)* 5, 1 (2014), 1–15.

- [30] Karen Kear. 2010. Collaboration via online discussion forums. *Online communication and collaboration* (2010), 30–33.
- [31] Paul A Kirschner, Simon J Buckingham-Shum, and Chad S Carr. 2012. *Visualizing argumentation: Software tools for collaborative and educational sense-making*. Springer Science Business Media.
- [32] Mark Klein. 2015. A Critical Review of Crowd-Scale Online Deliberation Technologies. Available at SSRN 2652888 (2015).
- [33] Mark Klein and Luca Iandoli. 2008. Supporting collaborative deliberation using a large-scale argumentation system.
- [34] Travis Kriplean. 2010. Building interpretable discussions: for effective public engagement. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*. 2971–2974.
- [35] Werner Kunz and Horst WJ Rittel. 1970. *Issues as elements of information systems*. Vol. 131. Citeseer.
- [36] Paula San Millan Maurino. 2007. Looking for critical thinking in online threaded discussions. *Journal of Educational Technology Systems* 35, 3 (2007), 241–260.
- [37] Elizabeth Murphy. 2004. Recognising and promoting collaboration in an online asynchronous discussion. *British Journal of Educational Technology* 35, 4 (2004), 421–431.
- [38] Manuela Neurauder-Kessels. 2013. *Impoliteness in cyberspace: Personally abusive reader responses in online news media*. Ph.D. Dissertation. University of Zurich.
- [39] Celestine A Ntuen, Eui H Park, and Kim Gwang-Myung. 2010. Designing an information visualization tool for sensemaking. *Intl. Journal of Human-Computer Interaction* 26, 2-3 (2010), 189–205.
- [40] Alexander Osterwalder and Yves Pigneur. 2010. *Business model generation: a handbook for visionaries, game changers, and challengers*. Vol. 1. John Wiley & Sons.
- [41] Fabio Paglieri and Cristiano Castelfranchi. 2010. Why argue? Towards a cost-benefit analysis of argumentation. *Argument & Computation* 1, 1 (2010), 71–91.
- [42] Trena M Paulus. 2009. Online but off-topic: Negotiating common ground in small learning groups. *Instructional Science* 37 (2009), 227–245.
- [43] Eugen Octav Popa, Vincent Blok, and Renate Wesselink. 2020. Discussion structures as tools for public deliberation. *Public Understanding of Science* 29, 1 (2020), 76–93.
- [44] Soumya Ray, Sung S Kim, and James G Morris. 2014. The central role of engagement in online communities. *Information Systems Research* 25, 3 (2014), 528–546.
- [45] Nancy Roberts. 1997. Public deliberation: An alternative approach to crafting policy and setting direction. *Public Administration Review* (1997), 124–132.
- [46] Daniel M Russell, Mark J Stefik, Peter Pirolli, and Stuart K Card. 1993. The cost structure of sensemaking. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*. 269–276.
- [47] Douglas Schuler, Anna De Liddo, Justin Smith, and Fiorella De Cindio. 2018. Collective intelligence for the common good: Cultivating the seeds for an intentional collaborative enterprise. , 13 pages.
- [48] Bryan Semaan, Heather Faucett, Scott P Robertson, Misa Maruyama, and Sara Douglas. 2015. Designing political deliberation environments to support interactions in the public sphere. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 3167–3176.
- [49] Ignacio-Jesús Serrano-Contreras, Javier García-Marín, and Óscar G Luengo. 2020. Measuring online political dialogue: does polarization trigger more deliberation? *Media and Communication* 8, 4 (2020), 63–72.
- [50] Aaron Shaw and Eszter Hargittai. 2018. The pipeline of online participation inequalities: The case of Wikipedia editing. *Journal of communication* 68, 1 (2018), 143–168.
- [51] Ruth Shortall, Anatol Itten, Michiel van der Meer, Pradeep K Murukannaiah, and Catholijn M Jonker. 2021. Inclusion, equality and bias in designing online mass deliberative platforms. *arXiv preprint arXiv:2107.12711* (2021).
- [52] Munindar P Singh. 2002. The pragmatic web. *IEEE Internet Computing* 6, 03 (2002), 4–5.
- [53] Kim Strandberg, Staffan Himmelroos, and Kimmo Grönlund. 2019. Do discussions in like-minded groups necessarily lead to more extreme opinions? Deliberative democracy and group polarization. *International Political Science Review* 40, 1 (2019), 41–57.
- [54] Cyril Velikanov and Alexander Prosser. 2017. Mass Online Deliberation in Participatory Policy-Making—Part I. In *Beyond Bureaucracy*. Springer, 209–234.
- [55] Yu-Mei Wang. 2019. Enhancing the quality of online discussion—Assessment matters. *Journal of Educational Technology Systems* 48, 1 (2019), 112–129.
- [56] Etienne Wenger. 1999. *Communities of practice: Learning, meaning, and identity*. Cambridge university press.
- [57] Yan Xia, Haiyi Zhu, Tun Lu, Peng Zhang, and Ning Gu. 2020. Exploring antecedents and consequences of toxicity in online discussions: A case study on reddit. *Proceedings of the ACM on Human-computer Interaction* 4, CSCW2 (2020), 1–23.
- [58] Amy X Zhang. 2017. Systems for improving online discussion. In *Adjunct Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. 111–114.
- [59] Weiyu Zhang, Xiaoxia Cao, and Minh Ngoc Tran. 2013. The structural features and the deliberative quality of online discussions. *Telematics and informatics* 30, 2 (2013), 74–86.

## 8 APPENDIX 1: THE OXFORD USE CASE ENGAGEMENT APPROACH: EXPLORING OPPORTUNITIES FOR EFFECTIVE ACTION TOGETHER

### 8.1 Getting started: the plan

**8.1.1 Initial engagement.** The key local stakeholder who was acting as the process owner was the Oxford Civic Society (OCS). Following an initial briefing meeting in which the scope of the project and communication pathways were agreed with OCS an introductory leaflet and invitation to participate were circulated to a list of potential participants through the OCS network via email. Included were representatives from heritage organisations, Oxford City Council, Oxford University and Oxford Friends of the Earth.

**8.1.2 BCAUSE engagement ethos.** The engagement ethos of BCAUSE is based on the principles of participatory research. A methodology that includes researchers and participants working together in dialogue to co-design the use case process, the discussion framework and tangible, actionable outcomes.

**8.1.3 The use case challenge.** Oxford has a rich and valuable legacy of historic buildings and conservation areas. As with other major cities nationally and globally Oxford faces the ‘zero-carbon’ challenge: how to adapt buildings and neighbourhoods so they can be more energy-efficient yet at the same time protect heritage assets that are a source of common heritage, civic pride and identity in the city.

The Oxford Use Case brings together a group of prestigious stakeholders involved in different aspects of heritage management and climate change to explore the conflicts of interests that exist between those seeking to protect the historic built environment and the need to mitigate the impact of climate change and carbon reduction. It seeks to find common ground with policymakers and create a tangible deliverable in the form of a set of pathfinder recommendations and actions going forward to engage the wider national consortium of heritage towns and cities.

Stakeholders are interested in participating saying that it appeals to them as a collaborative process that will bring together diverse sectoral interests to seek a way forward in the structured environment of the BCAUSE platform. Oxford Civic Society see this as an opportunity to become “trailblazers” in leading this initiative.

"The challenge is significant: there are perceived conflicts between carbon reduction and heritage protection. Planning regulations are complex and do not enable easy solutions to achieving the net-zero targets set out in the climate action plans of the City and County. This is an opportunity to bring interested parties together to create a way forward." *Ian Green, Director of Oxford Civic Society.*

**8.1.4 The use case objectives.**

- The real-life testing of the BCAUSE platform as a process to enable individuals and groups to better inform their decision making and consultation processes.

- To co-create solutions to complex problems by openly discussing them with others in a space designed for idea sharing and cooperation.
- To gain a mutual understanding of each other's points of view and encounter and compare different perspectives in a structured environment.
- To achieve an outcome that will be of value to participating stakeholders going forward.

## 8.2 The consultation process in action

*8.2.1 Start-Up.* OCS worked together with the BCAUSE team to bring together an initial group of pathfinding stakeholders, sending out invitations via the OCS network. Invitees were provided with a briefing note outlining the key concept behind the debate and the BCAUSE platform. With the initial stakeholder group assembled, the partially executed process plan was as follows.

*8.2.2 Workshop 1: The Pathfinder Workshop.* The first workshop introduced the BCAUSE platform and set the Use Case framework - defining the discussion challenge, the discussion topics and the potential process goals and outputs. Twelve participants attended the first in-person workshop and a set of key topics to frame the discussion were agreed by the stakeholder group.

- (1) What are the messages for decision-makers (government)?
- (2) How do we build knowledge and understanding and share it?
- (3) How do we communicate positively to encourage action and avoid disputes?
- (4) How do we tackle the skills gap?

The workshop included a practical demonstration of the platform and participants were also provided with a user guide. Following the workshop the OCS BCAUSE project, with some delay, went live with links to each of the four topics so that participants could choose topics of interest and a user guide.

*8.2.3 Empowered Online Discussion: Exploring the topics (3 weeks).* Using BCAUSE the stakeholder group were encouraged to collectively explore the topics chosen during the Pathfinder workshop and work towards topics, arguments - pros and cons and stakeholder positions. Initially, this process went well, but after a while it mostly came to a halt.

*8.2.4 Workshop 2: Consolidating the Debate: Sharing Feedback on Topics.* The second workshop was designed to bring participants back together to collectively discuss and reflect on the key topics that arose in the course of the discussions so far, add any additional questions and participants to invite, and identify emerging contours of process outcomes. Due to the limitations of the Covid pandemic and stakeholder commitments it was not possible to hold another physical discussion - instead participants provided feedback via a Zoom call.

*8.2.5 Empowered Online Discussion: Harvesting Ideas (3 weeks).* Participants were to work towards refining the discussion and preparing actions going forward to the final workshop. The on-line process that had already been fledgling before, now stopped completely.

*8.2.6 Workshop 3: Outcomes and Outputs.* The final workshop was planned as an interactive session in which participants were to choose the top ideas for impact to emerge from the debate. Participants then were to decide on the next steps and plan any dissemination outcomes such as publication, presenting findings to decision-makers or forming alliances to progress effective action. As the process had already died before, this planned workshop was not held. As a related real-world use case, the Edge debate use case was initiated.