**SFI COLLABORATIVE WORKING GROUP PROPOSAL**

Please complete this form and the appropriate budget spreadsheet and return to [hilary@santafe.edu](mailto:hilary@santafe.edu) for review.

You should consult with Hilary about potential meeting dates and room availability prior to submitting your proposal.

* 1. **Title of Meeting**:

The shapes and shaping of consensus in online conversations

* 1. **Type of Meeting** (Working Group or Workshop): Working Group

* 1. **Organizer(s) & Affiliations**:

N. Gizem Bacaksizlar (Santa Fe Institute)

Marco Pangallo (JSFM/University of Oxford)

Graham Heimberg (JSFM/Broad Institute)

Lynette Shaw (Michigan Society of Fellows/University of Michigan)

* 1. **Dates of Event** (actual meeting days):

August 24th – 27th, 2019

* 1. **Meeting Room Preference**:

Due to anticipated space constraints at SFI during this time, we propose that much of our meeting will be held offsite at a shared AirBnB.

* 1. **Budget:** See attached
  2. **Summary of Budget & Funding Sources**:

Total budget amount: $3,984

Amount/identity of other funding sources: James S. McDonnell Foundation 21st Century Science Initiative in Studying Complex Systems - Collaborative Award. (364)

Amount to be raised: $0

* 1. **Meeting Summary (1-2 paragraphs for SFI Website)**

Large scale, online discussions are a ubiquitous feature of contemporary society. Arguably, they have also become one of the most significant new forces in the shaping of public opinion. Many have been concerned that these new spheres of public discourse bear some responsibility for the well-documented rise of political polarization and cross-group antipathy.

To date, however, a systematic understanding of how different online discussion platforms capture the shape of existing opinion spaces or potentially influences their constitution remains to be developed. This project endeavors to begin addressing this gap by theorizing and empirically investigating how three different platforms affect the way online discussions are generated from, and potentially reshape, the latent opinion space of their participants.

* 1. **2-5 Page Research Proposal** (include participant list, affiliation and contact info)

***The shapes and shaping of consensus in online conversations***

**Motivation/Background**

Large scale, pseudo-anonymous online discussion forums are now ubiquitous throughout society. They have also arguably become one of the most significant new forces of our time in capturing and shaping mass public opinion. Many have been rightly concerned that these new spheres of public discourse bear some responsibility for the well-documented rise of political polarization and cross-group antipathy at the national level[[1]](#footnote-1)[[2]](#footnote-2)[[3]](#footnote-3)[[4]](#footnote-4)[[5]](#footnote-5). In response to these concerns, a new field of "Civic Technology" has emerged with an aim of exploring how digital platforms might be designed to more actively foster civility and consensus in such discussions. To date, however, a systematic understanding of how different online discussion platforms capture the shape of existing opinion spaces and potentially influence them remains to be developed. This project will endeavor to address this gap via investigating the following research questions:

1. **How do different online discussion platforms capture or contributing to shaping consensus and civility?**

Online discussion platforms provide data that we can use to characterize the nature of the ensuing conversations. We will use this data to develop a latent representation of an opinion-space and quantify opinion diversity within a conversation. Comparing these metrics across platforms will give us insight into how different platform designs influence the online dialog. We will also develop models of how online dialog in generated for insight into how different platform structures (e.g. upvote/downvote, moderated, anonymous) impact the diversity of opinions, civility, and polarization within a discussion.

1. **Is there a relationship between different types of potential targeted attacks on a given platform and the resulting effectiveness of said attacks on opinion polarization and civility?**

Along with differences in dialog structures, each online platform has different audiences, draws distinct types of attention, and enables or constrains different forms of interaction among participants. Can we understand how these features potentially make platforms more or less resilient to targeted attacks (e.g. bots)? And can we understand how these features make it more or less obvious when such conversations are being influenced said attacks?

**Research Approach**

*Modeling*

To interpret the data, we will build an opinion dynamics model. In this model, agents will have a predisposition towards two different political positions (e.g. pro or against different Women’s March priorities), and potentially change this position as they read and write comments on the online platform. If the comments are too far from their own political positions, agents may reinforce their position and be less likely to reach consensus.

We plan to start building an Agent-Based Model that is solved numerically and highlights key mechanisms while keeping enough details of the different platforms. The plan is then to build a stripped-down opinion dynamics model, based on spin models[[6]](#footnote-6) or a model with continuous opinion space[[7]](#footnote-7) (Deffuant et al. 2000).

*Data Analysis*

This project will analyze the opinion dynamics about the 2017 Women’s March discussions on the three different online discussion platforms, such as Pol.is, Twitter, and Reddit. The first platform Pol.is is an open source platform for facilitating online discussions on specific subjects among large groups. It shows previous user comments on a particular subject and a voting structure for each comment. The second, Twitter is a social media platform that encourages people sharing their opinions and emotions. Users should sign up to be able to tweet. It is a well-known microblogging site that has a large pull of users from the news media channels to public users, politicians, and celebrities. The third platform Reddit is a discussion website with social news and web content ratings. Only registered members can submit and/or comment on a content on Reddit with links, text posts, and images.

We will extract the 2017 Women’s March data over the three platforms to have consistent comparison analyses on opinion changes. The data analysis process will include analyzing user behaviors with various reaction mechanisms on different online discussion platforms, having basic statistical analysis of user activities and comments, and conducting sentiment analysis, topic modeling, and document embedding on comments. We will also investigate identifying opinion dynamics with metrics of civility, polarization, and opinion-space.

*Data and Analysis Descriptive Details*

* + - Reaction mechanisms according to the platform:
      * Pol.is: Agreement, Disagreement, and Neutral
      * Twitter: Likes and Retweets
      * Reddit: Upvotes and Downvotes
    - Consistent unique user analysis over time
      * Basic statistics on user activities
      * Basic statistics on user comments
      * Demographic metadata (when available)
    - Comments (texts) analysis over time
      * Sentiment analysis to detect the polarity of comments
      * Structural topic modeling of comments to identify themes
  + Metrics of civility, polarization, opinion-space
    - PCA dimension of opinion-space
    - Clustering for finding latent/stereotyped opinion profiles. Clustering can be done on vectors obtained from doc2vec.

**Logistics, and participants**

We are proposing to continue the research began during our SFI visit. Each group member is available to return to Santa Fe in the days leading up to our next JSMF-SFI conference, Aug 24th-27th. Due to potentially limited space at SFI, we will rent an airbnb to stay at as well as work during the days. Before our in person working group session, we will virtually meet over Skype on bi-monthly basis in order to perform preliminary data analyses and discuss different models for simulating online discussions.

* N. Gizem Bacaksizlar
  + ngizembacaksizlar@gmail.com
  + Santa Fe Institute / Postdoctoral Fellow
* Marco Pangallo
  + [marcopangallo@gmail.com](mailto:marcopangallo@gmail.com)
  + PhD student at University of Oxford (to start Postdoc at another institution by January 2020)
* Graham Heimberg
  + gheimberg@gmail.com
  + JSMF postdoc at Broad Institute
* Lynette Shaw
  + shawla@gmail.com
  + Center for the Study of Complex Systems, University of Michigan/Postdoctoral Scholar, Michigan Society of Fellows

1. Bail et. al. (2018) Exposure to opposing views on social media can increase political polarization. *Proc Natl Acad Sci USA* 115: 9216-9222. [↑](#footnote-ref-1)
2. Doherty, C. Kiley, J. Dimock, M. and Keeter, S. (2014) Political polarization in the American public: How increasing ideological uniformity and partisan antipathy affect politics, compromise, and everyday life (Pew Res Cent, Washington, DC). [↑](#footnote-ref-2)
3. Papacharissi, Z. (2004) Democracy online: civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society.* 6:259 - 283. [↑](#footnote-ref-3)
4. Iyengar S, Westwood SJ (2015) *Fear and loathing across party lines: New evidence on  
   group polarization.* Am J Polit Sci 59:690–707. [↑](#footnote-ref-4)
5. Farrell, H. (2012) The Consequences of the Internet for Politics. *Annu. Rev. Polit. Sci.* 15:35-52. [↑](#footnote-ref-5)
6. E.g. Galesic, M. and Stein, D.L. (2019) Statistical physics models of belief dynamics: Theory and empirical tests. *Physica A* 519:275-294 [↑](#footnote-ref-6)
7. Deffuant, G., Neau, D., Amblard, F., and Weisbuch, G. (2000) Mixing beliefs among interacting agents. *Advances in Complex Systems* 3:89 - 98 [↑](#footnote-ref-7)