Augmented Democratic Deliberation: Can Conversational Agents Boost Deliberation in Social Media?

Blue Sky Ideas Track

Rafik Hadfi
Kyoto University
Kyoto, Japan
rafik.hadfi@i.kyoto-u.ac.jp

Takayuki Ito
Kyoto University
Kyoto, Japan
ito@i.kyoto-u.ac.jp

ABSTRACT

Online social media are currently perceived as the new means of providing a platform for participation among citizens. This comes at a time when people are led to alternative spaces of political expression as traditional channels become strained. New technologies appear to have a transformational potential that could lead to social change and achieve deeper and wider mobilisation in political processes. However, people still face challenges when seeking information, disseminating information, or engaging in online deliberation. Allowing every citizen to participate in discussions and thus influence the final decision requires countless interactions that take considerable amounts of time and energy. This process is cognitively demanding due to linguistic barriers or when the problems on the table are multidisciplinary. We envision in this blue sky paper the development of autonomous and intelligent conversational agents that can augment the deliberative capacities of citizens in social media. To implement our vision, we start by proposing an approach that quantifies deliberation in online argumentative discussions. Then, we propose a methodology to optimise deliberation across discussion threads. The proposed concept is expected to pave the way to a form of augmented democratic deliberation built on the cooperation of humans and conversational agents.

KEYWORDS

Conversational Agents; Multiagent Systems; Collective Intelligence; Social Media; Online Discussions; Deliberation; Digital Democracy; Augmented Democracy; Natural Language Processing

ACM Reference Format:

Rafik Hadfi and Takayuki Ito. 2022. Augmented Democratic Deliberation: Can Conversational Agents Boost Deliberation in Social Media?: Blue Sky Ideas Track. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 5 pages.

1 INTRODUCTION

Online social media now have a transformational potential that can lead to social and political changes [2, 47]. Social networks with AI capabilities are poised to redefine how democratic processes are carried out [14]. Deliberation by the citizenry, in particular, plays a central role in modern democracy and is constantly changing under the influence of new communication technologies. Deliberation is an important process that promotes the involvement of larger parts of the population in decision-making processes. It is a communicative procedure that typically leads to collective decisions and promotes substantive, balanced, and civil discussion [34].

The deliberation process on large-scale platforms must overcome a number of challenges. Allowing every citizen to participate in discussions and thus influence the final decisions often results in countless interactions. Many citizens want to incrementally shape outcomes, while others want to propose their own solutions or criticize the solutions brought to the table by others. Under such constraints, it is less likely for everyone to have equal time to raise issues, develop ideas, receive feedback, or evaluate new options [32]. If debaters had the same degree of influence on the proposed solutions, then the deliberation process would likely take substantial amounts of time and energy. As an example, political discourses via live chat have significantly low deliberative quality [15]. The tediousness of the process itself often disincentivizes citizens to participate in it, which in turn renders the deliberation unfair. Moreover, deliberation processes become more difficult when the decision makers face complex multi-stakeholder, multi-disciplinary, or "wicked" problems [29, 30]. Other considerations, such as language barriers, may also pose problem for large-scale deliberations when participants try to communicate delicate decisions to a larger public [11]. In order to guarantee the scale (possibly thousands of citizens) and quality of deliberation (possibly different topics), it is crucial to understand how deliberation works on social networks [28]. If we could understand the mechanisms that underlie deliberation, we would be able to suggest effective ways to conduct democratic deliberation on social networks and even use conversational agents to help larger groups of citizens engage in public decisions. The vision of this paper is to augment the classical forms of deliberation using conversational multiagent systems (CMASs). Figure 1 illustrates the transition from the classical forms of deliberation to multiagent-powered deliberations. This transition is significant because it has the potential to advance the state of the art in deliberative democracy by algorithmically formulating deliberation using an objective index that could be applied to any deliberative process [14]. Once quantified, the proposed index of deliberation could be algorithmically integrated in conversational multiagent systems that interact with humans to optimise deliberative processes.

In this paper, we introduce an approach that starts by quantifying the notion of deliberation. Then, we propose an agent-based approach to augment deliberation in online discussions. This idea paves the way to a form of democratic deliberation that is built on the cooperation of humans and agents. In Section 2, we cover some of the related work on discussion platforms, their role in democratic

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

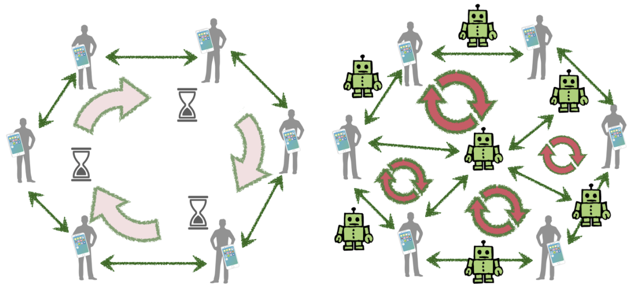

Figure 1: Classical forms of deliberation among humans are affected by cognitive limitations such as the inability to quickly read, process, and respond to ideas. Deliberative loops are either nonexistent or too weak to create effective deliberation. When intelligent conversional agents are introduced, deliberative processes could be amplified.

deliberation, and the potential use of conversational agents. In Section 3, we provide an operational definition of deliberation that can be easily implemented in argumentative conversational agents. In Section 4, we outline the overall methodology used to implement our vision. Finally, we summarize our concept and highlight the challenges it faces.

2 PROSPECTS FOR SOCIAL MEDIA

Social media are currently being used to integrate opinions and lead to improved social agreements [38, 39]. For example, the Collagree system was employed for idea gathering and city planning [25–27]. The CoLab system was used to collect the opinions of thousands of people worldwide to address global climate change [39]. The Deliberatorium is another platform where people submit ideas by following an argumentation map that frames the ideas within a given discussion structure [24]. Such platforms are also being used to empower citizens and help implement sustainable goals [43]. For instance, the D-Agree platform was employed to collect opinions on the implementation of Sustainable Development Goals in Afghanistan [20, 21].

Social media are also posed to strengthen the deliberative forms of democracy by generating a “positive supply-side shock to the amount of freedom in the world” [47]. We envision the use of artificial and autonomous conversational agents to augment the effect of social media on democratic processes. Empowering deliberative democracy by augmenting people’s cognitive abilities using intelligent conversational agents is a novel idea based on the recent progress in the areas of autonomous artificial agents and machine learning [22]. Digital platforms that seek to facilitate public deliberation over policy or regulations are not new [23, 29, 37, 44], but the challenges remain persistent. The main challenge is the conception of intelligent deliberative processes that allow decision makers to identify a wide range of options, assess their relevance, and develop epistemically responsible solutions [17]. For instance, the Cornell Regulation Room project attempted to address this challenge using a model of computationally assisted regulatory participation [36]. However, this type of approach relies heavily on human mediation to improve the quality of the comments, and thus it cannot serve as a general template for automatic deliberation processes. The approach advocated in this paper adopts a method based on conversational agents and a combination of deliberative [50] and facilitation mechanisms [18] to optimise human deliberation. Similar approaches have been proposed to increase cooperation among humans [46].

3 TOWARD AN OPERATIONAL DEFINITION OF DELIBERATION

Here, we depart from an operational definition of deliberation that could be later implemented within conversational agents on online platforms [18, 25]. To mathematically understand and algorithmically implement the notion of deliberation, we start by adopting an economic interpretation of deliberation [49] and its applicability to utilitarian agents in preferential domains. This notion can then be extended to linguistic domains where agents rely on natural language processing (NLP) to converse with humans in online debates. Our core idea is that the propagation of opinions in a discussion raises the possibility that some form of systematic behavior may emerge as a result of newly formed social relationships between the participants. To the extent that the opinions and interests of the participants are shared, groups may possess the ability to coordinate, in the sense that individual behaviors generate a form of rational behavior for the group. The presence of social influence in any type of network, coupled with the ability of individuals to modulate their opinions in response, renders it possible for individuals to engage in dialogue. Characterising deliberation is, therefore, a matter of accounting for influence cycles, or feedback loops, that enable individuals to engage in back-and-forth discussions and thereby modify their opinions. Accordingly, it is important to establish the necessary and sufficient conditions ensuring that cyclic and conditional opinions converge to stable opinions. We use this insight to quantify deliberation based on the existence of feedback loops in argumentative discussions.

4 METHODOLOGY

4.1 Representing Large Online Discussions

The first aim is to represent online discussions in a manner that allows the quantification of deliberative cycles. To this end, a promising candidate method is offered by the issue-based information system (IBIS), which is an argumentation-based formalism designed to model complex and ill-defined problems that involve multiple stakeholders [10, 31]. The formalism is based on issues, ideas, and arguments, as illustrated in Figure 2 (left). We extend this representation to allow time-dependent analysis, which is the paramount task in studying the evolution of deliberation processes in online debates. The problem with the standard IBIS model (Figure 2 (left)) is that it does not reflect the real temporal and spatial aspects of discussion data. The temporal aspect is characterised by the timestamps of the posts, while the spatial aspect refers to the author and location of the post. The target representation shown in Figure 2 (right) is aggregated as a temporal knowledge graph [35]. Such representation integrates and aggregates information into densely interlinked ideas, issues, arguments, timestamps, and user information. Such information is required for the quantification of deliberation and later the integration of agent actions within sub-threads of the discussion.

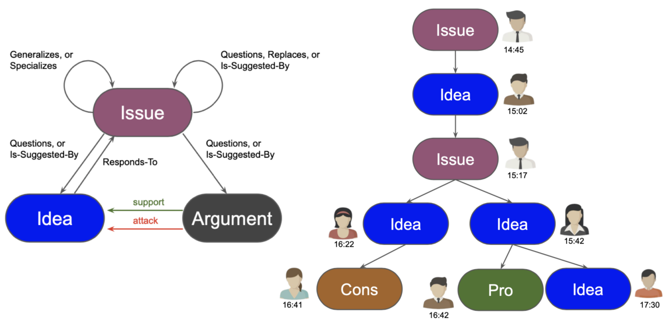

Figure 2: Issue-based information system (IBIS) and its temporal and user-centered representation

4.2 Quantifying Deliberation in Online Discussions

To quantify deliberation, we need to first account for the influence cycles that usually enable individuals to engage in back-and-forth discussions and thus modify their opinions. The aim here is to algorithmically identify such feedback loops in the temporal graphical representation mentioned in the previous section with respect to a topic model T that can vary across the threads of the discussion. Then, we measure how the textual content is propagated across the nodes that make up the discussion thread. For instance, in Figure 3, three debaters are involved in a loop relative to path P_1 concerning "climate" topics. Path P_2, on the other hand, does not constitute a relevant topical loop.

The presence of cyclic influences introduces the possibility of unstable behavior emerging as interminable discussions with no outcome, often characterised either as infinite regress or Byzantine discussions, and where opinions oscillate continuously. Another consequence of cyclic influences is found in echo chambers [51] or cascading effects [52]. In all cases, it is important to establish the necessary conditions that ensure that cyclic conditional opinions converge to stable opinions. It is possible to establish a theorem on deliberation processes as an extension to the Markov Convergence Theorem (MCT). This can be done by looking at the evolution of an IBIS time series over time, its limit distribution, and whether it corresponds to a definite deliberative outcome. This goal could be achieved by finding the deliberative feedback loops in the previously established discussion knowledge graph. Then, it would be possible to measure the propagation of topical utterances in the detected loops. A loop is defined as a closed path, where a path in the graph is a sequence of the form (1).

$$C(t_1, u_1, \theta_1) \rightarrow C(t_2, u_2, \theta_2) \rightarrow \ldots \rightarrow C(t_n, u_n, \theta_n) \qquad (1)$$

where $C(t_i, u_i, \theta_i)$ is a random variable that describes the textual content of the underlying node, given that it has an IBIS type θ_i and is posted by user i at time t. A feedback loop is therefore defined as a path with $C(t_i, u_i, \theta_i) \simeq_T C(t_{i+1}, u_{i+1}, \theta_{i+1})$ and where $\simeq_T$ is a semantic comparison operator that operates on embedded textual content [18]. After algorithmically extracting the loops from a given knowledge graph G, it is possible to quantify the directed propagation of information within the paths of G. This could be achieved using a class of measures defined as the sum of the conditional mutual information I over all feedback loops L of the knowledge graph G, given a topic model T and IBIS types $\Theta = \{Issue, Idea, Pros, Cons\}$. This measure is illustrated in (2).

$$\Delta(G,T) = \sum_{\substack{L \subset G \\ L = C_1 \rightarrow \ldots \rightarrow C_n}} I(C_1; \ldots; C_n | T, \Theta). \qquad (2)$$

The estimation of the function I is done using language models operating on the discussion corpora [56].

4.3 Optimising Deliberative Feedback Loops

First, the measure (2) is used to determine which of two discussions, A or B, has better deliberation with respect to a topic T. This is done by comparing $\Delta(G_A, T)$ and $\Delta(G_B, T)$, where G_A and G_B are extracted IBIS knowledge graphs of discussions A and B, respectively. To increase the deliberative nature of a discussion's subgraph γ, a conversational agent increases the deliberation by algorithmically optimising a function of the form $\arg\max_\gamma \Delta(\gamma, T)$.

In practice, conversational agents must use facilitation strategies [18, 25] that add nodes (posts, or messages) within the discussion threads (sub-graphs) to create more feedback loops. The grafted nodes depend on the target users and topics, in the sense that argumentative utterances have to be inserted in particular locations of the discussion graph. To illustrate this grafting process, let us take an example where three users u_1, u_2, u_3 are exchanging messages consecutively at t_1, t_2, t_3. The messages are classified to their IBIS types $\theta_1, \theta_2, \theta_3 \in \Theta$, with $\Theta = \{Idea, Issue, Pros, Cons\}$. The first step taken by a conversational agent $a_{1 \leq i \leq N}$ of a CMAS is to identify the path shown in (3) with respect to some topic model T. Once a similarity is established between $C(t_1, u_1, \theta_1)$, $C(t_2, u_2, \theta_2)$ and $C(t_3, u_3, \theta_3)$, the agent a_i generates a new utterance that is conditioned on $\theta_{1,2,3}$, adding it below node $C(t_3, u_3, \theta_3)$ as in (4).

$$C(t_1, u_1, \theta_1) \rightarrow C(t_2, u_2, \theta_2) \rightarrow C(t_3, u_3, \theta_3) \qquad (3)$$

$$C(t_1, u_1, \theta_1) \rightarrow C(t_2, u_2, \theta_2) \rightarrow C(t_3, u_3, \theta_3) \rightarrow C(t_4, a_i, \theta_4) \qquad (4)$$

The idea behind this mechanism is to constantly bring key ideas, issues, or arguments, to the attention of the participants. This method aims to reduce lengthy, disjoint, and insular threads of discussion [51]. Detecting topical similarities and dissimilarities is a tedious cognitive task that could be delegated to conversational agents that are constantly mining arguments, topics, and interacting with the participants to create more deliberative loops.

To implement this mechanism, it is possible to use automated forms of facilitation relying on the conversational multiagent system (CMAS) illustrated in Figure 4.

The conversational multiagent platform is composed of the conversational agents that are assigned to the branches of the discussion, an NLP engine for the extraction of the IBIS elements, an argumentative engine [8, 12, 16], and a fact checker that mines and validates the claims against ground truths. The argumentation

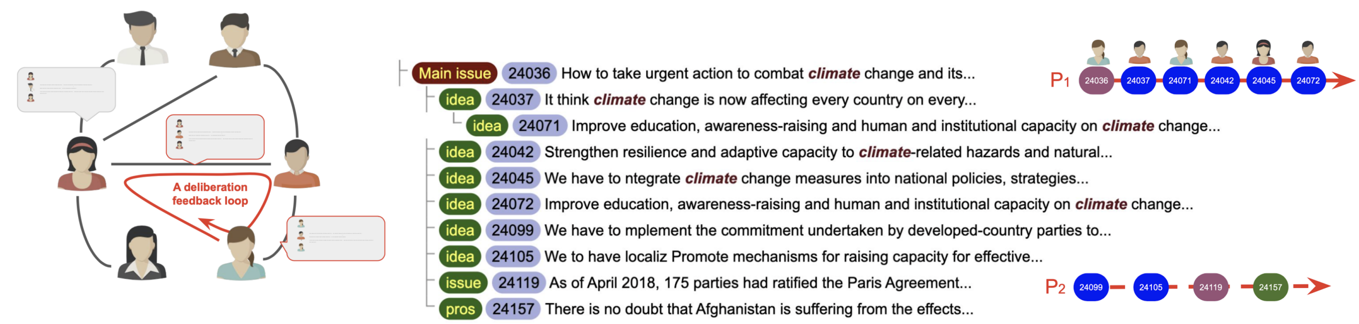

Figure 3: To find the deliberative loops (red arrows), the discussion tree (right) is initially analysed with respect to the topic on "climate". Path P_1 represents the propagation of posts that are related to the topic. Path P_2 is disjoint with respect to the topic of climate. The goal is to identify closed paths in any graph and quantify their topical strength.

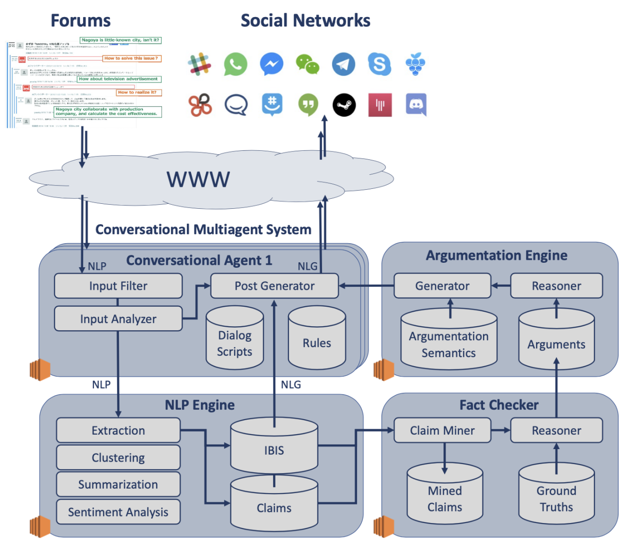

Figure 4: Argumentative conversational multiagent system

engine takes the constructed IBIS elements, constructs the arguments, defines the interactions between the arguments, evaluates the arguments, selects the acceptable arguments, and generates the corresponding utterances while trying to maximise the deliberative loops. The effects of the agents on the deliberation process could be studied with respect to convergence properties. Specifically, one could look at the evolution of the deliberation value Δ over time for the same threads of discussion.

5 SUMMARY AND CHALLENGES

We propose augmenting the classical forms of deliberation in online platforms using conversational multiagent systems (CMAS). We start by quantifying deliberation in argumentative discussions and then proposing an agent-based method to optimise a deliberation measure. The proposed idea is expected to pave the way to a form of online democratic deliberation that is built on the cooperation of humans and conversational agents.

The first challenge that a deliberative CMAS has to face is the scalability problem. Large-scale participatory and deliberative processes are known to result in conflicts [9, 32]. The produced content is also prone to misinformation and may not always be factual, and thus it requires some form of validation [33, 53]. Therefore, it is a challenge to incentivise wider populations to take part in lengthy deliberative processes.

The second challenge is the design choices and the biases they could generate when adopting specific forms of deliberation [48]. The adopted argumentative or deliberative schemes often follow a cultural style that might not be appropriate for certain populations. Western cultural standards are predominantly adopted in deliberation research [40]. However, different cultures have distinct conversation, argumentation, deliberation, and consensus styles [5, 6, 40, 45]. Designing deliberative CMAS must therefore account for a wider spectrum of cultural specificities, gender behaviours, communication styles, cognitive abilities, languages, social norms, and so forth [1].

The third challenge that must be addressed is the potential polarising effect of the CMAS. Several studies have investigated the effect of artificial agents on social media [3, 19, 42, 51]. In particular, the type of reinforcement created by the agents should not lead to echo chambers or cascading effects that cause certain threads to be insulated from others [51, 52]. The risks associated with this form of augmented deliberation calls for regulatory measures to cope with a new political sphere in which humans and conversational agents interact safely on social media. The third challenge raises more issues as to whether people should entrust deliberation processes to conversational agents. This comes at a time when social bots are negatively perceived due to their malicious activities [4, 7, 13, 54, 55]. Establishing trust between humans and agents could, for instance, benefit from consensus mechanisms that build Proofs of Trust (PoT) prior to or during deliberations [41].

ACKNOWLEDGMENTS

This work was partially supported by a JST AIP Challenge Program Grant (2021), a JST CREST Grant (JPMJCR20D1), and a JSPS Grant-in-Aid for Scientific Research (C) (JP20K11936).

REFERENCES

[1] Nardine Alnemr. 2020. Emancipation cannot be programmed: blind spots of algorithmic facilitation in online deliberation. *Contemporary Politics* 26, 5 (2020), 531–552.

[2] Emmanuelle Avril. 2014. Social networks and democracy: Fightbacks and backlashes in the world wide agora. In *Democracy, Participation and Contestation*. Routledge, 243–255.

[3] Christopher A Bail, Lisa P Argyle, Taylor W Brown, John P Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. 2018. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences* 115, 37 (2018), 9216–9221.

[4] Marco T Bastos and Dan Mercea. 2019. The Brexit botnet and user-generated hyperpartisan news. *Social science computer review* 37, 1 (2019), 38–54.

[5] AndréA BäUchtiger and Dominik Hangartner. 2010. When deliberative theory meets empirical political science: Theoretical and methodological challenges in political deliberation. *Political Studies* 58, 4 (2010), 609–629.

[6] Joshua Becker, Ethan Porter, and Damon Centola. 2019. The wisdom of partisan crowds. *Proceedings of the National Academy of Sciences* 116, 22 (2019), 10717–10722.

[7] Alessandro Bessi and Emilio Ferrara. 2016. Social bots distort the 2016 US Presidential election online discussion. *First monday* 21, 11-7 (2016).

[8] Claudette Cayrol and Marie-Christine Lagasquie-Schiex. 2005. On the acceptability of arguments in bipolar argumentation frameworks. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*. Springer, 378–389.

[9] Joshua Cohen. 2018. *Reflections on deliberative democracy*. Routledge.

[10] Jeff Conklin and Michael L Begeman. 1989. gIBIS: A tool for all reasons. *Journal of the American Society for Information Science* 40, 3 (1989), 200–213.

[11] Nicole Doerr. 2012. Translating democracy: How activists in the European Social Forum practice multilingual deliberation. *European Political Science Review* 4, 3 (2012), 361–384.

[12] Phan Minh Dung. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence* 77, 2 (1995), 321–357.

[13] Emilio Ferrara. 2017. Disinformation and social bot operations in the run up to the 2017 French presidential election. *arXiv preprint arXiv:1707.00086* (2017).

[14] Thomas William Flynn. 2011. *Debating deliberative democracy: how deliberation changes the way people reason*. Ph.D. Dissertation. University of York.

[15] Eleonore Fournier-Tombs and Giovanna Di Marzo Serugendo. 2020. DelibAnalysis: Understanding the quality of online political discourse with machine learning. *Journal of Information Science* 46, 6 (2020), 810–822.

[16] John Fox and Simon Parsons. 1997. On using arguments for reasoning about actions and values. In *Proceedings of the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning, Stanford*. Citeseer, 55–63.

[17] Paul Gölz, Anson Kahng, Simon Mackenzie, and Ariel D Procaccia. 2018. The fluid mechanics of liquid democracy. In *International Conference on Web and Internet Economics*. Springer, 188–202.

[18] Rafik Hadfi, Jawad Haqbeen, Sofia Sahab, and Takayuki Ito. 2021. Argumentative Conversational Agents for Online Discussions. *Journal of Systems Science and Systems Engineering* (2021), 1–15.

[19] Rafik Hadfi, Naoki Kawamura, Atsuya Sakai, Naoko Yamaguchi, and Takayuki Ito. 2020. A Study on the Polarisation Effects of Biased Conversational Agents in Online Debates. (2020), 2G6ES305–2G6ES305.

[20] Jawad Haqbeen, Takayuki Ito, Rafik Hadfi, Tomohiro Nishida, Zoia Sahab, Sofia Sahab, Shafiq Roghmal, and Ramin Amiryar. 2020. Promoting Discussion with AI-based Facilitation: Urban Dialogue with Kabul City. (2020).

[21] Jawad Haqbeen, Takayuki Ito, Sofia Sahab, Rafik Hadfi, Shun Okuhara, Nasim Saba, Murtaza Hofiani, and Umar Baregzai. 2020. A Contribution to COVID-19 Prevention through Crowd Collaboration using Conver-sational AI & Social Platforms. (2020).

[22] Pablo Hernandez-Leal, Bilal Kartal, and Matthew E Taylor. 2019. A survey and critique of multiagent deep reinforcement learning. *Autonomous Agents and Multi-Agent Systems* 33, 6 (2019), 750–797.

[23] Michael Herz. 2016. E-rulemaking's democratic transformation: Anticipated, actual, and potential. *Revue Internationale des Gouvernements Ouverts* 3 (2016), 195–208.

[24] Luca Iandoli, Mark Klein, and Giuseppe Zollo. 2007. Can we exploit collective intelligence for collaborative deliberation? The case of the climate change collaboratorium. (2007).

[25] Takayuki Ito, Rafik Hadfi, and Shota Suzuki. 2021. An Agent that Facilitates Crowd Discussion. *Group Decision and Negotiation* (2021), 1–27.

[26] Takayuki Ito, Yuma Imi, Takanori Ito, and Eizo Hideshima. 2014. COLLAGREE: A faciliator-mediated large-scale consensus support system. *Collective Intelligence 2014* (2014), 10–12.

[27] Takayuki Ito, Yuma Imi, Motoki Sato, Takanori Ito, and Eizo Hideshima. 2015. Incentive mechanism for managing large-scale internet-based discussions on COLLAGREE. *Collective Intelligence* 2015 (2015).

[28] Martin Karlsson. [n.d.]. What does it take to make online deliberation happen?: A comparative analysis of 28 online discussion forums. In *Fourth international conference on Online deliberation*.

[29] Mark Klein. 2011. How to harvest collective wisdom on complex problems: An introduction to the mit deliberatorium. *Center for Collective Intelligence working paper* (2011).

[30] Jon Kolko. 2012. *Wicked problems: Problems worth solving*. Ac4d Austin Center for Design.

[31] W. Kunz and H.W.J. Rittel. 1970. *Issues as Elements of Information Systems*. Number 131 in California. University. Center for Planning and Development Research. Working paper, no. 131. Institute of Urban and Regional Development, University of California.

[32] Cristina Lafont. 2019. *Democracy without shortcuts: A participatory conception of deliberative democracy*. Oxford University Press.

[33] Sung-Chul Lee, Jaeyoon Song, Eun-Young Ko, Seongho Park, Jihee Kim, and Juho Kim. 2020. Solutionchat: Real-time moderator support for chat-based structured discussion. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[34] Christian List. 2017. Democratic deliberation and social choice: A review. (2017).

[35] Qiao Liu, Yang Li, Hong Duan, and Yao Liu. 2016. Knowledge graph construction techniques. *Journal of computer research and development* 53, 3 (2016), 582.

[36] Michael A Livermore, Vladimir Eidelman, and Brian Grom. 2017. Computationally assisted regulatory participation. *Notre Dame L. Rev.* 93 (2017), 977.

[37] Martin Lodge and Kai Wegrich. 2015. Crowdsourcing and regulatory reviews: A new way of challenging red tape in British government? *Regulation & Governance* 9, 1 (2015), 30–46.

[38] Thomas W Malone. 2018. *Superminds: The surprising power of people and computers thinking together*. Little, Brown Spark.

[39] Thomas W Malone and Mark Klein. 2007. Harnessing collective intelligence to address global climate change. *Innovations: Technology, Governance, Globalization* 2, 3 (2007), 15–26.

[40] Seong-Jae Min. 2020. On the Westerness of deliberation research. *Journal of Deliberative Democracy* 10, 2 (2020).

[41] Du Mingxiao, Ma Xiaofeng, Zhang Zhe, Wang Xiangwei, and Chen Qijun. 2017. A review on consensus algorithm of blockchain. In *2017 IEEE international conference on systems, man, and cybernetics (SMC)*. IEEE, 2567–2572.

[42] Mert Ozer, Mehmet Yigit Yildirim, and Hasan Davulcu. 2019. Measuring the polarization effects of bot accounts in the us gun control debate on social media. In *Proceedings of ACM Conference (Conference'17)*.

[43] Paulo Savaget, Tulio Chiarini, and Steve Evans. 2019. Empowering political participation through artificial intelligence. *Science and Public Policy* 46, 3 (2019), 369–380.

[44] 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.

[45] Feng Shi, Misha Teplitskiy, Eamon Duede, and James A Evans. 2019. The wisdom of polarized crowds. *Nature human behaviour* 3, 4 (2019), 329–336.

[46] Hirokazu Shirado and Nicholas A. Christakis. 2020. Network Engineering Using Autonomous Agents Increases Cooperation in Human Groups. *iScience* 23, 9 (2020), 101438. https://doi.org/10.1016/j.isci.2020.101438

[47] Clay Shirky. 2009. *Here comes everybody: How change happens when people come together*. Penguin UK.

[48] 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).

[49] Wynn C Stirling. 2016. *Theory of social choice on networks: preference, aggregation, and coordination*. Cambridge University Press.

[50] Jennifer Stromer-Galley. 2020. Measuring deliberation's content: A coding scheme. *Journal of Deliberative Democracy* 3, 1 (2020).

[51] Cass R Sunstein. 2001. *Echo chambers: Bush v. Gore, impeachment, and beyond*. Princeton University Press Princeton, NJ.

[52] Cass R Sunstein. 2017. Deliberative trouble? Why groups go to extremes. In *Multi-Party Dispute Resolution, Democracy and Decision-Making*. Routledge, 65–95.

[53] Ilse Verdiesen, Virginia Dignum, and Jeroen Van Den Hoven. 2018. Measuring moral acceptability in E-deliberation: A practical application of ethics by participation. *ACM Transactions on Internet Technology (TOIT)* 18, 4 (2018), 1–20.

[54] Samuel C Woolley. 2016. Automating power: Social bot interference in global politics. *First Monday* (2016).

[55] Savvas Zannettou, Michael Sirivianos, Jeremy Blackburn, and Nicolas Kourtellis. 2019. The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans. *Journal of Data and Information Quality (JDIQ)* 11, 3 (2019), 1–37.

[56] Zhuosheng Zhang, Yuwei Wu, Hai Zhao, Zuchao Li, Shuailiang Zhang, Xi Zhou, and Xiang Zhou. 2020. Semantics-aware BERT for language understanding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 9628–9635.