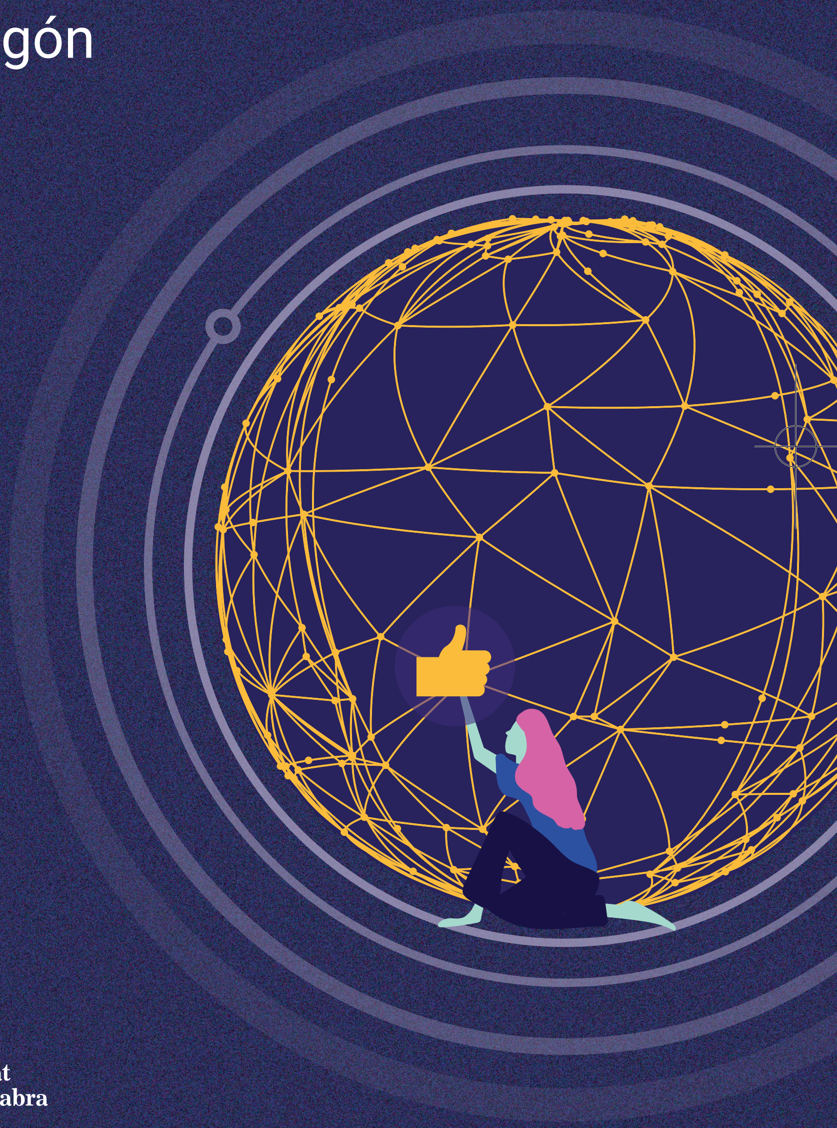
Universitat
Pompeu Fabra
Barcelona

Characterizing Online Participation in Civic Technologies
Pablo Aragón

Universitat
Pompeu Fabra
Barcelona

Characterizing Online Participation in Civic Technologies

Pablo Aragón



Universitat
Pompeu Fabra
Barcelona

DOCTORAL THESIS UPF / 2019

Dept. of Information and Communication Technologies

Characterizing Online Participation in Civic Technologies

Pablo Aragón

DOCTORAL THESIS UPF / 2019

THESIS SUPERVISORS
Dr. Vicenç Gómez
Dr. Andreas Kaltenbrunner

Dept. of Information and Communication Technologies

Universitat Pompeu Fabra Barcelona

Cover design: Berenice Zambrano (@darkolina)

To Marion,
for her unconditional love and support.

“Gracias a la vida, que me ha dado tanto”

— Violeta Parra

Acknowledgement

This manuscript is the result of a long process that was made possible thanks to the contribution of many very good people. Therefore, I wish to acknowledge their support, which goes beyond the scope of the thesis itself.

First, this research has benefited from the outstanding advice of my co-supervisors Vicenç Gómez and Andreas Kaltenbrunner, who have not only shared their experience and wisdom but have also shown great patience with an atypical and stubborn student.

I consider myself privileged to have carried out this PhD at the Information and Communication Technologies Department of Universitat Pompeu Fabra (DTIC - UPF). This ecosystem relies on people as committed as Aurelio Ruiz, whose work is essential to ensure that academic research returns to society; Lydia García, who makes any bureaucratic problem disappear; and Vanessa Jiménez, Magda Castellnou and Ruth Temporal, who always offered me solutions to any request. I would also like to thank my colleagues from Office 55.210 for adopting such a rara avis: Federico Franzoni, Rasoul Silab, Xavier Salleras, Federico Miatton, Zaira Pindado, Arantxa Zapico, Alexandros Zacharakis, and Javier Silva; and to the members of Volleybolud@s, a team capable of becoming runner-up in the DTIC beach volley tournament despite its great dysfunctionality.

A significant part of the PhD was done as a research scientist at Eurecat, Centre Tecnològic de Catalunya. In this regard, I am very grateful for all the efforts made to generate synergies between my research activity in the centre and the PhD. Among many people, I would especially like to thank David Laniado for everything he has taught me in so many years, Matteo Manca, and all my colleagues at the Digital Humanities / Data Science & Big Data Analytics unit.

To address new research challenges in this thesis, it was crucial to have joined the Oxford Internet Institute - University of Oxford as a visiting scholar. This unforgettable experience happened because of the hospitality and support of Scott Hale, and benefited from the contributions of Chico Camargo, Helen Margetts and Taha Yasseri. Also, I will always remember the time spent in Oxford with colleagues as wonderful as Diana Sanz Roig and Johannes Wachs.

Along with these experiences, this thesis has been enriched by collaborations with inspiring scientists like David García, Diego Saéz-Trumper, Miriam Redi and Alberto Lumbreras, and the exhaustive and generous review efforts of Jorge Saldívar, Helena Gallego, Iñaki Iturrate and Ana Valdivia.

In order to investigate technopolitical phenomena such as those analysed in this PhD, one has to come out of academic institutions to fully understand their complexity and, even more importantly, to contribute to them in return. In this sense, the Decidim (Barcelona) project has been an extraordinary life experience. Many of the co-founding members of this project were already great colleagues at DatAnalysis15M: Arnau Monterde, Xabier Barandiaran, Antonio Calleja-López and Juan Linares. Moreover, this project has allowed me to meet new allies who are now as important to me as Carol Romero and Andrés Pereira, along with the members of the Decidim Team and the Metadecidim community. It has also been a privilege to be part of the Collective Intelligence for Participatory Democracy Lab (Participa LAB) at Medialab Prado in Madrid. This unique project was made possible thanks to the work done by Yago Bermejo, Alejandra Baciero, Pedro Álvarez, Marcos García and a global community of fantastic people like Berenice Zambrano, who has also designed an impressive cover for this manuscript; Rebeca Díez, Saya Sauliere, and Alberto Abellán among many others. I would also like to acknowledge other very inspiring experiences from (no longer) remote places as the g0v community in Taiwan and the Mexican technopolitical community coordinated by Paola Ricaurte.

Besides research, these years would not have been the same without friends as great as Nuria Alarcón, Nuria Coma, Úrsula Marcilla, Gabriel Cisneros, Roberto Velasco, and Cristina Escuer.

While the interaction with colleagues and friends usually become the visible network of a doctoral project, the critical infrastructure of almost every PhD is family support. In that sense, this work would not have even begun without the hard work and efforts of my parents María del Carmen and José Antonio, my sister Lucía, and my grandparents Carmen and Tasio.

Finally, the person to whom I am most grateful for her loving support throughout these years is Marion. I could never reward her enough for all the sacrifices she has made...but I will strive for the rest of our lives.

Funding

This doctoral project has been supported by the Spanish Ministry of Economy and Competitiveness under the María de Maeztu Units of Excellence Programme (MDM-2015-0502).

Research on this thesis has also received complementary support from the EU project D-CENT (FP7/CAPS 610349) and the CIEN LPS-BIGGER project (UCTR150175, IDI-20141259), co-funded by Centro para el Desarrollo Tecnológico Industrial (CDTI) and Fondo Europeo de Desarrollo Regional (FEDER).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

This thesis constitutes one of the first investigations focused on characterizing online participation in civic technologies, a type of platform increasingly popular on the Internet that allows citizens new forms, on a larger scale, of political participation. Given the opportunities of civic technologies in democratic governance, it should be noted that their design, like that of any online platform, is not neutral. The ways in which information is presented or interaction between users is allowed can greatly alter the results of participation. For this reason, we analyze the impact of different interventions in civic technologies in relation to online conversation views, ordering criteria for ranking petitions, and deliberative interfaces. Since these interventions were carried out by the corresponding development teams, the analyses have required to develop novel computational and statistical methods, while also extending generative models of discussion threads to better characterise the dynamics of online conversations. Results of the different case studies highlight the social and political impact of these interventions, suggesting new directions for future research and the need to develop a paradigm of citizen experimentation for democracy.

Resumen

Esta tesis constituye una de las primeras investigaciones en caracterizar la participación online en tecnologías cívicas, un tipo de plataforma cada vez más popular en Internet que permite a la ciudadanía nuevas formas, a una mayor escala, de participación política. Dadas las oportunidades de las tecnologías cívicas para la gobernanza democrática, cabe señalar que su diseño, al igual que el de cualquier plataforma online, no es neutral. La forma en que se presenta la información o se permite la interacción entre las usuarias puede alterar en gran medida los resultados de la participación. Por este motivo, analizamos el impacto de diferentes intervenciones en tecnologías cívicas en relación a las vistas de las conversaciones online, los criterios de ordenación en rankings de peticiones e interfaces deliberativas. Dado que estas intervenciones fueron llevadas a cabo por los propios equipos de desarrollo, los análisis han requerido desarrollar nuevos métodos computacionales y estadísticos, a la vez que se han ampliado modelos generativos de hilos de discusión para caracterizar mejor la dinámica de las conversaciones online. Los resultados de los diferentes estudios de caso destacan el impacto social y político de estas intervenciones, sugiriendo nuevas líneas de investigación en el futuro y la necesidad de desarrollar un paradigma de experimentación ciudadana para la democracia.

Resum

Aquesta tesi és una de les primeres investigacions que té per objecte la caracterització de la participació en línia en tecnologies cíviques, un tipus de plataforma cada vegada més popular a Internet que permet a la ciutadania noves formes, a major escala, de participació política. Donades les oportunitats de les tecnologies cíviques per a la governança democràtica, cal assenyalar que el seu disseny, com el de qualsevol plataforma en línia, no és neutral. La forma en què com es presenta la informació o es permet la interacció entre les usuàries pot alterar en gran mesura els resultats de la participació. Per aquest motiu, analitzem l'impacte de diferents intervencions en tecnologies cíviques en relació amb les vistes de conversa en línia, els criteris d'ordenació en rànquings de peticions i amb interfícies deliberatives. Atès que aquestes intervencions van ser dutes a terme pels propis equips de desenvolupament, les anàlisis han requerit desenvolupar nous mètodes computacionals i estadístics, alhora que s'han ampliat models generatius de fils de discussió per caracteritzar millor la dinàmica de les converses en línia. Els resultats dels diferents estudis de cas destaquen l'impacte social i polític d'aquestes intervencions, suggerint noves línies d'investigació en el futur i la necessitat de desenvolupar un paradigma d'experimentació ciutadana per a la democràcia.

Resumé

Cette thèse constitue l'une des premières recherches sur la caractérisation de la participation en ligne à des technologies civiques, un type de plateforme de plus en plus populaire sur Internet qui permet aux citoyens de nouvelles formes, à plus grande échelle, de participation politique. Compte tenu des opportunités offertes par les technologies civiques dans la gouvernance démocratique, il convient de noter que leur design, comme celui de toute plateforme en ligne, n'est pas neutre. La façon dont l'information est présentée ou l'interaction entre les utilisateurs est permise peut grandement modifier les résultats de la participation. Pour cette raison, nous analysons l'impact de différentes interventions dans le domaine des technologies civiques par rapport à l'agencementaux des conversations en ligne, aux critères d'ordre de classement des pétitions et aux interfaces délibératives. Comme ces interventions ont été réalisées par les équipes de développement correspondantes, les analyses ont nécessité de développer nouvelles méthodes informatiques et statistiques, tout en élargissant les modèles génératifs de fils de discussion afin de mieux caractériser la dynamique des conversations en ligne. Les résultats des différentes études de cas mettent en évidence l'impact social et politique de ces interventions, suggérant de nouveaux axes de recherches futures et la nécessité de développer un paradigme d'expérimentation citoyenne pour la démocratie.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

The beginning of the 21st century was also the starting point of the so-called 'Web 2.0'. This paradigm shift transformed the original idea of the World Wide Web, based on interlinked resources distributed all over the Internet, by the emergence of platforms designed to facilitate online participation, e.g., social media, social networking sites, blogs and microblogging services, social news sites, wikis, etc. Tom O'Reilly referred Web 2.0 as 'the web as platform' in which the network effects from user contributions become essential [O'R05]. Furthermore, scholars like Jan Van Dijk and Manuel Castells have suggested that the revolution originated by Web 2.0 has led to the rise of the 'network society' [VD99, Cas04]. Innovation under this framework does not only relate to efficiency in how information flows but also to the emergence of new organizational forms. Yochai Benkler examined these structures of organization in his seminal work 'The wealth of networks' [Ben06] and concluded that they break the classical centralized hierarchical logics by promoting new decentralized commons-based production practices. Indeed, online platforms have transformed the practices of our modern society and their capabilities surpass the possibilities one could imagine just a few decades ago.

Despite the revolutionary opportunities provided by web platforms for the rise and constitution of the network society, this phenomenon implies that platforms are central to almost every online practice. For this reason, José van Dijck has indicated that we are actually witnessing the era of the 'platform society':

> *"Many people think of platforms simply as technological tools that allow them to do things online: chatting, sharing, commenting, dating, searching, buying stuff, listening to music, watching videos, hailing a cab, and so on. But these online activities hide a system whose logic and logistics are about more than facilitating: they actually shape the way we live and how society is organized."* [VDPDW18]

In fact, the way platforms are designed has proven to be of great relevance in promoting problems of modern life like the spread of disinformation [DVBZ+16, KWL16], algorithmic biases and discrimination [Tuf15, OKC18], hate speech and misbehaviour [DZM+15, CBDNML17, KCL17], or mass network surveillance [Fuc13, Zub19]. Therefore, to understand our modern society we need to characterize how platforms operate and how they affect us as human beings.

The above claim takes on even greater value as we observe the growing popularity of online platforms for political activities. The Internet itself was heralded for empowering citizens and challenging existing power structures by diversifying the relationship between governments and citizens [MR00, Cas08]. However, this effect is even more noticeable by the increasing number of platforms by civic organizations and national, regional and local governments to scale up citizen participation and to promote new forms of governance.

The potential of Web 2.0 platforms for civic purposes has drawn much attention from academia. Moreover, some existing platforms are the result of scientific research with the aim of enhancing the ability of policy-makers and citizens to make more informed decisions [SL15, GKSA19]. As noted by [Lin12], there is a plethora of competing labels for citizen coproduction in the age of social media:

collaborative government [McG], citizen sourcing [Tor07], wiki government [Nov09], government as a platform [O'R11], do-it-yourself government [MD13], participatory civics [Zuc14], etc. All these notions rely on online platforms for the citizenry to participate. Platforms of this nature are often referred to as 'civic technologies' ('civic tech'), a term motivated by the expected civic outcome of such technological approaches that is gaining popularity in academic and policy communities. Although the outcome from civic-inspired purposes is expected to be beneficial, we must bear in mind that the way in which platforms allow users to participate may differ. Based on the theory of affordances formulated by James J. Gibson [Gib77] which was later extended by William W. Gaver when referring to 'technology affordances' [Gav91], a recent comparison of features in the social platforms *Facebook* and *Twitter* concluded that interaction does not necessarily take place in the same way on them, since the options available to users are not the same:

> *"These are known as the platforms' affordances. In other words, the totality of possibilities these environments offer.* ***Platforms are not neutral.*** *Whether one can respond to a particular contribution or whether one has to address one's remarks to the group at large; whether the platform allows images or links to be posted; whether all participants have the same view of a given exchange and its dynamics – all these aspects play a part in determining the exchange and its outcome."* [Cam18]

The lack of neutrality of technological platforms gives great power and responsibility to the agents in charge of their design. We should be aware that the agent who controls a successful platform controls the interface between systems and participants, and dictates the rules of engagement [HY09]. This observation connects directly to the concept of 'nudge' proposed by Richard H. Thaler and Cass R. Sunstein: "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their incentives" [TS09]. Both Thaler and Sunstein, who became head of President Obama's Office of Information and Regulatory Affairs, supported the idea of 'libertarian paternalism' that can be summed up as nudging people to make better choices (as judged by themselves) without forcing certain outcomes upon anyone [BST14].

In contrast to the paternalistic principles of nudging, we consider that democratic governance requires citizens to be able to evaluate power structures and policy making, at present, to be able to assess the design of civic technologies as they might affect the way the citizenry participates online. Therefore, it becomes crucial to characterize online participation in civic technologies to shed light on the democratic potential and deficits of the platform society.

1.2 Research goals and context

The research scope of this thesis lies at the intersection of two different areas: computational social science and machine learning. The main goal is to provide a better understanding of how interventions in civic technologies influence user behavior. Given the many different forms of participation in platforms of this kind (citizen forums, online petitions, voting processes, crowdfunding, collaborative mapping, etc.) and the many existing case studies all over the world, this dissertation focuses on three specific civic technologies developed in Spain: the social news site *Menéame*[1], and the platforms for participatory democracy *Decide Madrid*[2] and *Decidim Barcelona*[3], which are developed by the municipalities of the two corresponding cities.

The reasons for choosing these three civic technologies are multiple. In the case of *Menéame*, this site is not only a very influential forum of political discussion but also an ideal opportunity to examine the impact of an intervention of great interest: the adoption

[1] `https://www.meneame.net/`
[2] `https://decide.madrid.es/`
[3] `https://www.decidim.barcelona/`

of conversation threading for the visual representation of discussion threads. In the case of *Decide Madrid* and *Decidim Barcelona*, this doctoral research project has been carried out in close collaboration with the development team of each platform. This context provides an unusual opportunity to get informed about the interventions in each site and the motivations for doing so. Interestingly, as suggested in a comprehensive survey on controlled experiments on the web [KLSH09], platform features have often been determined by the same way medicine was prescribed prior to World War II: by people who were regarded as experts, not by using scientific methods. Therefore, a practical goal of this doctoral research project has been to provide advice by measuring the effects of: (1) the change of the petition ranking in the home page of *Decide Madrid*, and (2) the deliberative platform design in *Decidim Barcelona*.

It is true that the characterization of the impact of platform interventions is often conducted through online field experiments like A/B testing. Indeed, there is abundant literature on best practices for these techniques [KLSH09, KDF+13, BEB14, CK15, MM18]. However, civic technologies like *Decide Madrid* and *Decidim Barcelona* have been developed to foster participatory policy-making in the corresponding cities. For this reason, approaches like A/B testing would present major ethical issues concerning the experimentation with real citizens participating in decision-making processes that directly affect their living conditions. As a consequence, research in this thesis relies on observational studies through mining data about participants without participant awareness or environmental manipulation (Type IV of Human Subjects Research [BT15]). Studying user behaviour in non-controlled online environments presents some limitations, even more for the case of *Menéame* since many details about platform interventions are unknown a priori. Nevertheless, these limitations have encouraged us to address a methodological goal: to develop statistical methods to infer events that might have affected user behaviour and to extend data-driven models to characterize the mechanisms governing the dynamics of online participation.

1.3 Outline and contributions

In this dissertation, we will characterize online participation in civic technologies through statistical and computational techniques. As shown in Figure 1.1, research is divided into two parts: the first one with two chapters focused on online discussions, and the second one with two chapters focused on online petitions[4]. Research chapters are preceded by this introductory chapter and the following chapter which reviews the state of the art and background of the thesis. The manuscript includes a final chapter presenting the conclusions, and two appendices with complementary research to this project.

Chapter 2: In this chapter we review related work on civic technologies, dynamics of online discussion threads, dynamics of online petitions, and platform interventions. The survey of generative models of online discussion threads is published in:

> [AGGK17] *Pablo Aragón, Vicenç Gómez, David García, and Andreas Kaltenbrunner. Generative models of online discussion threads: state of the art and research challenges.* Journal of Internet Services and Applications*, 8(1):15, 2017.*

and was the basis for tutorials given at:

- *The 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2018)*
- *The 13th International AAAI Conference on Web and Social Media (ICWSM-19)*

[4]The author of this dissertation is the main contributor of the publications listed in this section

PART I: ANALYSIS OF ONLINE DISCUSSIONS

We first analyze the structure of the discussion threads from *Menéame* to propose a methodology to identify platform interventions. Then, we model the impact of a specific intervention: the adoption of conversation threading as default criterion to represent discussions.

Chapter 3: This chapter addresses the challenge of automatically detecting events that affect online deliberation in online discussions. To do so, we develop a language-independent methodology based on regression discontinuity design, structural metrics of online discussions threads and data from *Menéame*. The work described in this chapter is included in:

> [AGK17a] *Pablo Aragón, Vicenç Gómez, and Andreas Kaltenbrunner. Detecting Platform Effects in Online Discussions.* Policy & Internet*, 9(4):420–443, 2017.*

Chapter 4: In this chapter we explicitly focus on the platform intervention in *Menéame* which replaced the linear interface of discussions with a hierarchical interface (conversation threading). We replicate our methodology based on regression discontinuity design using different metrics of social reciprocity. In addition to this, we propose a novel generative model of discussion threads which includes reciprocity as behavioral feature to explain the structure and growth of online discussion threads. Results of this chapter are published in:

> [AGK17b] *Pablo Aragón, Vicenç Gómez, and Andreas Kaltenbrunner. To Thread or Not to Thread: The Impact of Conversation Threading on Online Discussion. In* Proceedings of the International AAAI Conference on Web and Social Media*, 2017.*

and were also presented at *The 3rd Annual International Conference on Computational Social Science (IC2S2 2017).*

PART II: ANALYSIS OF ONLINE PETITIONS

We first characterize petition growth in *Decide Madrid* and evaluate the improvement originated by changing the algorithm for ranking petitions in the home page. Then, we assess the effectiveness of the deliberative platform design for petitions in *Decidim Barcelona*.

Chapter 5: This chapter presents a study of petition growth in the platform *Decide Madrid* based on time-series analysis and clustering. In particular, we examine the spillover effects between multiple simultaneous participatory processes in a same platform, and detect undesired effects of the original ranking algorithm in the home page. Furthermore, we present how our analysis has motivated a new ranking algorithm in *Decide Madrid* and the impact of such intervention in signing behaviour.

The work of this chapter was presented as a non-archival poster at *The 13th International AAAI Conference on Web and Social Media (ICWSM-19)* and the resulting manuscript is under review.

Chapter 6: In this chapter we focus on the case study of *Decidim Barcelona*, in which petitions can be discussed with an interface that combines threaded discussions and comment alignment with the petition. This innovative approach allows to examine whether neutral, positive or negative comments are more likely to generate discussion cascades. The results of this chapter are published in:

> [AKCL+17] *Pablo Aragón, Andreas Kaltenbrunner, Antonio Calleja-López, Andrés Pereira, Arnau Monterde, Xabier E. Barandiaran, and Vicenç Gómez. Deliberative Platform Design: The Case Study of the Online Discussions in Decidim Barcelona. In* Social Informatics*, pages 277–287, Cham, 2017. Springer International Publishing.*

and were also presented at *The European Symposium on Societal Challenges in Computational Social Science 2019: Polarization and Radicalization.*

Chapter 7: We finally conclude this dissertation by summarizing and discussing the main findings and suggesting new directions for future research.

APPENDICES

This manuscript also includes two appendices with the results of complementary research to the work.

Appendix A: This appendix presents the exploration of a dataset we built with more than 350K petitions from *Avaaz.org*. In particular, we examine how social media campaigning is related to the success of petitions, as well as some geographic and linguistic findings about the worldwide community of *Avaaz.org*. We conclude with example research questions that could be addressed with our dataset. The results of this appendix are published in:

> [ASTR+18] *Pablo Aragón, Diego Sáez-Trumper, Miriam Redi, Scott Hale, Vicenç Gómez, and Andreas Kaltenbrunner. Online Petitioning Through Data Exploration and What We Found There: A Dataset of Petitions from Avaaz.org. In* Proceedings of the International AAAI Conference on Web and Social Media*, 2018.*

Appendix B: In this appendix we describe two knowledge discovery systems for *Decide Madrid*. These systems are the result of a fruitful collaboration between 2016 and 2019 with the Collective Intelligence for Participatory Democracy Lab (Participa LAB). The details of these systems are published in:

> [AGK16] *Pablo Aragón, Vicenç Gómez, and Andreas Kaltenbrunner. Visualization Tool for Collective Awareness in a Platform of Citizen Proposals. In* Proceedings of the International AAAI Conference on Web and Social Media*, 2016.*

[ABGK18] *Pablo Aragón, Yago Bermejo, Vicenç Gómez, and Andreas Kaltenbrunner. Interactive Discovery System for Direct Democracy. In* 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 601–604, 2018.*

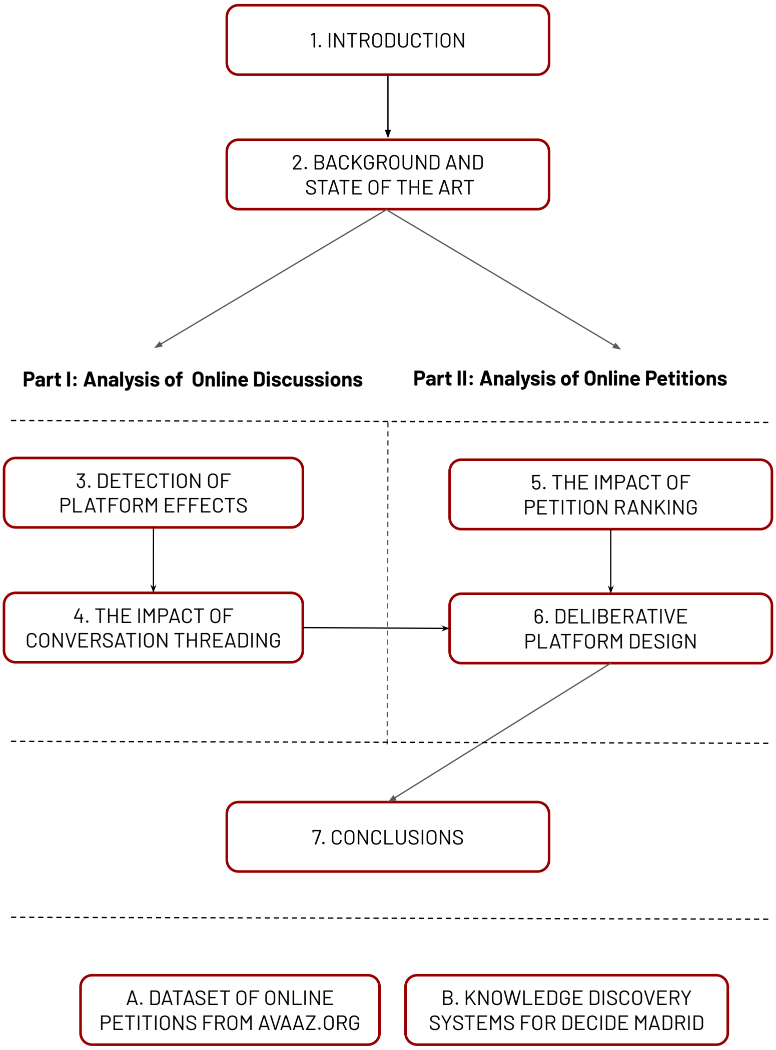


Figure 1.1. Conceptual scheme of this manuscript.

2 Background and State of the Art

2.1 Introduction

In this chapter we provide an overview of research on civic technologies, dynamics of online discussion threads, dynamics of online petitions, and the effects of platform interventions. This review is intended to make readers familiar with the background of this thesis.

2.2 Civic technologies

This thesis focuses on a specific type of online platforms with a especially evident political dimension: 'civic technologies'. Many definitions of platforms have focused on their technical side, e.g., *"providers of software, (sometimes) hardware, and services that help code social activities into a computational architecture"* [VD13] or *"programmable digital architectures designed to organize interactions between users – not just end users but also corporate entities and public bodies"* [VDPDW18]. However, as introduced in Chapter 1, platforms are more than just technological tools, the social logics they offer turn them into socio-technical systems. For this reason, we agree with the approach by Tarleton Gillespie who considered not only the computational and architectural aspects of platforms, but also their role as intermediaries between the participants, having to understand platforms in a political sense [Gil10].

Regarding online platforms precisely built for political processes, one of the first studies about the emergence of civic technologies is the report by the Knight Foundation in 2013 which defined them as those technologies *"used to inform, engage and connect residents with government and one another to advance civic outcomes"* [PSGH13]. That work divided civic technologies into two clusters:

- Open government: data access and transparency, data utility, decision making, resident feedback, mapping and visualization, and voting.

- Community action: tools for civic crowd funding, community organizing, information crowdsourcing, neighborhood forums, and peer-to-peer sharing.

Interestingly, a recent survey of civic technology for social innovation [SPA$^+$19] also reported this dichotomy in academic literature:

- Government-centric definitions, e.g., *"use of technology by cities for service provision, civic engagement, and data analysis to inform decision making"* [Lin12].

- Citizen-centric definitions, e.g., *"platforms and applications that enable citizens to connect and collaborate with each other and with government"* [Cit12].

That survey then unified both approaches through their common element: *"enabling participation in democratic governance (i.e., the many activities citizens undertake to negotiate living together in society)"* [SPA$^+$19]. This definition better captures the context of this thesis as it does not limit the scope to technologies requiring a government to get involved and, especially, it focuses on technologies that foster citizen participation.

The extent to which citizens hold power in participatory processes has been traditionally categorized through the Arnstein's Ladder of Citizen Participation [Arn69] which ranges from none to real citizen control. Because of the emergence of civic technologies, new frameworks have been proposed, e.g., the e-Participation Ladder [Kin02] (see Figure 2.1) in which participants benefit from the bidirectional communication of these platforms to collectively discuss and decide online. This is the case of the three civic technologies of this thesis, since participants have to discuss and make decisions online: news to appear in a front page (*Menéname*), and public policies to be implemented by a city council (*Decide Madrid* and *Decidim Barcelona*).

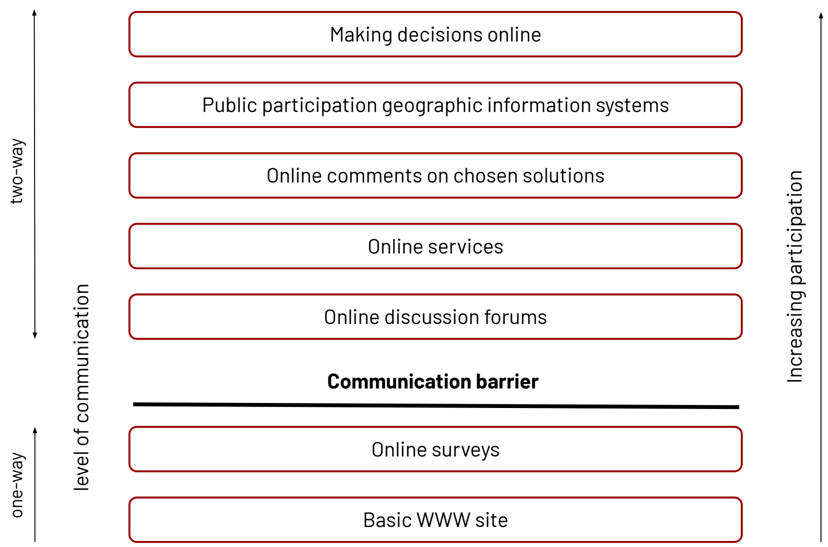


Figure 2.1. e-Participation Ladder. Adapted from: [Kin02]

2.3 Dynamics of online discussion threads

Discussions in online platforms commonly occur as an exchange of written messages among two or more participants. In this way, conversations are often represented as threads, which are initiated by a user posting a starting message (hereafter post) followed by (possibly other) users sending replies to either the post or the existing replies. Therefore, given this sequential posting behavior, online discussion threads follow a tree network structure as shown in Figure 2.2.

The properties of this network structure represent the typical features of the dynamics of online discussion threads. *Popularity* is a classic feature which expresses that the more connected a node is (the most commented messages), the more likely a node is to attract new edges (new replies). This feature is usually introduced through the preferential attachment process [BA99], also referred to as the Yule process [Sim55]. This process is a common property of many social networks and establishes that the probability that one of the links of a new node connects to certain node depends on its degree. Thus, node 1 in Figure 2.2 is the most popular message, followed by node 2, then nodes 5 and 6 and, finally, the rest of the nodes. Besides the popularity of nodes, it is usually expected that the newest comments are the most attractive messages. Therefore, another relevant feature is *novelty*, i.e., nodes 1 and 2 are the most popular ones but also the oldest ones, which might reduce the arrival of new replies. Moreover, notice an important difference: node 1 is the initial post while node 2 is a comment, like the rest of the nodes. Some users might be interested in replying directly to the post while some other users might be interested in replying to comments and getting engaged in the discussion. This is the reason why certain models establish a *root-bias* as a feature. In addition to this, users often engage in discussions between pairs of users resulting into chains of messages (segments), e.g., nodes 5 and 9 are likely to be posted by the same user. In consequence, other informative features are *segment lengths* or *reciprocity* by also considering the authorship of messages, i.e., users tend to reply to comments that replied to their previous messages. Furthermore, the consideration of authorship allows the definition of features related to *social influence* and *user roles*, i.e., certain users might follow a specific behavior. Apart from all these structural features of online discussion threads, features can also consider linguistic patterns from the messages like the occurrence of certain text expressions.

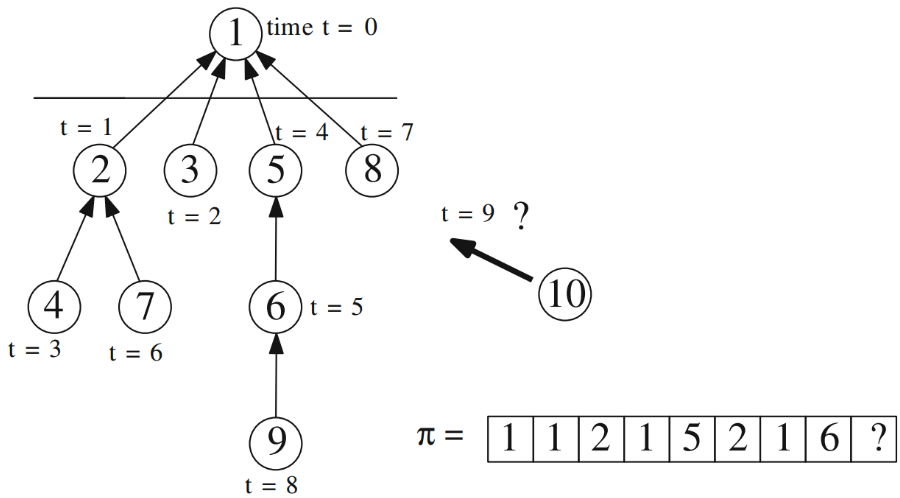


Figure 2.2. Example of a discussion thread represented as a tree: at time-step $t = 9$, node (comment) number 10 is added to the thread. The parent of each node is sequentially stored in vector π. Source: [GKLK13].

2.3.1 Generative models of online discussion threads

The features described above are essential to model the dynamics of online discussion threads. We refer to modeling as defining a mathematical description of a process that generates online discussion threads, as illustrated in Figure 2.3.

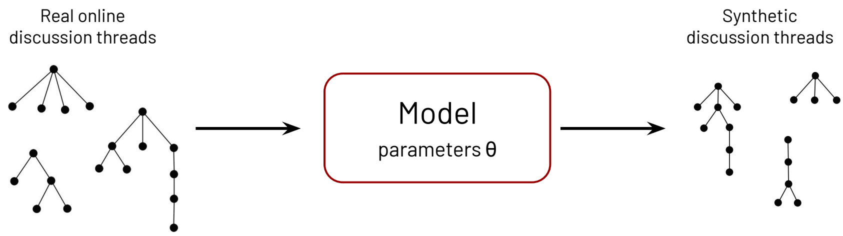


Figure 2.3. Modeling approach for discussion threads considered in this work: the model (box in the middle) represents a mechanism or procedure that describes how discussion threads are formed. It is usually governed by a set of parameters θ which are typically learned from real data composed of real discussion threads. This learning step involves some type of optimization. For given parameters θ, the model can be used to generate synthetic threads that reproduce the properties of the real discussion threads.

We consider data-driven models, that is, models that try to capture some phenomena of interest of a given dataset [SS98, Mur12]. Such models are constrained by the nature of the data and the type of phenomena that they try to explain. For example, a model can be defined at the fine-grained level of the individual text of a comment or it may abstract an entire conversation from the content of the messages. Also, it can describe the precise timing when comments are sent/received or it can completely disregard any temporal aspect. In particular, we will focus on models that incorporate a fundamental ingredient of online discussion threads: their reply structure.

Contrary to purely discriminative approaches, generative models are also able to produce instances of the objects of interest, in our case, synthetic instances of discussion threads. Generative models provide more insights and explain better the formation process of online discussion threads than purely descriptive approaches [Mit04].

The behavior of a model depends on its parameters θ, that are adjusted to fit the data. It is important to differentiate between fully data-driven models and those that are largely constrained using prior knowledge. Fully data-driven models usually depend on a large number of parameters and are used as a *black-box*. They are typically trained end-to-end to optimize some measure of predictive performance. Conversely, parsimonious models try to explain phenomena with as few parameters as possible. An example would be a model with a single parameter that is able to generate conversation threads with the same degree distribution than the real ones. Although the latter type of models may perform worse than fully data-driven models in terms of predicting power, they tend to be more interpretable [BC+92] and can thus provide a better understanding of the governing mechanisms of online discussions.

To estimate the model parameters, the most common approach is to optimize a likelihood function, which quantifies how well the model explains the data as a function of its parameters [Mur12]. While this optimization has an analytical solution for very simple models, it can be in general computationally challenging. The complexity of such an optimization depends on the model complexity. Models with a large number of parameters compared to the size of the data may fail to generalize and their predictions may be not valid for new data. In these cases, adding some form of regularization and partitioning the dataset into different subsets for training, validation and testing (cross-validation) helps. In any case, it is necessary to take into account the statistical assumptions in the data generating process. For example, whether data points are independently distributed and under stationary conditions between training and testing conditions, which is often not the case.

A model is expected to be *identifiable*, i.e., as the number of data increases, the true parameter values must converge. For a learned identifiable model, distinct parameter values θ should correspond to distinct models. In contrast, a model is said to be *non-identifiable* when different parameter values result in the model. This can occur, for example, when flat directions exist in the likelihood landscape. As a sanity check, a good strategy to evaluate identifiability is to:

1. generate synthetic data with some parameter values θ^* from the estimation,

2. train the model with those data,

3. evaluate whether the model estimates consistently the same parameter values θ^*.

This could be done for different choices of θ^*. The validation of generative models, in particular, network formation models, typically examines whether the structural properties of generated data are comparable to the original properties of the empirical data.

Typical *structural properties* of discussion threads are:

- *size*: the number of messages,

- *width*: the maximum number of messages at any reply level,

- *depth*: the length of the largest exchange of messages,

- *users*: if the message authorship is known, number of users who authored at least one message.

Usually, statistical tests are used for validation, e.g. the Kolmogorov–Smirnov (KS) test [Kol33, Smi48, AE11], which measures the maximum punctual distance between the empirical cumulative distribution function (CDF) $F_e(x)$ and the CDF of the generated synthetic data $F_g(x)$, defined both on observations x of the structural property of interest:

$$\text{KS-stat} = \sup_x |F_e(x) - F_g(x)|.$$

The use of such tests provides strong statistical evidence favoring a model. Very often, however, these tests can be too strict due to finite-size effects or other artifacts present in the data. In these cases, an alternative qualitative validation, for example, using visualization techniques, can be also satisfactory.

Before presenting each generative model, we should introduce the Galton-Watson branching process for its history and relevance in modeling random trees [WG75]. Indeed, this model is often used as a baseline to compare against other models, e.g. [KMM10]. It starts with a single root node, i.e., the post, and evolves at discrete time-steps. To generate the nodes at time-step $t+1$, each node originated at time-step t generates independently a certain number of children deg according to a fixed probability distribution $p(deg)$. This process is repeated until no new children are generated, i.e., the discussion is over. This is a very simple model that can be estimated very efficiently from the data, since it just requires fitting the empirical distribution $p(deg)$. Although the classical branching process is able to reproduce certain features of online discussion threads such as the degree distribution, it is not a generative model that can explain the mechanisms underlying the dynamics of online discussions. Because, it uses a fixed probability distribution p at each node, it may fail to capture other relevant structural properties such as the depth distribution, and it disregards the authorship and the arrival timestamp of the messages.

Kumar et al. (2010) The limitations of the classical branching process are addressed in *Kumar et al. (2010)* [KMM10] by incorporating the novelty upon the preferential attachment model. That is to say, messages not only attract replies according to the number of previous replies, i.e., degree, but also to their time-stamp. At time-step t, either the thread terminates with some fixed probability $p_f \in (0,1)$ or a new comment is attached to an existing comment k. At time t, the probability of attachment depends on two features: the popularity or degree $\deg_k$, and the novelty, or elapsed time since

k was written, r_k. These features are parametrized by α and τ, respectively. Formally, let the random discrete variable X_t denote the label of the parent node at time t. If a new node is attached, an existing node $k \in [0, \ldots, t]$, is chosen with probability proportional to a linear combination of the two previous features

$$p(X_t = k|\alpha, \tau, p_f) = \frac{\alpha \text{deg}_k + \tau^{r_k}}{\sum_{k'} (\alpha \text{deg}_{k'} + \tau^{r_{k'}}) + p_f}, \tag{2.1}$$

where $\alpha \geq 0$, $\tau \in (0,1)$ and $p_f \in (0,1)$ are $\in R$ and the model parameters to be optimized for a given dataset.

Kumar et al. also propose an authorship model to determine the author of a comment based on the observation that users tend to reply users who had previous replied to their messages. In this model, the author of a new message is selected from the path between this new message and the root node with some probability, and otherwise randomly. However, this model is limited in the sense that the structure and growth of the discussion thread do not depend on the authorship of the messages.

Wang et al. (2012) An alternative framework for modeling the dynamics of online discussions is to consider a continuous-time model. This approach is more convenient when one is interested, for example, in understanding phenomena related to reaction times or lifespan of online conversations. Continuous-time models typically use counting processes as generative models [DVJ03]. Here, we focus on *Wang et al. (2012)* [WYH12], in which commenting behavior is analyzed together with the topic (post) exposure duration to understand user attention to news items. Their generative model is motivated by conflicting observations of previous studies that report significant differences in the probability distribution of thread sizes (the total number of comments of a conversation). While the threads sizes analyzed in some studies followed heavy-tailed distributions [KMM10, GKK11a], other previous studies reported distributions with a light tail [Ogi08, TWdR09]. *Wang et al. (2012)* first observes that the waiting time between two consecutive comments from a user follows a upper truncated Pareto distribution. Based on this observation, they proposed a model that explains the discrepancies of previous studies by means of the topic exposure duration distribution. The growth of attention eventually saturates because the old topics are replaced with newly generated contents. One important assumption is that users share the same microscopic behaviors, i.e., the waiting time for different users comes from the same distribution. In doing so, they are able to model the process of M users as M independent concurrent counting processes. For sites like Digg and Reddit with a short exposure distribution, the predicted distributions of sizes are also light-tailed, whereas in other sites, like Epinions, with longer exposure durations, the obtained sizes are heavy tailed. *Wang et al. (2012)* focuses on reproducing the in-degree distribution of the comments. This is achieved by considering a preferential attachment process. In this model, t denotes the exact time passed since the creation of a topic. Let $\deg_k(t)$ denote the in-degree at time t for a comment k and let p_0 be the fixed probability to comment a post/comment with no replies. The probability of a new comment attaching to comment k is given by

$$\frac{\deg_k(t) + p_0}{(1 + p_0)\gamma M t^{c_0 - c}} \tag{2.2}$$

for constants γ and c and a positive exponent c_0 that measures the combined impacts of factors such as resonance and social influence.

Gómez et al. (2013) This discrete-time model extends previous generative models for discussion threads [KMM10, GKK11a]. Besides popularity and novelty features (parameterized with α and τ respectively), it considers an additional feature (a root bias) that makes explicit the difference between the process of writing to the post (with id 0) node and to a descendant (user comment). This feature is parametrized with β, a positive real number. Instead of having a parameter p_f to terminate the generation of a thread, as in

Kumar et al. (2010), this model generates threads of a given size, drawn from the empirical distribution of a dataset of conversation threads. Formally, the next parent node X_t is chosen according to the following probability:

$$p(X_t = k|\alpha, \tau, \beta) = \frac{\alpha \deg_k + \tau^{r_k} + \beta\delta_{0,k}}{\sum_{k'} \alpha \deg_{k'} + \tau^{r_{k'}} + \beta\delta_{0,k'}} \tag{2.3}$$

where $\delta_{0,k}$ is the Kronecker delta function, i.e., β is a free parameter for the root node, and zero otherwise. The relation between the model of Kumar et al. and this one is made clear by looking at both numerators of equations (2.1) and (2.3).

In *Gómez et al. (2013)*, a model comparison was also done to show the relevance of each feature in every dataset. This statistical test was performed by considering the likelihoods of different reduced models that neglect each of the features separately. To clearly illustrate the differences among model features, we present in Figure 2.4 synthetic threads generated by:

(a) Null model ($\alpha = 0, \ \beta = 0, \ \tau = 1$)
New messages reply to any message in the thread.

(b) Popularity-based model ($\alpha = 2, \ \beta = 0, \ \tau = 1$)
New messages tend to reply to messages with many replies.

(c) Root-bias-based model ($\alpha = 0, \ \beta = 2, \ \tau = 1$)
New messages tend to reply to the initial post.

(d) Novelty-based model ($\alpha = 0, \ \beta = 0, \ \tau = 0.1$)
New messages tend to reply to recent messages.

Backstrom et al. (2013) The next model under consideration, *Backstrom et al. (2013)*, is not strictly speaking a generative model of discussion threads, but proposes how to predict structural properties of a thread (e.g. size) by combining features of different nature. In this model, the representation of discussion threads differs from

(a) $\alpha = 0, \ \beta = 0, \ \tau = 1$ **(b)** $\underline{\boldsymbol{\alpha = 2}}, \ \beta = 0, \ \tau = 1$

(c) $\alpha = 0, \ \underline{\boldsymbol{\beta = 2}}, \ \tau = 1$ **(d)** $\alpha = 0, \ \beta = 0, \ \underline{\boldsymbol{\tau = 0.1}}$

Figure 2.4. Synthetic threads generated with *Gómez et al. (2013)* using different values for the model parameters. Values different from ones of the null model are bold and underlined.

the previous models. Here threads are represented as sequences of arrivals of comments regardless of the reply relationship among them. This decision might be explained by the linear conversation view of the platforms used for the evaluation of the model, e.g. *Facebook*. The model focuses on the authorship of the first comments of the sequence in order to predict, among other purposes, the final size of the thread. Each thread is represented with ρ, a sequence of non-negative integers in which the ρ_t is equal to the number of distinct users arriving to the discussion thread before the author of comment at time-step t ($\rho_t = 0$ if the author wrote the initial post). The data structure λ is then used to assess whether the five possible length-two patterns (0,0),(0,1),(1,0),(1,1),(1,2) have predictive value of the (macro-averaged) thread size.

The predictive model of the thread size is then built using these arrival patterns along with some additional features:

- Social influence: Number of links between the user and users who previously commented, and number of links between the user and the user who published the post.
- Novelty: Elapsed (continuous) time for the first comments to be published.
- Text-based: The occurrence of terms like 'comment', 'agree', etc.
- Miscellany: Number of words, characters, and question/exclamation marks in the comment, and number of links in the post before and after the comment is posted.

Nishi et al. (2016) This model was proposed for reply trees on *Twitter*. The model is motivated by observing that the structure of discussion threads is characterized by some long path-like reply trees, large star-like trees, and long irregular trees. Actually, some of the previous models already denoted long path-like reply trees as 'skinny' in *Kumar et al. (2010)* or 'focused' in *Backstrom et al. (2013)*, and large star-like trees as 'bushy' in *Kumar et al. (2010)* or 'expansionary' *Backstrom et al. (2013)*. Because many of the previous models are based on the branching process model which does not capture appropriately long chains of messages in discussion threads, the depth distribution is often underestimated, as noted in *Kumar et al. (2010)* and *Gómez et al. (2013)*. This last model proves that the branching process model produces unrealistic fractions of long path-like trees or large irregular trees (combination of star-like and path-like structures). Therefore, the authors introduce the concept of segments: maximal chains without branching in a discussion thread. Formally, a segment of length λ is defined by $\lambda + 1$ connected nodes (replies) such that the $\lambda - 1$ inner nodes have in-degree equals to 1. For example, the discussion thread in Figure 2.2 is composed by 5 segments of $\lambda = 1$ $(1-2,\ 1-3,\ 1-8,\ 2-4,\ 2-7)$ and a segment of $\lambda = 3$ $(1-5-6-9)$. Thus, the model adds to the branching process model: (1) the distribution of segment length l, (2) the correlation between λ and the degree of the root, and (3) the correlation between the degree of root and the the degree of the end node of segments. The assessment of this extension shows the ability to capture the fraction of long path-like trees but not large irregular trees. According to the authors, results are explained because a large λ value in one branch implies a relatively high probability of large λ values in other branches. This effect is solved by a final extension which allows λ to be correlated among segments starting from the same node.

Lumbreras (2016) This generative model is part of a doctoral thesis about automatic role detection in online forums [Lum16]. It is motivated by observing that the growth of discussion threads in previous generative models, in particular *Kumar et al. (2010)* and *Gómez et al. (2013)*, is irrespective of the user who is writing a new message. This new model proposes that there might exist different roles which categorize users who participate in the discussion threads. To this end, the model builds upon *Gómez et al. (2013)* and introduces latent types of users or roles. In this model, a role u, $u = 1, \ldots, U$, corresponds to specific values $\theta_u = (\alpha_u, \beta_u, \tau_u)$ associated to the popularity, root-bias and novelty influence, respectively, of a type of user. Let z_k denote a binary vector of U entries indicating the role membership of the author of the k-th comment, i.e., $z_{ku} = 1$ if author of comment k belongs to role u, and zero otherwise. This is the latent variable not present in the data. Let q_u denote its marginal distribution, i.e., $p(z_{ku} = 1) = q_u$ with $q_u \geq 0$ and $\sum_{u=1}^{U} q_u = 1$. In this model, the next parent node X_t is chosen according to the following joint probability:

$$p(X_t = k, z_k|\theta) = \prod_{u=1}^{U} q_u^{z_{ku}} p(X_t = k|\theta_u)^{z_{ku}}, \qquad (2.4)$$

where $p(X_t = k|\theta_u)$ is the same as equation (2.3). The existence of the latent variables z prevents to optimize a complete likelihood function defined using equation (2.4). Therefore, the expectation-maximization algorithm is used as an optimization procedure. The number of roles U (model selection) is computed using Bayesian Information Criteria. Roles are finally used to analyze their predictive power, i.e., the capability of this model to predict the parent message of arriving messages in comparison to *Gómez et al. (2013)*, and two minimal models: one based on popularity [BA99] and the other on novelty.

To sum up, we present their main characteristics in Table 2.1. We observe the heterogeneity of these approaches in relation to features (popularity, novelty, reciprocity, root-bias, arrival patterns, text expressions, social influence, segment lengths, user roles), structure of threads (tree, array), temporal dimension (discrete, continuous), and the structural properties for validation (size, depth, degree, shapes). Also, the evaluation of the models is done with real data from online discussion platforms of very diverse nature: online forums (*Y! Groups*, *Usenet*), social news (*Slashdot*, *Barrapunto*, *Digg*, *Reddit*, *Menéame*), peer production (*Wikipedia*), social networks (*Facebook*, *Google Plus*), and microblogging services (*Twitter*).

2.4 Dynamics of online petition signing

The right of petition is a citizen right with several centuries of history [Bro99, Fox12]. However, online platforms have led to a strengthening of this political practice since they substantially reduce the effort required to propose and sign petitions, as well as to disseminate them through social media channels. In addition to this, the digitization of petitioning also provides the possibility to analyze digital traces generated by proposers and signatories and, therefore, to characterize the explaining factors for the success of online petition signing.

Table 2.1. Main characteristics of the generative models of online discussion threads: the features, whether the predicted thread is a tree-like structure, whether threads grow in discrete or continuous time, and the datasets and structural properties used for the parameter estimation and the validation of the model.

Model	Ref.	Features	Structure	Time	Datasets	Str. properties
Kumar et al. (2010)	[KMM10]	Popularity, Novelty, Reciprocity	Tree	Discrete	Y! Groups, Usenet, Twitter	Size, Depth, Degree
Wang et al. (2012)	[WYH12]	Popularity	Tree	Continuous	Digg, Reddit, Epinions	Size
Gómez et al. (2013)	[GKLK13]	Popularity, Novelty, Root-bias	Tree	Discrete	Slashdot, Barrapunto, Wikipedia, Menéame	Size, Depth, Degree
Backstrom et al. (2013)	[BKLDNM13]	Novelty, Arrival patterns, Text expressions, Social influence	Array	Continuous	Facebook, Google+, Wikipedia	Size
Nishi et al. (2016)	[NTO$^+$16]	Popularity, Segment lengths	Tree	Discrete	Twitter	Size, Depth, Shapes
Lumbreras et al. (2016)	[Lum16]	Popularity, Novelty, Root-bias, User Role	Tree	Discrete	Reddit	Size, Depth, Degree

One of the first studies examining the dynamics of petitions focused on the online platform of the German Parliament and found that signing activity was concentrated in a small group of petitions [JJ10]. Interestingly, some other petitions benefited from coinciding in time with these popular ones. A later analysis of petitions from the same platform also analyzed this spill-over effect and confirmed that petitions coinciding with a successful petition were able to obtain almost twice as many signatures per day than the others [SJ14]. The temporal growth of petition signing was examined for the first time in a study of the UK Government platform [HMY13]. That work considered diverse factors to affect growth (e.g. the topical category of the petition, the day of the week in which it was published, the number of petitions started on the publication day) and revealed that the number of signatures a petition gathers on its first day was the most significant. The temporal dynamics of online petitions have been also characterized by focusing on the elapsed time between successive signatures to a same petition [BWMB17]. Results of that study, based on data from *openPetition*, found that petitions with many signatures are less bursty than ones with few signatures.

Some other approaches to predict petition growth and success have considered linguistic features. A study applied named entity recognition and topic modeling to the message of petitions from the White House's *We the People* platform [HHU$^+$15]. Their findings were that informativeness, named location, and several topics significantly correlated with the final number of signatures. That work was then extended to discover that petition popularity is negatively associated with language extremity and with the occurrence of many names, while popular petitions usually contain topics familiar to the public or about important social events [HHU$^+$16]. In contrast to these observations, a recent work analyzed petitions from the *Russian Public Initiative* project (including pro and against voting data) and found that informativeness was a significant predictor only for votes "against" [Por18]. Topic modeling has also been applied on a dataset of petitions from *Change.org* to find topical niches and to suggest that, although petitions have to compete over limited resources of signatories, specialized or focused petitions do not perform better in concentrated topics [TSH17]. This finding is consistent with other study of the Finnish platform *Adressit.com* that concluded that the more specific a petition is, the fewer signatures it receives [Ber17]. Other works have applied sentiment analysis techniques to validate that *Change.org* petitions with positive emotional tone tend to be more successful than others with heavy cognitive reasoning emphasizing moral judgment [EDK16, CDK$^+$19].

The role of social media in disseminating online petitions has been the focus of increasing interest from academia. A first work that included interviews with creators of successful petitions showed that many petitioners used social networking and discussion forums [Wri12]. Another study examined discussions on *Twitter* about online petitions from *We the People* [DLH$^+$15]. Authors of that study then extended their methodology to focus on the campaign on *Twitter* for a petition from *Change.org* and concluded that tweeting and certain forms of online media are associated with the willingness to sign online petitions [HDD$^+$17]. In a similar way, a study of petitions from *Care2* confirmed that the number of users posting about the petition and the number of retweets correlated positively with petition success [PACM16]. Nevertheless, high activity on social media is not always an indicator of gathering many signatures as shown in a recent analysis of online petitions from the UK Parliament platform [ABS17]. The ease of signing petitions in non-governmental platforms (e.g., *Change.org*, *Care2*) versus the greater difficulty on government platforms which often require a verified citizen account might explain these differences.

2.4.1 Models of online petition signing

Just as there are models to simulate the growth of online discussion threads, some models have also proposed in recent years to provide a better understanding of the governing mechanisms of online petition signing. We review below three models whose main features are summarized in the Table 2.2.

Chan et al. (2017) This model was designed to infer virality and diffusion structure from online petition signing in the White House's *We the People* platform [CLHD17]. The notion of virality is inspired by a previous approach that analyzed how items propagate on *Twitter* [GAHW15]. Because the diffusion network of a petition from *We the People* is not available, this work first detects peaks of activity to infer events and to then measure virality (similar to the notion of popularity in online discussion threads) as the excess of signatures on the second day with respect to the first day. This allows them to discover that (1) petitions with later global peak days obtain on average more signatures than petitions with earlier peaks, and (2) successful petitions were more likely to obtain more signatures on their second day than on their first day. The model is based upon these observations. At each time step, a petition has a small probability of attracting a number of users drawn from a log-normal distribution. In parallel, to incorporate virality, each new signer has a small probability of spreading the petition to other new users which is parametrized by the basic reproduction number (R_0).

Yasseri et al. (2017) This model is motivated by a previous analysis of the daily growth of petitions from the UK Government online platform [HMY13]. Because that study found the number of signatures on the first day of a petition as the most significant factor, the model is an adaptation of a very well-known model of novelty and collective attention that was originally proposed for the popularity of Digg stories [WH07]. The idea behind this approach is to consider the rapid rise and decay in petition signing by combining a multiplicative process and an outreach factor that decays very fast. The number of signatures of a petition p after t hours since it was published is expressed as:

$$s_p(t) = s_p(t-1)(1 + \mu_i r(t-1)) = s_p(0) + e^{\mu_i \sum_{t'=0}^{t'=t-1)r(t')}} \tag{2.5}$$

where μ_i is the multiplication factor specific of each petition and $r(t)$ the decay factor that is fixed for all petitions:

$$r(t) = \frac{\mathrm{E}[\log(s_p(t))] - \mathrm{E}[\log(s_p(t-1))]}{\mathrm{E}[\log(s_p(t))]} \tag{2.6}$$

Therefore, the multiplication and decay factors in this model correspond to the parametrization of popularity and novelty features respectively.

Proskurnia et al. (2017) The analysis of *Care2* petitions on *Twitter* reviewed above [PACM16] was immediately extended to build a predictive model of the success of online petitions leveraging multidimensional time-series [PGK+17]. This model takes into account different temporal factors. First, it considers the rise and decay of petitions but also includes circadian rhythms. In particular, the number of signatures of a petition p after t hours is expressed as:

$$s_p(t) = \left\{ a_p + b_p sin(\frac{2\pi}{T}(t + \phi_p)) \right\} t^{k_p} e^{-t/\tau_p} \tag{2.7}$$

where a_p is the intensity, b_p is the amplitude of the oscillation, ϕ_p its phase (cycle of $T = 24h$), τ_p is the novelty decay parameter, and k_p is the initial rise in the signing activity.

More advanced features are then added to the model: popularity (denoted as self-excitation), and external influence which considers the diffusion of the petition in social media and the exposure of the petition in the rank of top 10 petitions in the homepage of the platform. To do this, Equation 2.7 is extended as follows:

$$s'_p(t) = s_p(t) + \sum_{i=0}^{T_{mem}} (c_{self}(i) s_p(t-i) + c_{sm}(i) n_{sm}(t-i) + \frac{c_{front}(i)}{n_{srank}(t-i)}) \tag{2.8}$$

where $n_{sm}(t)$ is the number of tweets multiplied by the average number of the authors' followers, $n_{srank}(t)$ is the position at the rank in the homepage (if not featured, the value is 1000), T_{mem} is the size of a memory window of 10 hours, and c_{self}, c_{sm}, c_{front} express importance of self-excitation, social media influence, and ranking influence, respectively.

The main difference between this model and the two previous ones is its deterministic nature. However, the authors suggest as future work to check its predictive power if a stochastic Hawkes process were incorporated.

Table 2.2. Main characteristics of the models of online petition signing: the features, the temporal unit, and the datasets used for the parameter estimation and the validation of the model.

Model	Ref.	Features	T. unit	Datasets
Chan et al. (2017)	[CLHD17]	Popularity	Signature	We the People
Yasseri et al. (2017)	[YHM17]	Popularity Novelty	Hour	UK Government We the People
Proskurnia et al. (2017)	[PGK^{+}17]	Novelty Seasonality Popularity External effects	Hour	Care2

2.5 Platform interventions

The dynamics of participation in any online platform are affected by many and diverse factors. An essential element is platform design since it defines and shapes the forms of participation that technologies provide. In fact, the design of interfaces between users and technologies is the basis of human-computer interaction, a research discipline at the intersection of design, computer science, behavioral sciences, media studies and other research fields, that has received widespread attention in the last decades [CNM83, Jac12].

It is nearly impossible to find online platforms today that have not undergone significant changes since their first release. Algorithms and interfaces are frequently modified in order to improve the affordances of online participation platforms. To evaluate performance and to guide the development process, much research has been carried out revealing relevant findings about the effects of platform interventions of very different kind. For example, interventions in field experiments could consist of actions carried out by volunteers who set social norms in online discussions [Mat19a] or who moderate comments [Mat19b]. However, the focus of this thesis are interventions related to the deployment, modification or removal of a platform technical feature. We present below a review of illustrative works by distinguishing between (1) experimental approaches based on controlled environments, and (2) observational approaches based on uncontrolled environments, as the ones of this dissertation.

2.5.1 Experimental studies

The design of experiments is a systematic and rigorous approach to validate an hypothesis, in this context, to validate and measure the expected effects of an intervention in an online platform. Among the different works about how to design controlled experiments on the web [Eis05, McF12, KDLX14, XCF$^+$15, CK15], a prominent reference is the survey and practical guide elaborated by the Analysis and Experimentation team at *Microsoft* [KLSH09]. This resource describes techniques like A/B testing (see Figure 2.5), best practices and practical lessons and it is the basis of successive research works with relevant findings on how to increase the trustworthiness and the scale of online experiments [KDF$^+$12, KDF$^+$13], and the discussion of rules of thumb and pitfalls for experimentation in web environments [KDLX14, DFG$^+$16]. Another relevant example of designing and deploying online experiments is the work done at the social media platform *Facebook*, which released an open source framework for online field experiments [BEB14]. This research project has also contributed with findings of relevance on diverse topics: the role of social networks in information diffusion [BRMA12], the importance of social influence for advertising [BEYR12] and political mobilization [JBB+17], or the estimation of peer effects, i.e., how individual behaviour is affected by the behavior of others [EKB16]. In relation to this, it is worth mentioning that research done by *Facebook* has been heavily criticized by many sectors of academia and society when they carried out a massive experiment manipulating the emotional expressions that users received in their news feed [KGH14].

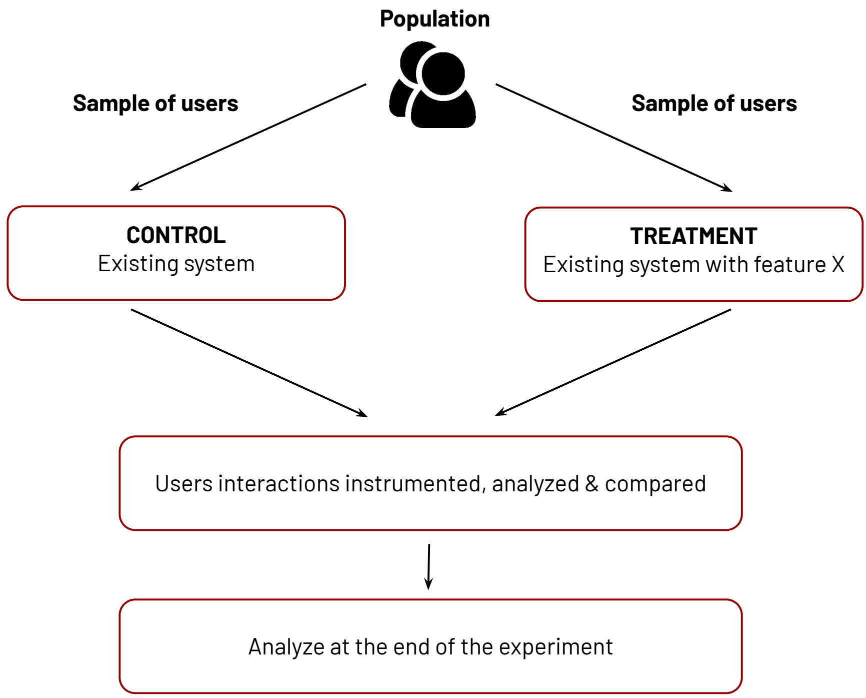


Figure 2.5. High-level flow for A/B test. Adapted from: [KDF+12]

Experimentation with online discussion platforms Several experiments with participants that evaluate platform features are reviewed in a survey of research and empirical exploration of student contribution in asynchronous online discussion [HCN10]. These experiments were proven effective in detecting missing features like the inability to flip back and read through discussion postings while composing a message or to edit and delete messages [MC04], as well as technical problems related to the login [HTB+05]. Indeed, there is much interest in experimenting with online discussion platforms in the context of e-learning, e.g., to empirically validate that forum experience improves by introducing a reputation system within the platform [CFHH14]. In an experiment about argumentation in threaded discussions, participants were split into three groups: the first one posted their messages using response categories and labels, the second one only response categories, and the third one was not provided with these features while posting messages [JJ04]. The study found that message labeling reduced the proportion of arguments, i.e., the proportion of arguments elaborated with explanations. Some other experiments have focused on the way discussions are presented to participants. The adoption of the typical threaded interface of online forums was examined in an experimental chat to reveal that users' patterns of interaction are equally effective, but different (and possibly more efficient) than interactions in standard chat programs [SCB00].

Experimentation with online petition platforms First works based on field experiments relied on simulated environments. An early experiment with participants recruited from Amazon Mechanical Turk examined whether signing an online petition increased or decreased subsequent contribution to a charity [LH13]. Results revealed that participants who signed the online petition were significantly more eager to to make a donation to a related charity, while participants who did not sign the petition usually donated more money to a charity not related to the petition. Another example is the a controlled experimental setup to analyze whether the personality of petition starters and followers might predict signing activity [MJHR15]. The experiment, inspired by models of collective behaviour [Sch78, Gra78, GS83] and the Big Five personality traits [Wig96], served to find a significant association between willingness to start with both extraversion and internal locus of control, while the tendency to follow was associated with agreeableness. To the best of our knowledge, the only three works reporting experiments in a real online petition platform were carried out in collaboration with *Change.org*. Two of these experiments were setup to confirm that early growth of online petitions increases the attractiveness of new signatories [vdRKRP14, vdRAWF16]. The third experiment was designed to examine the relevance of emergent processes to identify the mechanisms involved [VTA+15]. Results indicated that petition growth was highly erratic since many revivals were impossible to be predicted and the authors of the study concluded that the dynamics of online petitioning were affected by processes of "accidental activation".

2.5.2 Observational studies

Although controlled experimentation is the most suitable methodology to measure the impact of a platforms intervention, this setup is impracticable in many cases. First, most of the experiments presented above relied on laboratory environments recruiting small groups of participants. Therefore, user behavior may be conditioned since they do not correspond to a real online participation scenario. Large-scale experiments with users in real-time, as the ones carried out by *Microsoft* [KDF+13] or *Facebook* [BEB14], require a privilege out of reach to most researchers: to have control of the platform to be intervened. Furthermore, interventions of platform features, like the manipulation of the news feed on *Facebook* to study emotional contagion [KGH14], might violates ethical research standards. These standards are particularly important for civic technologies since platforms are intended to facilitate political participation. Therefore, studies based on observational data collected from online platforms are becoming increasingly popular in computational social science [OKC18, KS19].

While causality in experimental approaches relies on the random assignment of participants to a control group and a treatment group, this cannot be done within observational approaches. Nevertheless, some empirical studies consist on scenarios in which individuals are exposed to the experimental and control conditions by nature, i.e., without the control of researchers. These approaches, denoted as quasi-experiments or natural experiments, become a suitable methodology for assessing interventions when the platform is not controlled or they violate research ethics. For instance, external researchers were able to run a natural experiment on *Facebook* to observe how the introduction of the “People You May Know” recommendation feature drastically affected the number of links and triangles of the friendship network [ZGR$^+$14]. This example inspired a later work which proposed the notion of platform effects as the way “the design and technical features of a given platform constrain, distort, and shape user behavior on that platform” [MP16] and served to raise critical questions when conducting observational studies with social media data [MMLMB19]. To mitigate the typical problems of observational studies like confounding factors or selection bias, that study of platform effects relied on applying regression discontinuity design [LL10]. As shown in Figure 2.6, this technique consists of measuring the local average treatment effect on an outcome variable by assigning a cutoff above or below which the intervention is assigned. Other popular approach in observational studies is propensity score matching [RR83]. This technique predicts the probability that an item will belongs to the control or treatment group using logistic regression on confounding covariates and then uses each probability as a score to create a counterfactual group by matching similar pairs of control/treatment items (see Figure 2.7).

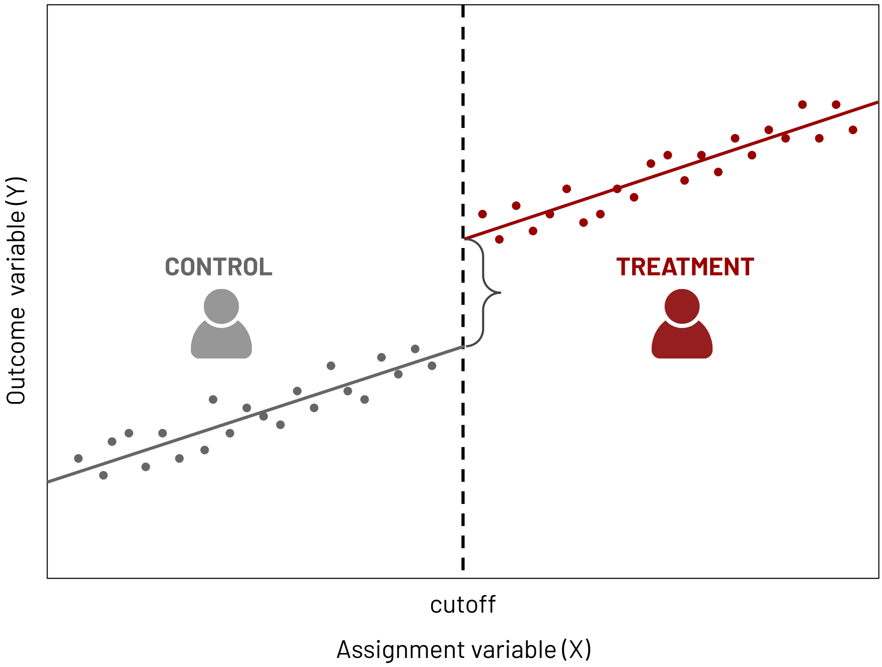


Figure 2.6. Regression Discontinuity Design. Adapted from: [LL10]

Observing online discussion platforms Some observational studies have focused on how conversations are affected by the length allowed by specific online discussion platforms. For instance, discussions in *Facebook* were affected when the size of the reply window decreased resulting in users posting shorter replies, faster, and more frequently [Sel11]. A similar example is the change done in 2017 on *Twitter* when extending the 140 character limit 280 characters [Ros17]. An observational study of this intervention focused on tweets with nearly or exactly 140 characters before and after the change [GAW18]. It was found that tweets before the change tend to be more successful than similar-length tweets after the constraint, which suggests that the length constraint improved quality of mes-

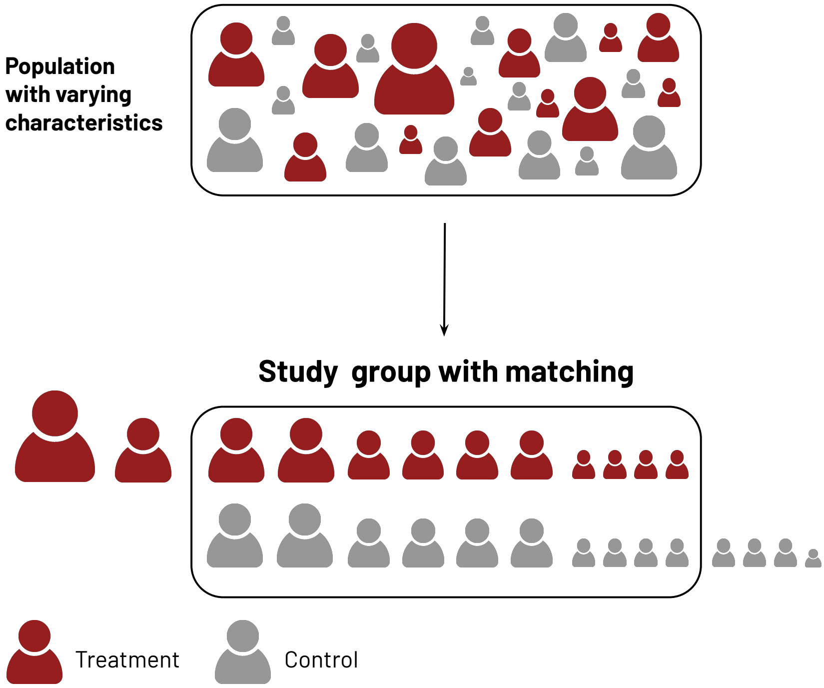


Figure 2.7. Propensity Score Matching. Adapted from: [Sum]

sages on *Twitter*. Another platform feature in online forums which has been observed is anonymity, with the aim of assessing whether it affects the quality of discussions. Different natural experiments carried out in the Huffington Post forums [FMN15], a South Korean forum [Cho13], and a platform of the US Army [KH05] indicated that removing anonymity produced more civil discussions and reduced anti-social behavior. However, as reported by [SDCC19], there is no consensus on whether anonymity favors or harm participation, given other research works that found anonymity promoting participation in online communities, e.g., [SK86, BMHH$^+$11, Pap04, KBK13].

Observing online petition platforms To the best of our knowledge, the only observational study about online petitions is the natural experiment carried out with data from the UK government platform before and after the introduction of a list of trending petitions in the homepage [HJMY18]. This intervention was expected to affect the dynamics of petition signing. Results showed that, although the daily activity does not increase, the distribution of signatures across petitions became significantly more unequal. This effect varied across petitions trending at different ranks on the homepage and among different profiles of users.

Part I

Analysis of Online Discussions

3 Detection of Platform Effects

3.1 Introduction

Nowadays, millions of citizens interchange messages in online discussion platforms. A large part of these discussions are related to political talk which is attracting an increasing number of citizens to go online to engage in political processes [Bim03, Cha06]. This might be explained by the potential of the Internet to create a new public space for political discussion [Pap]. Thus, information and communication technologies have been noted to facilitate the participation of citizens in democratic communication [HVD00] and, ultimately, the construction of an online deliberative public sphere [Dah]. Public sphere, seen as "a society engaged in critical public debate" [HHM91], is the essence of deliberative processes. Although the definition and implications of deliberation are far from conclusive [CCJ04], the many approaches to deliberation [Bar84, Hab96, Els98, Fis97, Cha03], have all a common denominator: the relevance of communication in detriment of direct voting. In this regard, public sphere and deliberation are influential concepts in the relationship of democracy and information communication technologies [Cha09].

Despite an important fraction of research indicating the benefits of online public spheres, some other studies have adopted a more critical position regarding the potential of the Internet in facilitating deliberative processes. Early work on comparing face-to-face and online deliberation reported considerable resemblance between both types [LFI04, Min, GSaH]. However, some recent analyses have found that discussions on the Internet generate more negative emotions and, therefore, consensus is less likely to be obtained [BWDC12]. The lack of consensus is commonly associated with scenarios of group polarization, which commonly occur in online discussion platforms [ABA$^+$96, Sun01, Sun02]. In this regard, uncivil attitudes in online discussions, which are contrary to deliberation by rational-critical discourse [Hab84], have been proven to play a major role in promoting polarized scenarios [ABS$^+$14]. Given that interactions between individuals are not always civil and rational [Wil00], some researchers have concluded that discussions on the Internet do not necessarily lead to online deliberation [Dah05].

The contradictions between findings from online deliberation analyses have motivated the examination of which features in online platforms might affect their deliberative potential. A feature that has received large attention is the moderation of messages. Although online moderation can be seen as a form of censorship and a threat to freedom of speech, some studies have defended that moderation by skilled users is a relevant feature to promote deliberation [CG01, Edw02, WS07, Wri09]. Another feature of interest is the anonymity of users. On the one hand, this feature is likely to improve online discussions because users feel no pressure of conventional cultural cues [Kim06, BT03]. On the other hand, another study indicated that this lack of pressure is precisely the reason behind the emergence of uncivil and non rational attitudes[FKH00]. The type of discussion, i.e., synchronous (e.g., chats) versus asynchronous (e.g., online forums), has also been examined and results often indicate that asynchronous discussions better promote deliberation [JK05, SJSN09]. Finally, online deliberation might be also conditioned by the topics under discussion. Political discussions in Slashdot, an online discussion platform which has been defined as "a form of online public sphere" [Poo05], were found much more deliberative than discussions of other topics, e.g., online gaming [GBKB10].

In general, most studies of online deliberation have examined one or a few features in one single online platform and, therefore, results are limited to individual characteristics of the online community and the platform itself. This research gap has been recently addressed by comparing different technical features (e.g., moderation, synchrony of discussions) in a news forum, three news websites, and *Facebook* news pages [EFE17]. Their results show that while moderation has a positive effect on online deliberation, this was not found for asynchronous discussions. That study, as many others, is focused on a subset of potentially relevant features while others are not considered, e.g., anonymity. More importantly, there could be events at a specific moment in time which produce durable effects on deliberation, e.g., the deployment/change of technical features or the emergence of new topics. Given that, to our best knowledge, previous research on online deliberation has not considered the effect of events of this nature, we aim to answer the following research question:

- *Is it possible to automatically detect events which affect online deliberation in online discussions?*

To answer this question, we have chosen an online discussion platform that we suspect to have been affected by specific events. In particular, we have collected the discussion threads over five years from *Menéame*, the most popular Spanish social news site. Two candidate events to have affected online deliberation are:

- **E1: The 15M movement.**
  In May 2011, the 15M grassroots movement (also known as the Indignados movement) occupied the main squares of the largest cities of Spain in order to advocate for a *real* democracy. This movement has made a significant impact on Spanish politics.

For instance, grassroots parties which emerged from the 15M movement [AVLK16], like Barcelona en Comú, Ahora Madrid and Zaragoza en Común, are currently ruling the city councils of many of these cities. The origin of the 15M movement is explained in [TCLM+15] by the emergence of technopolitical practices, many of which occurred in *Menéame*. Some other studies stated that *Menéame* played an essential role in the difussion of the call for the initial demonstration [OS12] and, furthermore, the construction of an online space that generated many of the claims and messages adopted by the 15M movement [Men11]. These effects were confirmed in [Pos14] which proved that aggregators and link recommendation sites, especially *Menéame*, experienced unprecedented traffic growth during the 15M movement. Therefore, although many links in early years were related to science and technology, the eruption of the 15M movement turned *Menéame* into one of the most relevant online discussion platforms in Spain about social and political issues.

- **E2: Change of the conversation view.**
  Since the first version of *Menéame*, directly inspired by *Digg*, many changes have taken place. Regarding features of online discussion, we highlight the change of the *conversation view*, i.e., the way in which the discussion threads are presented. The original conversation view of *Menéame* displayed the comments of a thread linearly in a chronological order, regardless of reply relationships. In January 2015, this design changed and, by default, messages are now displayed hierarchically following the tree structure of the discussion thread. Figure 3.1 shows both interfaces: a thread from 2011 about the rise of the 15M movement[1] presented in a *linear* conversation view (Figure 3.1a),

[1] https://www.meneame.net/story/junta-electoral-madrid-prohibe-concentracion-convocada-acampada

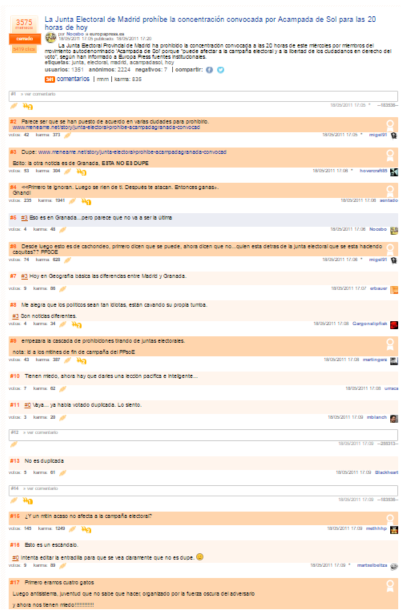
and a thread from 2015 about the victory of the grassroots party Barcelona en Comú in the local elections[2] presented in a *hierarchical* conversation view (Figure 3.1b). Comments posted by the story's author are blue, comments scored negatively are white (text is hidden unless clicked upon), and the rest of the comments are orange using intensity to indicate the voting score (i.e., the better, the darker).

These two candidate events are motivated by different reasons. For E1, different studies confirmed that *Menéame* played a key role in the communication dynamics of the 15M movement [TCLM+15, OS12, Men11, Pos14]. This resulted in a great increase of political talk within the platform. Given that discussion threads about politics in a similar platform exhibited higher level of deliberation [GBKB10], the increase of political talk might have affected deliberation in *Menéame*. For E2, hierarchical conversation views are the typical interfaces of asynchronous discussions, which better promote deliberation [JK05, SJSN09]. Furthermore, this type of view has been proven useful to improve different components of communication, e.g., construction of knowledge [McV07], context of the discussion [FPDL06, VN03], and coherence [SCB00]. Given that communication is the essence of deliberative processes, the change of conversation view from linear to hierarchical might have also affected deliberation in the platform.

As we detail in the next section, most studies on online deliberation have examined the principles of rational-critical discourse [Hab84, HHM91] with a specific coding scheme, e.g. [Tré04, SG07, CHH02, FE15]. Such approaches have always relied on the human examination of linguistic features. On the one hand, these classical approaches benefited from the reliability of using human coders. On the other hand, their cost is unaffordable for large datasets as the one of *Menéame*. For this reason, we will measure online deliberation using the computational model in [GBKB10], which is based on the structural complexity of discussion threads.

[2]https://www.meneame.net/story/ada-colau-gana-barcelona

We should remark that the main objective of this study is to automatically detect events that significantly produced durable effects on online deliberation. Although we suggest two candidate events, the statistical methods of this study will examine every other possible moment in time as a possible event. Therefore, our methodology is not limited to these two events.



(a) Linear.

(b) Hierarchical.

Figure 3.1. Comparison between (a) the linear conversation view of a discussion thread from 2011 about the 15M grassroots movement occupation of Puerta del Sol Square in Madrid, and (b) the hierarchical conversation view of a discussion thread from 2015 about the victory of the grassroots party Barcelona en Comú in the local elections.

3.2 Related work

We now present previous research that motivates our methodology based on platform effects [MP16] and structural indicators of online deliberation [GBKB10].

3.2.1 Impact of events in online platforms

Previous work has examined how the activity of online platforms is affected by the arrival of different types of events. We will distinguish between events originated by (a) the emergence of new topics under discussion, and (b) the deployment of new features in the platform.

The impact of the first type of events has been analyzed in different social media platforms. An epidemic model was defined in [CS08] to prove that viewing activity on *YouTube* can be explained by different factors, e.g., new popular topics. A later study also found that the popularity of videos can be predicted by, among other factors, the occurrence of external events; e.g., the video being massively posted in other online social networks and blogs [FBA11]. The impact of similar events, defined with data from *Google Trends*, was also analyzed in *Wikipedia*, concluding that trending topics notably affect the popularity of articles [RFM10]. Moreover, the impact of trending topics has also received much attention in the context of microblogging services, in particular, *Twitter*. The factors defined in [CS08] were proven effective to characterize trending topics [RFM10]. The impact of this type of events was also found to influence the usage of mechanisms on *Twitter*; i.e., the average number of URLs and hashtags within the tweets [NBG11]. Moreover, as shown in [LGRC12], peaks of activity originated by trending topics can also provide a semantic characterization of the discussions.

While the above type of events mostly corresponded to new topics under discussion, activity in online platforms can be also affected by technical changes the platform itself. Indeed, the measurement of interventions in online platforms is a well-studied issue in software development through A/B testing [KL17], which requires to control the change and deployment of technical features, a condition that cannot be assumed by external researchers. This is the motivation of a recent approach for causal inference using Bayesian structural time-series models [bro]. However, this approach requires, at least, two time series: one affected by the platform intervention, and another in which the intervention made no effect in order to construct a counterfactual. If there is no access to counterfactual information, an alternative approach is the use of experimental pretest-posttest design, which expects to infer the treatment effect of an intervention through regression discontinuity design on a time series. This has been proposed recently to measure platform effects, i.e., "the design and technical features of a given platform which constrain, distort, and shape user behavior on that platform" [MP16]. Thus, this methodology can be used used to detect events (the deployment of new features in the platform and also the emergence of new topics under discussion) that might affect indicators of online deliberation.

3.2.2 Measurement of online deliberation

The extent to which online tools enhance the quality of discussion and decision-making has attracted increasing interest from researchers and practitioners [DG09]. Different studies have analyzed how online deliberation occurs in online discussion platforms of diverse nature, e.g., Usenet newsgroups [Wil98], online forums [CHH02], media sites [ZCP08, EFE17], and online social networks [HG13].

To measure deliberation in online discussion platforms many different approaches have been proposed. The ideal requirements summarized in [Dah01b] to facilitate online deliberation are exchange and critique of reasoned moral-practical, validity claims, reflexivity, ideal role taking, sincerity, discursive inclusion and equality, and autonomy from state and economic power. The coding scheme defined in [Tré04] to measure the deliberativeness of online discussions using eight dimensions: equality, rationality, respect, constructiveness, interactivity, personal experience, emotional balance, and reflexiveness. These dimensions are similar to the ones from the coding scheme in [SG07]: reasoned opinion expression, disagreement, equality, topic, and engagement. As one could observe, many dimensions of these methodologies are essentially features of rational-critical discourse in consistency with the conceptualization of "public sphere" defined in [HHM91]. This observation is even explicit in other schemes, e.g., the model in [GW03], and was already observed in [CHH02]: "most researchers of online deliberation have opted to use content analysis as a means of measuring the quality of discussion, operationalizing their own conceptions of what good communication looks like". However, we should note that measuring online deliberation with content analysis has always relied on the examination of online discussions by trained human coders. Therefore, these approaches are unfeasible in large datasets. Some recent methodologies to measure online deliberation are including features that can be automatically extracted or inferred from datasets. A deliberative analysis of Wikipedia concluded that the network structures of different groups could be useful in quantifying features like equality, influence, and group member roles [BWCD11]. The empirical model proposed in [FE15] for the analysis of online deliberation uses three levels: input, throughput, and outcomes. Some of the dimensions of these levels can be automatically inferred from the texts (e.g., emotional talk with computational sentiment analysis) while some other dimensions still require the intervention of human coders (e.g., civility and constructiveness).

To the best of our knowledge, the only model to automatically measure online deliberation is the one presented by [GBKB10]. This approach is based on a previous Madisonian model by [AF04] in which deliberative processes are categorized in two dimensions: representation and argumentation. The model quantifies online deliberation without examining content features. In particular, the model in [GBKB10] uses network indicators based on the network topology of online discussions, i.e., the more complex the discussion threads, the greater the level of deliberation.

3.3 Dataset from *Menéame*

The analysis of this chapter uses data from online discussions in *Menéame*, the most popular Spanish social news website (130^{th} most visited domain in Spain according to *Alexa*[3]). This platform, developed in 2005, includes typical features of social news sites (e.g., *Digg*, *Slashdot*, *Reddit*) such as social bookmarking, blogging, and web syndication. Indeed, the developers of *Menéame* acknowledge *Digg* as an inspiration of the first version of the platform[4], and aimed to provide a similar service for the Spanish blogosphere.

The functioning of *Menéame* is broadly as follows. Users are able to submit links to blog posts / news (hereafter *stories*) that will appear in a queue of pending stories. Then users vote and discuss each story in a discussion thread in order to promote the most interesting stories to the front page of the platform. The selection of stories for the front page is done by an open source collaborative filtering algorithm based on multiple criteria, e.g., the voting score of the story, and the reputation index of the users who have voted the story.

The collaborative nature of this platform has several social and political implications, as observed in previous studies. As shown in [Tri10], many media outlets in Spain includes a *Menéame* sharing widget which illustrates the relevance of this platform in Spanish online media. Another study found that, although *El País* (the most visited media outlet in Spain) was the media source with most submitted stories to *Menéame*, other media outlets exhibited a greater impact within the platform [MO$^+$09]. Thus, the social design of *Menéame* allows users to build a social and collaborative agenda-setting opposed to the notion of agenda-setting of traditional media [Men11]. The true value of discussion and collaborative filtering in *Menéame* has been suggested to be the possibility to build a space of debate [FR09]. Indeed, it was that 67.6 percent of users said that they use *Menéame* not only to read stories but also to par-

[3] `http://www.alexa.com/siteinfo/meneame.net` (accessed Feb 6, 2017)
[4] `https://www.meneame.net/faq-es`

ticipate in the discussion threads, while 31.5 percent of users were only interested in reading stories [FR11]. The implicit social network of user interactions through comments has been investigated in [KGRDQV11] and the heterogeneity of user behavior in *Menéame* was also analyzed in [Mar15] from an ethnographic perspective. In particular, *Menéame* has been described as a virtual community that has developed a particular cyber-culture based on social structures and their own code of practices [Mar15]. Therefore, the development of this inner culture might be the result of the response of the online community to different events.

To generate the dataset of this study, we run a crawling process that collects all the stories in the front page of *Menéame* from 2011 to 2015 (both years included). We then perform a second crawling process to collect every comment from the discussion thread of each story. From both crawling processes we obtain 72,005 stories and 5,385,324 comments. For each of them, we keep associated metadata such as the id, url, user name, time-stamp, text message and received votes.

Finally, we should remark that messages in discussion threads of *Menéame* have to be posted as replies to either the story or another reply. For each message, the two conversation views of *Menéame always* indicate the id of the message being replied to (see Figure 3.2). Therefore, to automatically generate the tree structure of each discussion thread, we also collect the parent id of every comment to comment.

3.4 Analysis of platform effects

To better understand the activity in *Menéame* between 2011 and 2015, we first make a preliminary exploration of our dataset. Then, we present our statistical approach to detect events that have affected online deliberation in *Menéame*. Finally, we describe the results of the analysis.

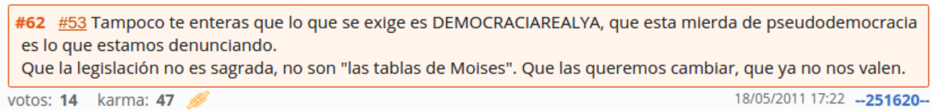


Figure 3.2. Example of how a comment (#62) replies a previous comment (#53). Comments to another comment are identical in both linear and hierarchical views. Thus, every comment starts with the symbol # followed by the id. If the comment is replying to another comment, not to the story, it automatically shows the symbol # followed by the id of the comment being replied to.

3.4.1 Preliminary exploration

We first analyze the posting and voting activity. Figure 3.3 presents a scatter plot of the number of stories and the number of votes to stories for every day in the dataset. As one could expect, the plot shows a strong correlation between both variables ($R = 0.821$). Nevertheless, we detect that some days (red markers) exhibit an abnormally higher level of activity than the rest of the days, especially in the sum of votes to the stories posted on these days. The inspection of the corresponding stories reveals that these were prominent days of the 15M movement:

- **17-19/05/2011** The rise of the 15M movement
  On May 15, 2011, the first demonstration took place in the largest cities of Spain. At the end of the demonstration in Madrid, a group of 40 protesters decided to camp in Puerta del Sol Square (*Acampada Sol).* The next day, although police forces attempted to evict the camp, more protesters joined Acampada Sol and around 200 people also decided to camp in Catalunya Square in Barcelona (*Acampada BCN*). This trend continued in the following days and the main squares of cities in Spain were occupied for weeks under the motto '15M movement'.

- **27/05/2011** Violent police eviction of Acampada BCN
The City Council of Barcelona sent 350 police officers to dismantle the protesters in Acampada BCN early in the morning. This action resulted in a violent clash between police and citizens. New calls to protest emerged in all the squares yet occupied in Spain and, in the evening, protesters rebuilt Acampada BCN.

- **21/02/2012** 15M Outbreak in Valencia (*Primavera Valenciana*)
Inspired by the actions of the 15M movement, schoolchildren and university students in Valencia started a rally of daily protests against the Spanish Government because of corruption scandals and the austerity measures proposed for debt control.

- **11/07/2012** Asturian miners' march
Coal miners from Asturias organized a march in Madrid in order to protest against the plans of the Government to reduce subsidies for 40 mines. Asturian miners arrived to Puerta del Sol Square and received the support of thousands of citizens.

- **25-27/09/2012** Encirclement of the Parliament (*25S Rodea el Congreso*)
On September 25, 2012, protesters from the 15M movement decided to surround the Spanish Parliament to claim against austerity measures, the tax system and the overall Spanish political system. Protests resulted into riots between police forces and citizens and, two days later, new surrounding actions were made by protesters.

- **31/01/2013** Podemos' anti-austerity march (*Marcha del Cambio*)
Podemos, emerging political party founded in the aftermath of the 15M movement, organized an anti-austerity march in Madrid. Tens of thousands of citizens attended the event, hosted in Puerta del Sol Square.

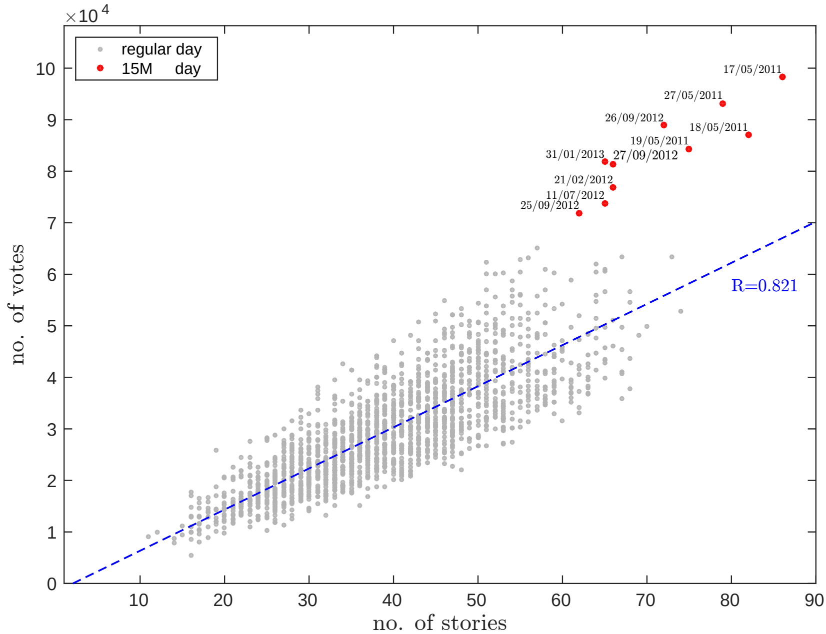


Figure 3.3. Scatter plot of days in the dataset of *Menéame* (2011-2015). Each day is represented by a dot with coordinates consisting of the number of stories in the front page (X-axis) and the sum of the votes to these stories (Y-axis). Although both dimensions are highly correlated, prominent days in the Spanish 15M movement (red markers) exhibit an abnormally high level of activity.

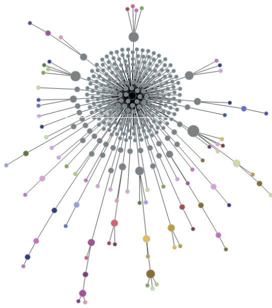
We then explore discussion threads to better understand the typical tree structures in *Menéame*. We adapt an existing thread visualization tool [AGK16] to examine differences in the structural properties between threads from 2011-2014 (i.e., when the conversation view was linear) and threads from 2015 (i.e., when the conversation view was hierarchical). We summarize our findings by illustrating two paradigmatic examples in Figure 3.4 (the two threads from Figure 3.1). In these visualizations, a discussion thread is represented as a radial tree in which nodes are messages and edges are the reply relationship between messages. The root node is the initial message (the story) and comments are expanded radially to indicate their depth in the discussion thread. The size of each node is related to the number of replies and the color of the node is:

- Black: Root of the thread, i.e., the story.
- Grey: First level comments.
- Random: Comments to another comment. To identify authorship, we set the same random color to comments published by the same user.

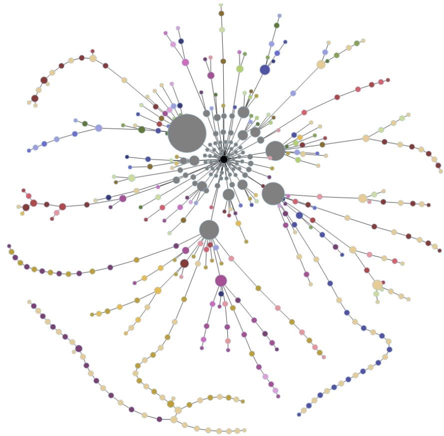
Although both examples of discussions show some similar features, such as chains of two users that alternate messages (i.e., chains of nodes of two alternating colors), there are clear differences. The thread from 2011 (about the emergence of the 15M movement) contains much more direct comments to the original post than the thread from 2015 (about the victory of Barcelona en Comú). Furthermore, the thread from 2015 shows that comments attract often many replies and originate new sub-discussions, an effect that rarely occurs in the thread from 2011. In particular, the left thread has a star-like structure (width=256, depth=5, h-index=4) while the right thread has a much more complex structure (width=110, depth=37, h-index=8). Summing up, we observe that complex discussion structures are more likely when users discuss with the hierarchical conversation view.

3.4.2 Statistical methods

The previous preliminary exploration showed evidence of the relevance of both events, the 15M movement (E1), and the change of the conversation view from linear to hierarchical (E2). To statistically detect events that affected online deliberation and to quantify their effect, we propose a technique inspired by the methodology suggested in [MP16], based on regression discontinuity design (RDD). RDD is



(a) Thread from 2011 discussed with the linear conversation view.



(b) Thread from 2015 discussed with the hierarchical conversation view.

Figure 3.4. Visualization of a discussion thread from 2011 about the 15M grassroots movement occupation of Puerta del Sol Square in Madrid (left), and a discussion thread from 2015 about the victory of the grassroots party Barcelona en Comú in the local elections (right). Nodes (i.e., comments) are sized based on the number of replies. To identify authorship, we set the same color to comments published by the same user except for the root node (black) and the comments at the first level (gray).

a statistical quasi-experimental technique commonly applied in economics to evaluate the causal effects of interventions. In [MP16], an intervention is defined as a time-stamp in a time series (i.e., when an event occurred, hereafter the *cutoff*) and to observe the local average treatment effect on an outcome variable. Given a cutoff c, a (linear) regression is defined as:

$$Y_i = \omega_0 + \omega_1 \cdot x_i + \omega_2 \cdot \mathbf{1}(x_i > c) + \omega_3 \cdot x_i \cdot \mathbf{1}(x_i > c) + \epsilon_i, \quad (3.1)$$

where x_i is the time-stamp (bin size = seven days), Y_i the average value of the outcome variable, $\omega_{0...3}$ the coefficients of the regression, and ϵ_i a random error term. Thus, RDD fits data in two different linear regression functions, before and after the intervention, in order to measure the difference between both functions at the cutoff. The null hypothesis is that $\omega_2 \approx 0$ and $\omega_3 \approx 0$, i.e., the intervention generated no effect.

The purpose of our study is not to measure the effect of a given intervention but to detect from data when an intervention occurred, i.e., an event which significantly affected online deliberation. Therefore, instead of setting an arbitrary cutoff (e.g., the rise of the 15M movement, the change of the conversation view), we apply an F-test, as suggested in [LL10], in every time-stamp of the time series. This approach allows us to find the most significant time-stamp based on the average values of the outcome variable before and after that cutoff.

To detect and measure events that affect online deliberation in discussion threads, our outcome variable is a metric suggested in [GBKB10], which conjugates the two following tree network metrics:

- width: maximum number of comments at any reply level,
- depth: number of reply levels.

To illustrate these two metrics, we present in Figure 3.5 an example thread using a radial tree. For this example, width = 14 (number of comments at the first level) and depth = 3. Width and depth of discussion threads act as good proxies for representation and argumentation, respectively [GBKB10]. This statement is based on the implicit assumption that users tend to follow a sequential posting behavior in discussion threads, i.e., replies explicitly indicate the message being replied to. Therefore, width approximates the number of different users involved in the discussion (to what extent the community is represented in the discussion), and depth indicates the number of messages of the longest chain of messages exchanged between users (how long argumentation lasts in the discussion).

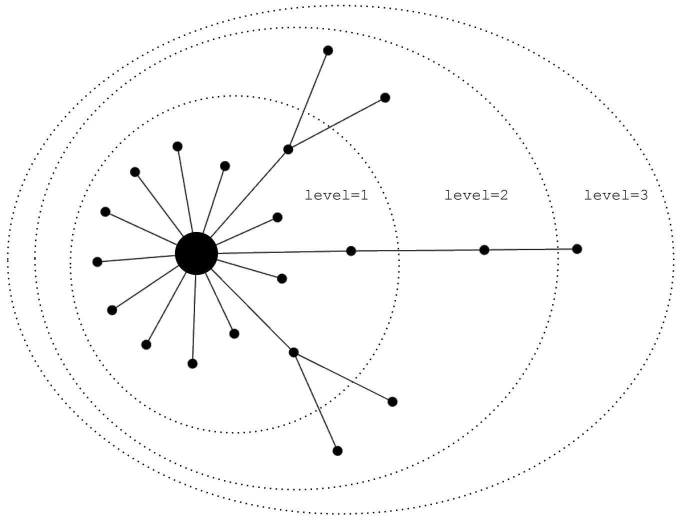


Figure 3.5. Example of a discussion thread presented as a radial tree. Width=14, because is the maximum number of comments at any reply level. Depth=3 because is the number of levels. *h-index*=2 because is the maximum h level in which there are, at least, h comments; i.e., there are more than two comments at the second level but less than three comments at the third level.

To illustrate this approach, we present in Figure 3.6 the four types of discussions defined in [GBKB10], using real threads from our *Menéame* dataset:

- **Type I**. Wide and deep discussion: high levels of argumentation and representation and, therefore, deliberation.
- **Type II**. Deep but not wide discussion: high levels of argumentation but low levels of representation.
- **Type III**. Neither wide nor deep discussion: low levels of argumentation and representation.
- **Type IV**. Wide but not deep discussion: low levels of representation but high levels of argumentation.

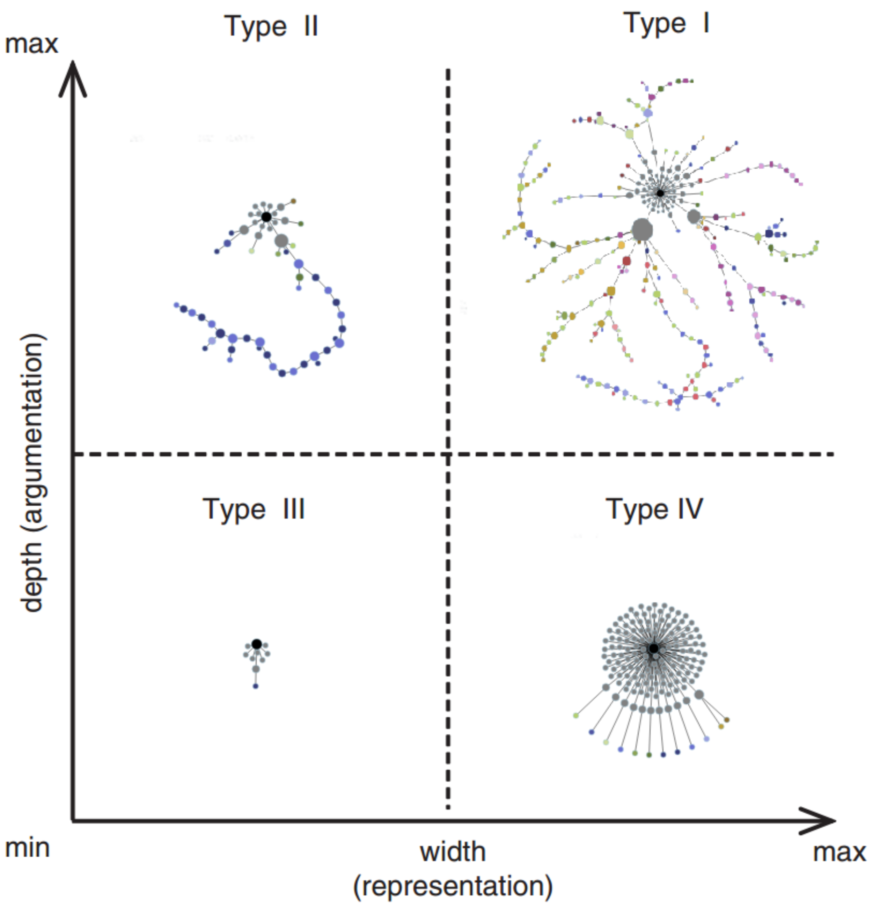


Figure 3.6. The four types of discussions defined in [GBKB10] using real threads from our *Menéame* dataset. Type I presents the best conditions for online deliberation: representation (width of the thread) and argumentation (depth of the thread).

The combination of the width and depth of a thread is then measured through the *h-index* of the discussion thread. This metric, defined in [GKL08], is inspired by the one proposed in [Hir05] which assigns an h index to a researcher who has authored at least h papers with at least h citations each. In a discussion thread, the *h-index* corresponds to the maximum h level in which there are, at least, h comments, i.e., $h + 1$ is the first level in which there are less than h comments. In our illustrative radial tree from Figure 3.5, there are more than two comments at the second level but less than three comments at the third level, therefore, its *h-index*=2.

3.4.3 Results

We first analyze whether the *h-index*, our measure for online deliberation, is affected by any event detected by our proposed method. Figure 3.7 (left) shows the longitudinal F-test statistic as a function of time. The best cutoff appears on 10/01/2015 and corresponds to the exact moment when the original linear conversation view was replaced with a hierarchical one[5] (E2). The regression discontinuity analysis corresponding to that cutoff is shown in Figure 3.7 (right). The discontinuity shows a notable increase in both the *h-index* itself and in the slope of the regression, indicating an acceleration after the intervention. In particular, the break at the cutoff is 0.28 ($\omega_2 = -2.550$; $\omega_3 = 0.0134$).

Since the *h-index* is a non-trivial combination of the width and depth of a discussion thread, we also examine these two metrics separately. The width as a function of time is shown in Figure 3.8. In this case, we observe a strong coupling of the width with a seasonal pattern, possibly reflecting the drop of activity during winter holidays. This prevents direct application of RDD using a linear model. Alternatively, we present a symmetric moving average of 24 weeks, to indicate cyclic activity, and 52 weeks, to completely detrend the time

[5] `https://www.meneame.net/notame/2002188`

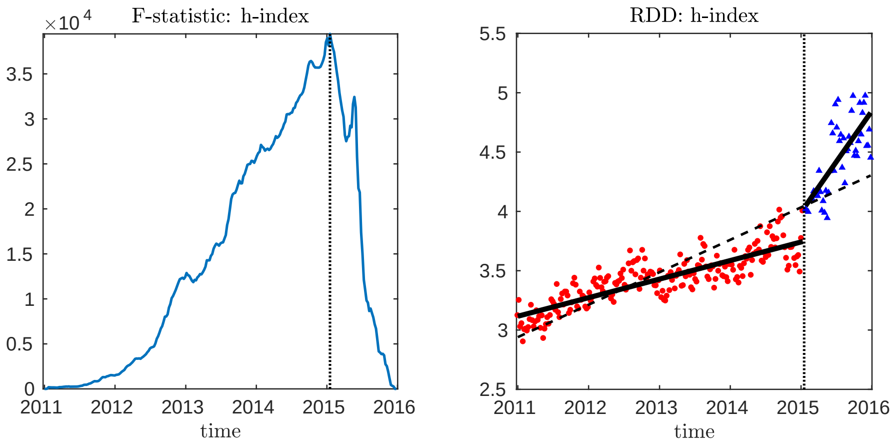


Figure 3.7. Regression discontinuity design applied to the *h-index* (bin size = seven days). The vertical line denotes the optimal cutoff obtained through an F-test. Red circles and blue triangles correspond to discussion threads before and after the optimal cutoff, respectively. The solid line is the result of the discontinuous linear regression and the dashed line corresponds to the linear regression of the null model.

series. This reveals a progressive decay trend in 2013, not related to a sudden change.

Figure 3.9 shows the results corresponding to our analysis of the discussion depth. Unlike the width, this metric does not exhibit a seasonal pattern and it is amenable for RDD using a linear model. As before, in Figure 3.9 (left) we show the F-test value as a function of time. In this case, the global maximum coincides with a local maximum of the *h-index*, four months after the hierarchical view was introduced. Interestingly, by looking for possible explanations of such a change, we found that the hierarchical conversation view was modified at that time. In particular, the maximum depth of the visualized discussion was increased from four to size levels[6]. The RDD results for the thread depth are presented in Figure 3.9 (right). The break at the cutoff is 1.614 ($\omega_2 = -0.488$; $\omega_3 = 0.009$) and confirms the discontinuity, while the null hypothesis does not capture such effect.

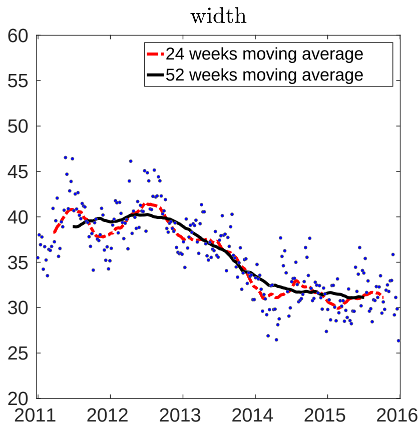


Figure 3.8. The width of discussion threads as a function of time (bin size = seven days). Results show that the width is affected by a seasonal pattern. The dashed lines are symmetric moving average using 24 weeks (red) and 52 weeks (black). A decrease is observed during 2013.

Finally, we look at the relation between depth and width, also over time. In Figure 3.10 we show scatter plots, with horizontal and vertical axes corresponding to the width and depth, respectively (bin size = seven days). The color gradient in Figure 3.10 (left) goes from the oldest threads (blue) to the most recent ones (red). We observe that the first discussion threads are characterized by wide but not deep structures, as in our example thread of 2011 presented in Figure 3.4 (left). Threads progressively acquire more depth and reduce width. This trend changes abruptly in January 2015, when

[6]`https://github.com/gallir/Meneame/commit/b35a6b2`

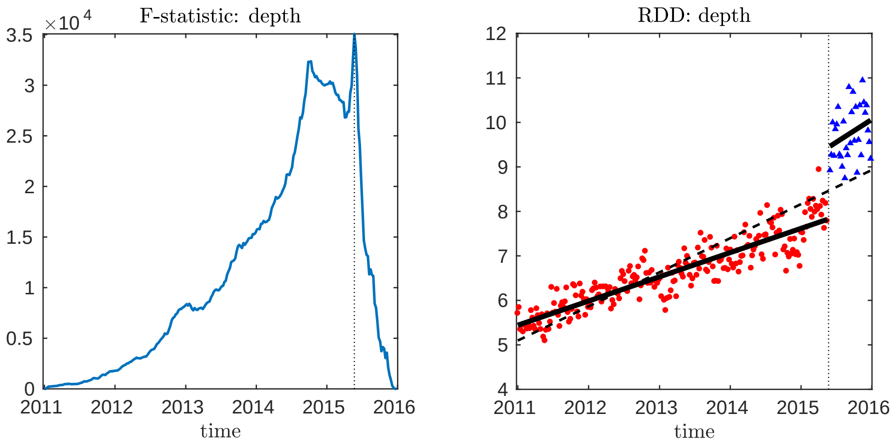


Figure 3.9. Regression discontinuity (RD) design applied to the thread depth (bin size = seven days). The vertical line denotes the optimal cutoff obtained through an F-test. Red circles and blue triangles correspond to discussion threads before and after the optimal cutoff, respectively. The solid line is the result of the discontinuous linear regression and the dashed line corresponds to the linear regression of the null model.

the hierarchical view replaced the original linear view. Subsequently, the width remains stable while the depth grows much faster, especially with the second version of the hierarchical view in which the maximum visual depth is increased. This may explain why the slope of the *h-index* increased in Figure 3.7 (right): the second version of the hierarchical conversation view induced much deeper conversations. Figure 3.10 (right) makes explicit this segmentation using different colors for each period: blue for the linear conversation view, yellow for the first version of the hierarchical view, and red for the second version hierarchical conversation view (increased maximum visual depth).

From these results we can conclude that our methodology detects the change of the conversation view (E2) as the most significant event in *Menéame* in terms of promoting deliberation, since the intensity of argumentation in the discussion threads is increased, an effect which is accentuated with the second version of the hierarchical view.

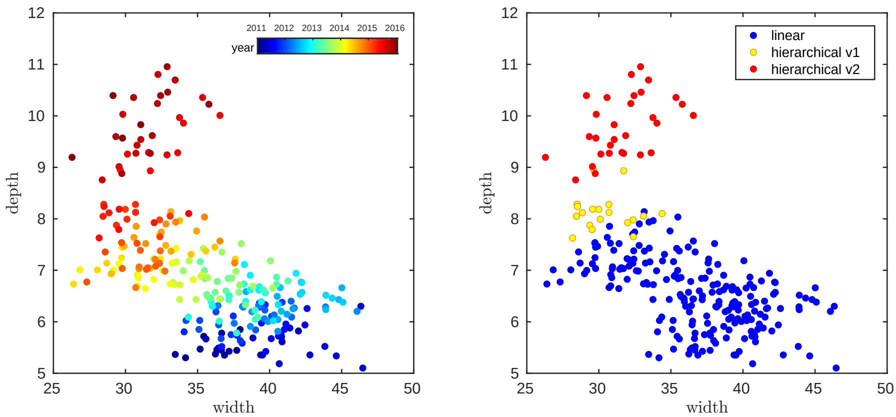


Figure 3.10. Scatter plot of width versus depth of the discussion threads (bin size = seven days). Left: dots are colored in a scale which indicates time. Right: blue dots are threads presented in a linear conversation view, yellow dots are threads presented with the first version of the hiearchical view, and red dots are threads presented in the second version hierarchical conversation view (increased maximum depth). From Jan 2011 to Jan 2015, depth increases while width decreases. Then, the linear conversation view is replaced by the a hierarchical one and the depth grows much faster while the width remains stable.

3.5 Discussion

This analysis of *Menéame* aimed at automatically detect events that affect online deliberation.

Our first candidate event was the rise of the 15M movement (E1). This was motivated by two observations: (a) the relevance of *Menéame* in the communication of this social movement [TCLM+15, OS12, Men11, Pos14] which led to an increase of political talk in the platform, and (b) political subforums in similar social news sites have exhibited greater levels of deliberation than other subforums [GBKB10]. The preliminary exploration of the dataset revealed outstanding levels of activity when actions from the 15M movement occurred. However, our statistical analysis of online deliberation did not find any significant effect induced by E1. Although politics might get relevance by the emergence of this movement, this did not affect the deliberative structure of online discussions.

Our second candidate event was the change of the conversation view from linear to hierarchical (E2). Previous studies indicated that hierarchical views helped to improve communication [SCB00, VN03, FPDL06, McV07]. Indeed, this is the typical interface of asynchronous and deliberative discussions [JK05, SJSN09]. Our statistical methodology allowed us to detect the change from linear to hierarchical in January 2015. When discussion threads started to be displayed hierarchically, the indicator of deliberation (*h-index*) suddenly increased, i.e., discussion structures became much more complex. Therefore, E2 was a significant event. Given that this change occurred in isolation, i.e., no other features took place at that specific time-stamp, this confirms that the adoption of hierarchical conversation views has a positive effect in online public spheres.

As we have indicated in the introduction of the chapter, our methodology was not limited to these two events, any time-stamp was susceptible to be the most significant cutoff in the regression discontinuity design. In particular, we found an additional significant event when the visual depth of online discussions was increased.

Given that we were unaware of this event in the design of the experiment, this finding illustrates the flexibility of our statistical approach.

In general, the increase of the depth of discussion threads (associated with higher argumentation) is induced by the long chains of reciprocal interaction between users, as shown in Figure 3.4. Since reciprocity, sometimes referred to as interactivity or mutuality, is one of the most common features when measuring online deliberation [Dah, Tré04, BWCD11, FE15], future work will focus on whether reciprocity is also affected by these technical changes of the conversation view. In this context our results may prove useful to understand how design of online platforms — in terms of what social information they present — may shape our decision-making environment [Mar17]. Despite the significance of the results, we should reflect on both the benefits and limitations of detecting events through the structural indicators of online deliberation defined in [GBKB10]. The decision of measuring online deliberation using the complexity of discussion threads, while disregarding content, will allow academics to easily replicate this methodology on large datasets from online discussions platforms of very diverse nature. In addition, our characterization of the structure of reply structures is aligned, in part, with [Dah01a] which suggested to focus on the contestation rather than emphasizing communicative rationality. Nevertheless, we are aware that our approach represents a confrontation with the Habermasian conceptualization of public spheres [HHM91] and the existing language-based coding schemes for online deliberation, e.g. [Tré04, SG07, CHH02, FE15]. Language-independent approaches to online deliberation, as applied here, examine the strength of exchanges rather than content. Therefore, these cannot characterize whether users back their comments in a respectful manner or simply fight with a flaming or trolling behavior. Given that these coding schemes are too expensive for large datasets and inspired by [BWCD11], which showed that some features can be automatically inferred, future work might address these limitation with natural language processing techniques to also compute linguistic features.

4 The Impact of Conversation Threading

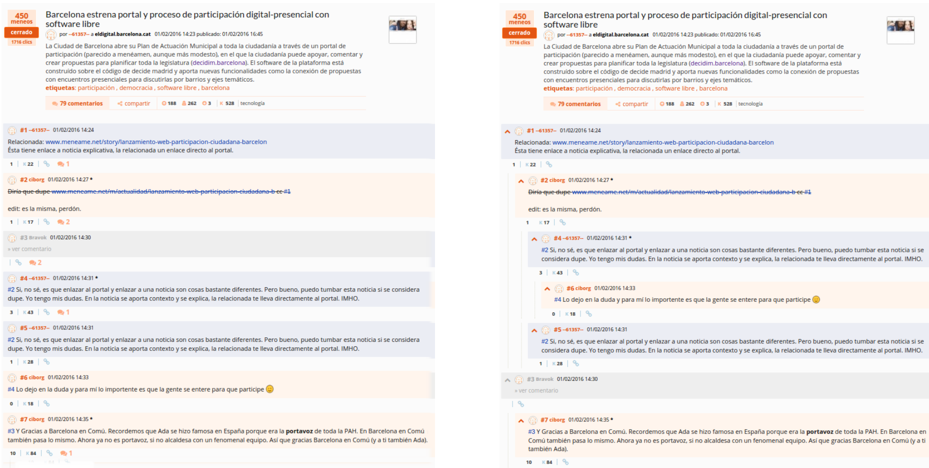
4.1 Introduction

The interaction between users in social media platforms has enabled the emergence of online communities. In these communities, online discussion is essential for the communication and collaboration. Since they are commonly built by strangers, trust between users is only possible when reciprocity occurs [Sea10], for example in the form of a strong exchange of messages between users. Such reciprocity has been traditionally seen as a sign of an inward focus and vigorous debate [FSW06] and some theories have also suggested a relationship between reciprocity and captivating/engaging communication [RS97]. Furthermore, reciprocity is a necessary condition for deliberative purposes because it allows to gain knowledge of the perspectives of others [Hab84]. Thus, many approaches to measure deliberation include reciprocity [Sch97, Jen03, GW03] in order to quantify the degree to which a conversation is a real discussion [JK05].

Although online discussions are simply characterized by an exchange of messages, there are many ways in which a discussion can be presented to a user. Discussion threads are collections of messages posted as replies to previous messages. Therefore, many platforms like email clients and online forums have adopted a *hierarchical* view, also known as *conversation threading*, i.e., messages are arranged close to their replies in a tree-like structure. With this type of view, reciprocal interactions between users are explicitly shown. In contrast, some other platforms show messages regardless of reply relationships with a *linear* view. The sorting criteria of messages with this view is typically chronological to indicate how a discussion thread evolves over time.

Previous work has examined the performance of experimental tools with a specific form of view, either linear or hierarchical. Results usually indicate some benefits of using the hierarchical view, favoring knowledge construction [McV07] or providing better local context [VN03]. Conversation threading mitigates the so-called co-text loss problem [FPDL06], i.e., the inability of readers to "identify which of the previous messages provides the elements that are necessary to understand the message that is being read" [PFdL03]. Co-text loss occurs when interactions are presented separately (e.g., with a linear view) and users are not able to distinguish the earlier message to which a particular message is replying to. A comparative study of both views in an experimental chat found that coherence was also improved thanks to the hierarchical view but, in contrast, participants reported better user experience when interacting with the linear view [SCB00].

These previous studies are based on small groups of recruited participants instead of an existing community, and they do not address how reciprocity is affected. Furthermore, they do not include a modeling approach, thus their theoretical insights about the observed behavioral differences are limited. Generative models of discussion threads have been proposed to explain the structure and growth of online discussion by means of behavioral features, such as popularity or novelty of messages [KMM10, WYH12, GKLK13]. They are language-independent approaches and can thus successfully reproduce many of the structural patterns observed in online discussions of very diverse nature. However, despite reciprocity patterns commonly



(a) Linear **(b)** Hierarchical

Figure 4.1. The two types of conversation views in *Menéame* for an example thread: (a) linear view, before the platform change in January 2015, (b) hierarchical view, after the change. Note that, in both views, every reply to a comment contains the symbol # followed by the id of the comment it replies to. Blue comments are written by the post's author, orange comments by other users. Color intensity is associated with a comment's voting score. Comments scored negatively are shown in white without text unless clicked upon.

emerge in online discussion networks, the state-of-the-art models do not incorporate this as a feature. Therefore, it is unclear whether reciprocity is either a behavioral feature or a resulting effect of discussion dynamics.

In this work we want to increase our understanding of the impact of conversation threading on online discussion. For that, we first measure how the specific type of conversation view affects reciprocity, and then how a modeling approach can capture the interplay between conversation view, structure of discussions, and reciprocity. The reciprocation of interactions plays a primary role since users are motivated to contribute to the community expecting useful help and information in return [KS02, PCW07, GWN14]. Existing theories have established that reciprocity is a defining attribute of online communities [WG99] and a behavioral indicator for their emergence [HBKG04]. Its absence leads communities to fail [Har93]. Given that reciprocity is essential in online communities and conversation threading makes explicit reciprocal interactions between users, as opposed to a linear view, our research questions are:

- **RQ1**: *How does conversation threading affect the reciprocity within the discussion of an online community?*
- **RQ2**: *Is reciprocity a key behavioral feature when modeling the structure and growth of discussion threads?*
- **RQ3**: *How does conversation threading affect the behavioral features when modeling the structure and growth of discussion threads?*

Answering these questions represents a methodological challenge, mainly because of the difficulties and limitations of performing a controlled experiment. We overcome this challenge using data from *Menéame*[1], the most popular Spanish social news networking service (the 2nd most visited site of this type in Spain after *Reddit*[2]). The website interface changed in January 2015. The original conversation view presented the comments of a thread linearly in a chronological order (see Figure 4.1a). Since that change, the comments are displayed by default hierarchically following a tree-structure (see Figure 4.1b). This platform intervention occurred in isolation, which allows us to analyze the impact of such a change with a reduced influence of possible confounders that may also affect the community and the originated discussions. For this reason, *Menéame* becomes an ideal opportunity to measure the impact of the two types of conversation view on a real and large online community.

[1]`https://www.meneame.net/`
[2]`http://www.alexa.com/siteinfo/meneame.net`

4.2 Dataset from Menéame

Menéame is the most popular Spanish social news networking service. Social news websites, like *Reddit*, Slashdot or *Digg*, feature user-posted stories which are discussed in threads, and voted to be ranked based on their popularity within the community. The selection process of featured stories is made by an open source collaborative filtering algorithm similar to the one in *Reddit*. Besides the change of the conversation view (from linear to hierarchical), some other reasons make *Menéame* a platform of interest in our study:

- The community of *Menéame* consists of thousands of users who daily debate hundreds of stories (links to news and blog posts).

- The platform was released in 2005 and therefore *Menéame* is a large and mature community of users which have developed their own culture of practices.

We collected all the stories which were promoted to the front page between 2011 to 2015 (both years included) and every comment from the discussion thread of each story. The reasons for focusing on the promoted stories is because they are more appealing to the community of *Menéame* and to guarantee a sufficiently large volume of comments per story. In total, we obtained 72,005 posts and 5,385,324 comments.

For each comment, we kept the associated meta-data such as the id, the id of the post/comment it is being replied to, the url, the user name, and the time-stamp. We should remark that, as shown in Figure 4.1, both the linear and the hierarchical interface display at the beginning of every reply to a comment contains the symbol # followed by the id of the comment it replies to. Therefore, discussions threads in *Menéame* can *always* be mapped into a tree, which is implicit in the linear view and becomes explicit when the view is hierarchical.

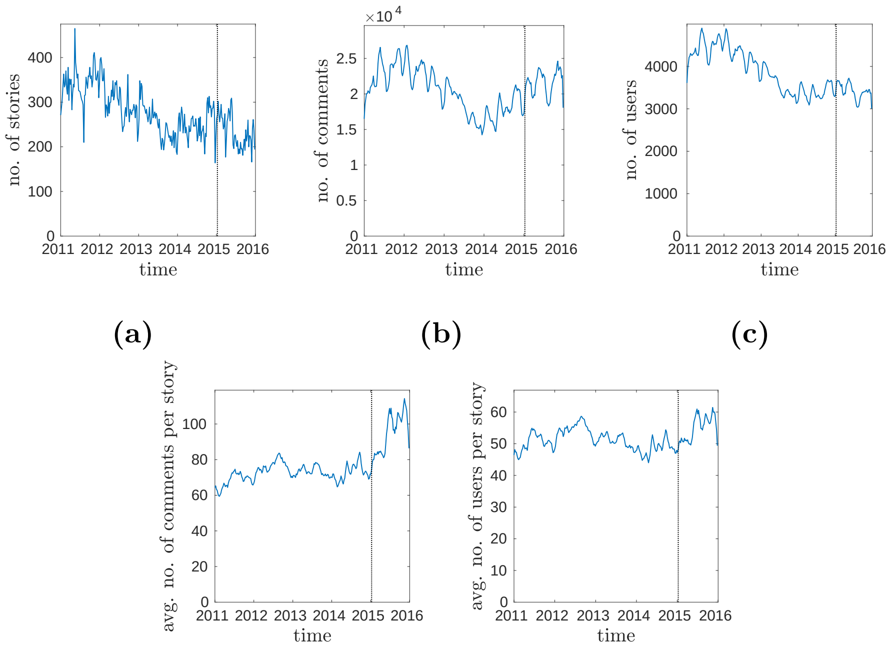
4.3 Measurement of conversation threading effects

In this section we present our statistical analysis and results on the dataset of *Menéame*. We first describe a preliminary analysis and then introduce our methodology based on regression discontinuity design (RDD). We then define how to characterize mathematically reciprocity and describe our results.

4.3.1 Preliminary analysis

To better understand the evolution of discussions in *Menéame*, we first examined the temporal profile of some global activity indicators of the platform. Results are shown in Figure 4.2 with a vertical line indicating the change to conversation threading (January 2015).

We observe that, although the number of stories in the front page (Figure 4.2a) decreases over time, the total number of comments (Figure 4.2b) first decreases from 2011 to 2014 but then increases from 2014 to 2016. The number of unique users (Figure 4.2c) also decreases from 2011 to 2014 but then remains stable. These trends are coupled with a seasonal pattern with activity drops during summer and winter holidays. These cyclic patterns are corrected when one normalizes the binned data by the number of threads. Interestingly, the average number of comments per thread (Figure 4.2d) and unique users per thread (Figure 4.2e) show a sustained increase with an apparent abrupt change in the beginning of 2015, i.e., when the conversation view was modified from linear to hierarchical.



(a) **(b)** **(c)**

(d) **(e)**

Figure 4.2. Number of stories (a), comments (b), unique users (c), average number of comments per story (d), and unique users per story (e). The vertical line indicates the change of the conversation view (from linear to hierarchical).

4.3.2 Impact of conversation threading on reciprocity

To quantify the impact of the change of the conversation view, we apply regression discontinuity design (RDD). RDD is a statistical test used in econometrics to estimate treatment effects in a quasi-experimental setting, where treatment is determined by whether an observed *assignment* variable exceeds a known cutoff point [TC60, LL10]. This technique has been applied recently in previous studies to measure how the design and technical features of a given platform constrain, distort, and shape user behavior on that platform [MP16, HJMY18].

In this work, we use RDD to assess statistically the impact of conversation threading, since we only have observational data in a non-experimental setting. We start from temporal measurements of our variable of interest, in our case, the reciprocity, which we define mathematically in the next paragraph. RDD fits two different functions to this temporal data, before and after the cutoff point (when conversation threading was adopted in *Menéame*) and allows to quantify the break between both fitted lines at the cutoff. The null hypothesis is that the reciprocity is not affected by the release of the new conversation view. . In the linear case the regression is:

$$Y_i = \omega_0 + \omega_1 \cdot x_i + \omega_2 \cdot \mathbf{1}(x_i > c) + \omega_3 \cdot x_i \cdot \mathbf{1}(x_i > c) + \epsilon_i, \quad (4.1)$$

where x_i is the time-stamp of a bin, Y_i is the average value of bin i (bin size = one month), ω_i are the coefficients of the regression, ϵ_i is a random error term, and c is the cutoff.

Linear RDD fits two different linear functions, before and after the cutoff, and allows to quantify the break between both fitted lines at the cutoff. The null hypothesis is that there is no discontinuity (the metric is not affected by the release of the new conversation view), i.e., $\omega_2 \approx 0$ and $\omega_3 \approx 0$.

The cutoff in classical RDD is the intervention given in the experiment. In the context of platform effects for our study, the cutoff is expected to be the time when the conversation view was modified in *Menéame*. However, we should note that, by definition, a discontinuous regression with a cutoff at midpoint of the time series is likely to better fit data than a continuous regression. Therefore, to enhance the robustness of our analysis and to prove the statistical significance of the change of the conversation view, we use an F-test, as suggested in [LL10], to set the cutoff as the most significant point in the time series.

In all RDD reported results, we will prevent biased estimates of the treatment effect by checking that the linear model represented a good model using a statistical analysis of the residuals.

To formally characterize reciprocity, we consider the directed network of replies between users in each discussion thread. In this network, each node correspond to a user and a directed edge between user u and v exists if user u replied to user v in the discussion. The weight of that edge is the number of times u replied to v in that thread. Given a directed network of N nodes, reciprocity is traditionally defined as follows:

$$r = \frac{E^{\leftrightarrow}}{E}, \tag{4.2}$$

where $E^{\leftrightarrow}$ corresponds to the number of bidirectional edges and E corresponds the total number of edges. This approach is limited in the sense that it does not consider the relative difference of reciprocity in comparison to a random network with the same number of nodes and edges. The definition in [GL04] overcomes this problem and defines the *corrected* reciprocity as

$$\rho = \frac{r - \bar{a}}{1 - \bar{a}}, \tag{4.3}$$

where $\bar{a}$ is the network density, i. e. the ratio between the number of existing edges and the total number of possible edges $\bar{a} = E/(N(N-1))$. The previous definitions of reciprocity do not take into account the weighted nature of edges in the reply network, i.e., the number of times that two users interchange messages within a thread. The proposed definition in [SPRG13] for reciprocity of weighted networks

$$r_w = \frac{W^{\leftrightarrow}}{W} = \frac{\sum_u \sum_{v \neq u} w_{uv}^{\leftrightarrow}}{\sum_u \sum_{v \neq u} w_{uv}}, \tag{4.4}$$

where u, v are nodes indexes, w_{uv} is the weight of the edge from u to v, and $w_{uv}^{\leftrightarrow}$ is the minimum weight between the edge from u to v and the edge from v to u.

We construct one network of replies between users for each conversation and compute the three previous reciprocity indicators in each of these networks. In the following, we omit results using r because they are indistinguishable from the results using ρ. This is explained because the constructed reply networks are very sparse and the density $\bar{a}$ is low. We then average these indicators at a time resolution of one month, which defines the bin-size in our analysis. The bin size is an arbitrary choice, we experimented with several sizes but observed no significant differences.

We show in Figure 4.3 how both corrected reciprocity ρ and weighted reciprocity r_w change over time, together with the results of the RDD test. We first note that both reciprocity measures show a sustained increasing trend, which suggests that captivating/engaging communication increases over time. Furthermore, if reciprocity is a defining attribute of an online community, as proposed in [WG99], the increasing trend can be interpreted as a positive indicator of the performance of *Menéame*. The weighted measure is slightly higher than the non-weighted metric, which suggests that the frequency of replies between the same users is important. However, both profiles are very similar, so this frequency is not qualitatively determinant.

We should remark that we use an F-test in every point in the time series to establish the most significant cutoff. Our analysis identifies January 2015 as the optimal cutoff, which corresponds exactly with the transition of the interface. This is indicated in Figure 4.3 by a black dashed line that separates the data before (in red) and after (in blue) the cutoff. The results show a notable impact for the both corrected reciprocity and weighted reciprocity. The obtained values in the RDD for the corrected reciprocity ρ are $break = 0.019, \omega_2 = -0.171$ and $\omega_3 = 0.004$. The corresponding values for the weighted reciprocity r_w are $break = 0.021$, $\omega_2 = -0.192$ and $\omega_3 = 0.004$. This means that the null hypothesis can be rejected and, therefore, there is a significant effect in reciprocity when *Menéame* transitioned from a linear to a hierarchical conversation view. It is also important to mention that the slope increased after

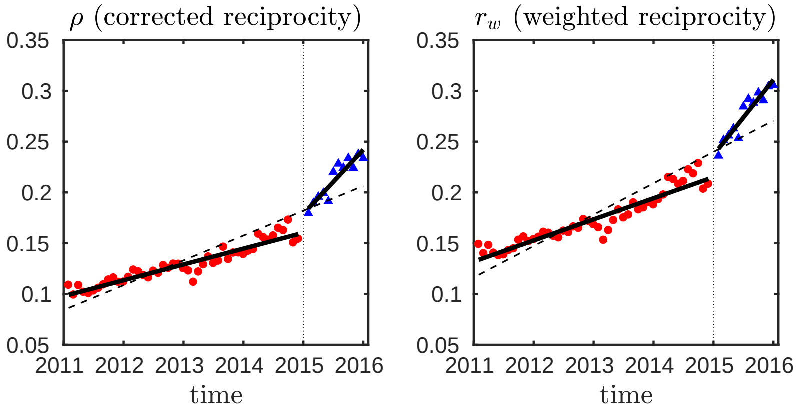


Figure 4.3. Regression discontinuity design for the metrics of reciprocity in the discussions (bin size = one month). Red circles and triangles data are points before and after the cutoff (vertical line). Solid line is the discontinuous linear regression, the dashed line is the continuous linear regression of the null model.

the change, indicating that reciprocity, not only changes abruptly after the adoption of conversation threading, but also increases at a higher speed during the period of available data considered. We will further discuss the impact of these findings in the discussion section.

4.4 Modeling reciprocal online discussions

We now take a modeling approach to gain understanding of the interplay between the structure and the evolution of the discussions, the reciprocity as an abstract feature, and the type of representation. In the next subsection, we characterize informally the network structure of discussion threads. We then describe an existing generative model of online discussions and present our extension which incorporates an authorship model and reciprocity. We show that our proposed extension better explains the observed data. Finally, we perform RDD within the model features.

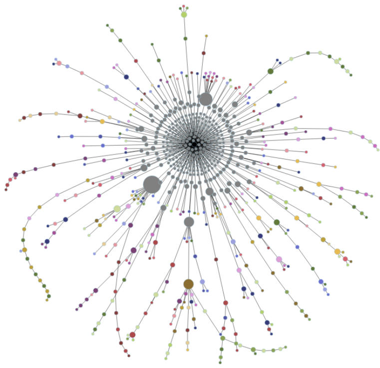
4.4.1 Structure and growth of discussion threads

To illustrate the typical structure of discussion threads in *Menéame*, we use a thread visualization tool [AGK16]. Note that these networks differ from the reply networks analyzed in the previous section, since nodes here corresponds to comments, instead of users.

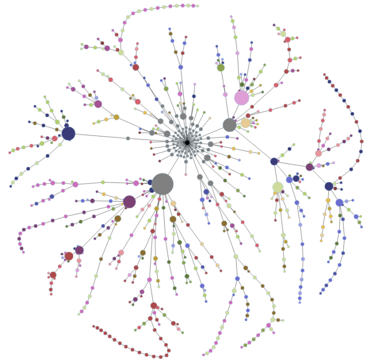
In Figure 4.4, we present two popular discussion threads that took place before and after the platform change. The first one is from 2013 (left) and the second one is from 2015 (right). Node color follows the criteria: black (root of the thread, i.e the story), gray (first level comments) and random color (replies to comments). We observe that every reply written by the same user gets an identical random color. These criteria allow us to observe that both threads share some similarities, such as long chains of two users that alternate reciprocal interactions (i.e., chains of nodes of two alternating colors). This finding is consistent with previous work on modeling the structure and evolution of discussion cascades using data from *Menéame* [GKK11b]. Node size corresponds to the number of received comments (except for the root) and shows that replies (colored nodes) in the thread of 2015 often attract themselves many replies and originate new subdiscussions within the thread. This effect is not that pronounced in the 2013 thread, in which comments usually belong to chains of two users and rarely trigger a discussion cascade. In summary, we observe that the thread from 2013 is closer to a star-like structure (i.e., contains many more direct comments to the original post) while the thread from 2015 is more complex with higher branching probability at deep levels of the discussion.

4.4.2 A generative model of discussion threads

To measure the impact of using a hierarchical view in the evolution of the discussion threads, we build on the model introduced in [GKLK13], which has proven to be successful in capturing the structural properties and the temporal evolution of discussion threads present in very diverse platforms, e.g., *Slashdot*, *Barrapunto*, *Wikipedia*



(a) Thread in 2013. `https://www.meneame.net/story/1860558`



(b) Thread in 2015. `https://www.meneame.net/story/2484585`

Figure 4.4. Visualization of two example threads before (a) and after (b) the conversation view was modified. Black node is the root of the thread (the post). Gray nodes are first level comments. The other nodes are replies to comments where comments written by the same user get the same color. Node size corresponds to the number of received comments, except for the root.

and the same Menéame (before conversation threading was adopted). It is a parametrized mathematical model that generates growing trees in discrete time. At each time-step, a new comment (node) arrives to the thread and each of the following structural features is considered for each node in the discussion:

- The *popularity* or number of replies. A node will *attract* replies proportionally with factor α to the number of replies received so far.

- The *novelty* or the elapsed time since it was written. Recent comments will tend to be more replied than old comments. Novelty decays exponentially with parameter τ.

- The *root-bias* or tendency to write more comments to the root node. This differentiates between the original post (root node), which attracts replies with factor β, and ordinary comments (non-root nodes).

Formally, the discussion thread at time-step t is represented as a vector of parent nodes $\pi_{1:t} = (\pi_1, \pi_2 \ldots, \pi_t)$, where π_t indicates the parent of the node written at time t. When a new comment arrives to the discussion, it is attached to an existing node $j \in 1, \ldots, t$ with probability proportional to its *attractiveness* function $\phi_j(\cdot)$, defined as a combination of the features $\theta = (\alpha, \tau, \beta)$

$$\phi_j(\pi_{1:t}; \theta) := \alpha \text{deg}_j(\pi_{1:t}) + \tau^{t+1-j} + \beta \delta_{j,1}$$
$$p(\pi_{t+1} = j | \pi_{1:t}; \theta) \propto \phi_j(\pi_{1:t}; \theta), \tag{4.5}$$

where $\text{deg}_j(\pi_{1:t})$ is the degree of node j in the tree $\pi_{1:t}$ and δ is the Kronecker delta function, i.e., β is only relevant for the root node.

The model parameters are estimated through maximum likelihood given a dataset composed of M threads $\mathcal{D} = \{\pi^{(1)}, \ldots, \pi^{(M)}\}$ corresponding to a particular period of time.

The previous generative model may fail in describing some structural properties, such as the average depth of a comment, which tends to be underestimated, as noted in [GKLK13]. This is actually the case in *Menéame*, which is characterized by very deep threads with long chains of messages between two alternating users, as shown in Figure 4.4. We postulate that the original model fails to capture precisely that commenting behavior tends to be reciprocal, i.e., users tend to reply comments that are replies to their previous comments. In the next section, we extend the original model with an authorship model and introduce a new feature: the reciprocity.

4.4.3 Extending the model

We now represent a conversation thread with the parent vector $\pi_{1:t}$ together with a vector of respective authors $a_{1:t} = (a_1, a_2, \ldots, a_t)$. The authorship vector will grow depending on the structure of the discussion, which in turn will depend on the authorship of the messages.

Our author model does not allow two consecutive comments to be written by the same user. Furthermore, a user cannot self-reply a comment made by herself. Let U denote the number of different users that participated in the conversation so far. At time $t+1$, a new comment is originated from a new user with id $U+1$ with probability p_{new}, or otherwise from an existing user v chosen according to how many times user v has been replied in the thread, r_v. Our author model is described as

$$p(a_{t+1} = v | a_{1:t}, \pi_{1:t}) = \begin{cases} p_{new}, & \text{for } v = U + 1 \\ \frac{(1-p_{new})2^{r_v}}{\sum_{i=1}^{U} 2^{r_i}}, & \text{for } v \in 1, \dots, U \end{cases} \tag{4.6}$$

We set p_{new} empirically to $p_{new} = t^{-1/k}$ and estimate k from the data ($k \approx 7$). Notice that the preferential attachment process that selects authors is multiplicative. This is required to capture well the probability distribution of the number of comments per unique author in a thread. Once the author a_{t+1} is decided, the new comment is attached to an existing comment j proportionally to the extended attractiveness function $\phi'_j(\cdot)$, which now depends on the vector of authors $a_{1:t}$ and the parameters $\theta' = (\alpha, \tau, \beta, \kappa)$

$$\begin{aligned} \phi'_j(\pi_{1:t}, a_{1:t}; \theta') &:= \phi_j(\pi_{1:t}; \theta) + \kappa \delta_{a_{\pi_j}, a_{t+1}} \\ p'(\pi_{t+1} = j | \pi_{1:t}, a_{1:t}; \theta') &\propto \phi'_j(\pi_{1:t}, a_{1:t}; \theta'), \end{aligned} \tag{4.7}$$

where the additional term $\kappa \delta_{a_{\pi_j}, a_{t+1}}$ is non-zero for reciprocal comments only and $\phi_j(\cdot)$ is the original (author-independent) attractiveness function given in Equation (4.5).

The new parameter κ determines how strong reciprocal comments are weighted. Only those replies to comments authored by the selected author, i.e., $a_{\pi_j} = a_{t+1}$, will contribute to the κ-term. Thus, for $\kappa = 0$ the new feature will play no role in the evolution of the thread whereas very large values of κ will make all comments of corresponding users reciprocal. The additional parameter κ can be optimized using maximum likelihood together with α, β and τ.

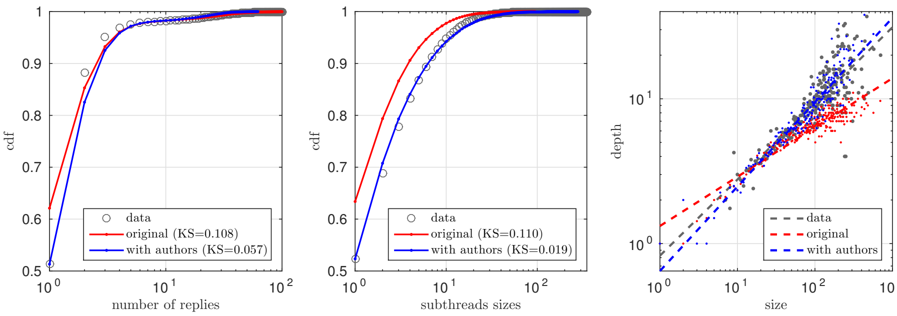


Figure 4.5. Comparison between the original model (in red) and the extended model with authorship and reciprocity (in blue) in terms of how well they reproduce the real discussion threads (gray circles). The plots show the cumulative distribution function (cdf) of the degrees (left), subtree sizes (center) and the correlation between depth and number of comments (right). The curves where obtained from $2 \cdot 10^3$ threads generated from both models after optimization of their respective parameters. Dashed lines in the right subplot correspond to linear fits in the logarithmic domain. KS indicates Kolmogorov-Smirnov test value (the lower the better).

We first compare the original model and the proposed extension and then we analyze how the change in the interface affects the model parameters. To show that the extended model not only reproduces better the depths, we also compare the two models using the same indicators as in [GKLK13]. Figure 4.5 shows that the distribution of the number of replies, subthreads sizes and the relation between the thread sizes and depths are reproduced significantly better thanks to the authorship model and the reciprocity feature.

Figure 4.6 shows the empirical probability distributions (pdf) of the depth of a comment calculated from the real threads and from synthetic ones generated from both models after optimizing their respective parameters. Whereas the resulting depths using the original model are underestimated (red curve), the extended model is able to generate deeper threads and to reproduce better the depth distribu-

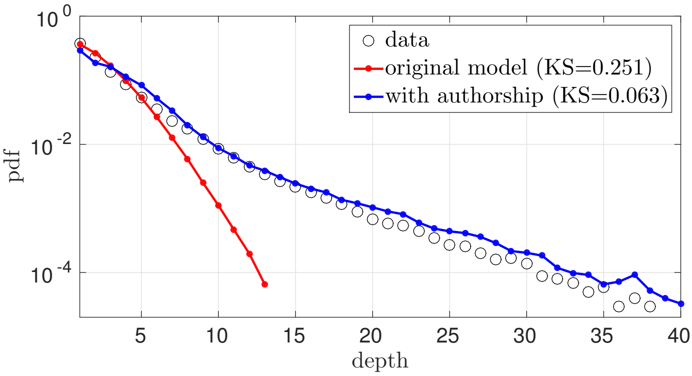


Figure 4.6. Probability distribution of the comment's depths. The original model fails to capture the long tail created by reciprocal message chains whereas the proposed model is able to reproduce the data accurately. The curves where obtained from $2 \cdot 10^3$ threads generated from both models after optimization of their respective parameters. KS indicates Kolmogorov-Smirnov test value (the lower the better).

tion. In particular, it captures the tail behavior accurately and the observed discrepancies are only minor. The KS test accepts the model hypothesis at the 5% confidence level (p-value = 0.0041). The synthetic threads also contain chains of messages with alternating users, as in the original data. We thus conclude that by increasing minimally the complexity of the model with the authorship model and the reciprocity, the overall descriptive power of the model is greatly improved.

4.4.4 Impact of conversation threading on behavioral features

We now analyze how the platform change affected the evolution of the threads by fitting the extended model to data from different periods of time. In Figure 4.7 we present the results of the RDD on the four estimated parameters, each of them corresponding to one of the features. We obtained the following RDD values for the reciprocity feature κ (break = 51.28; $\omega_2 = -287.78$; ω_3=7.06), the popularity α (break = 0.12; $\omega_2 = -0.35$; ω_3=0.01), the novelty τ (break = 0.08; $\omega_2 = -0.15$; ω_3=0.004), and the root-bias β (break = 3.72; $\omega_2 = -1.90$; ω_3=0.12).

Globally, we observe notable increases in all the parameters after the platform change. The most noticeable change corresponds to the reciprocity feature, parameterized by κ (see the change of order of magnitude in Figure 4.7). Once the hierarchical view is active, users behave significantly more reciprocally and tend to engage more in dialogues. These findings are consistent with the above one for the corrected and the weighted reciprocity metrics.

The other features also show an abrupt increase after the platform change, but to a lesser extent. We emphasize that the interplay between the features may be nontrivial, even mediated by a hidden, not modeled feature, since the relative weights differ between the two conditions. Nevertheless, since reciprocity is only relevant at the later stages of the discussions, where comments are written from existing authors that have already been replied, their relevance is also increased after the platform change. Finally, it is interesting to mention that the same analysis performed in the original model was unable to detect a significant change in parameters β and α at the time of the platform intervention.

4.5 Discussion

We have presented a study about the impact of conversation threading in online discussions. While previous studies in this field [McV07, FPDL06, VN03, SCB00] had relied on experiments recruiting small groups of participants, our findings are observed in an existing, large and mature community with over five years of online discussion data.

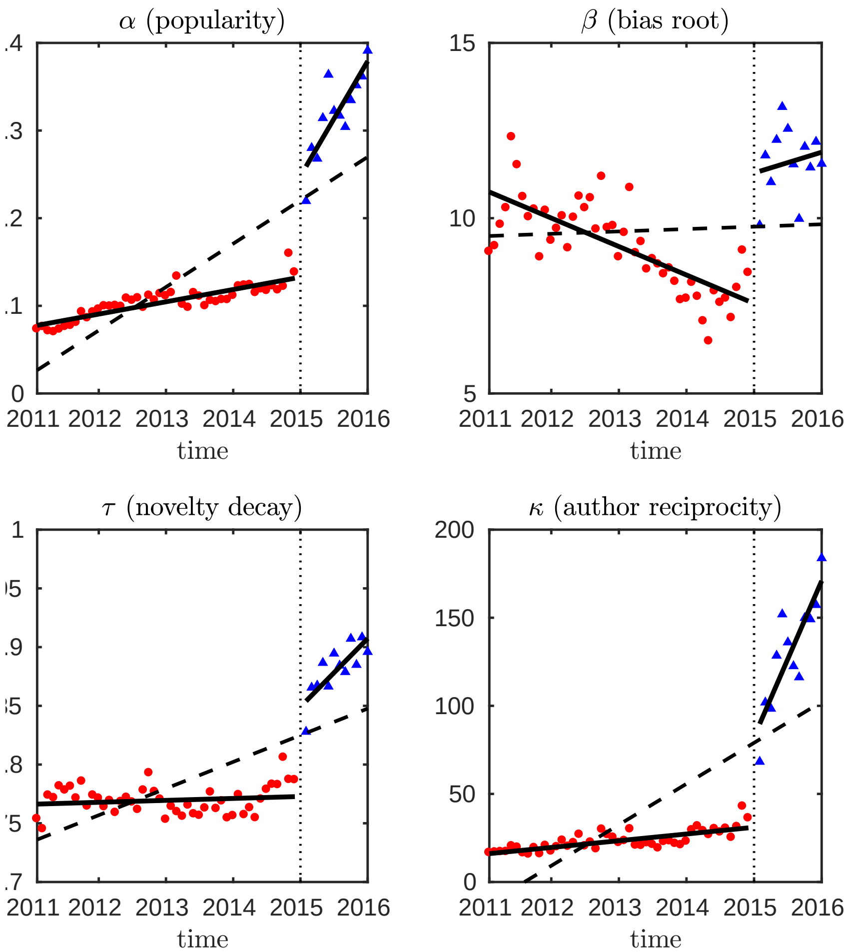


Figure 4.7. Regression discontinuity design for the behavioral features of the model (bin size = one month). Red circles and triangles data are points before and after the cutoff (vertical line). The solid line is the discontinuous linear regression, the dashed line is the continuous linear regression of the null model.

We first analyze how the implementation of conversation threading affects the reciprocity in the discussion of an online community (*RQ1*). One would expect reciprocity to increase since a hierarchical conversation view emphasizes the exchange of messages between users. Indeed, although we already observe a natural increase of reciprocity over time, as suggested in [FSW06], the adoption of this type of interface triggers an additional boost leading to even higher levels of reciprocity. This is a positive behavioral indicator for online communities [WG99, HBKG04, PCW07], and it is aligned with previous work on the benefits of hierarchical views for constructing knowledge [McV07], providing better context of the discussion [FPDL06, VN03] and improving coherence [SCB00]. Our results have implications for the characterization of user roles in online discussion. Reciprocity has been used to distinguish different types of users in online forums; e.g *taciturns* with low tendency to reciprocate interactions and *grunts* with relatively higher levels of reciprocity [CCC06]. Given that users in social media change their role over time [GGKST^{+}14], this opens interesting research directions like assessing whether the distribution of user roles is affected by changing the conversation view.

The relevance of reciprocity in online discussion leads us to reflect on the role of this behavioral pattern in the formation of discussion threads. The existing generative models of discussion threads [KMM10, WYH12, GKLK13] include features from messages like the popularity (number of incoming replies) or the novelty (timestamp). However, the tendency of users to reply to the replies to their messages has only been considered indirectly, *a posteriori*. For example, although the authorship model in [KMM10] establishes authors of messages to promote the reciprocity of replies, it does it after the discussion thread is generated and, therefore, reciprocity is ignored during the growth of the discussion. Furthermore, all of these models fail in modeling accurately the depth of discussion threads which can be explained by the occurrence of long chains of reciprocal interactions between two users, as postulated in [PB12] and empirically shown in Figure 4.4.

This leads to our second research question (*RQ2*) about whether reciprocity can improve the descriptive power of models of discussion threads. To answer this question we extend the model in [GKLK13] by incorporating authorship and establishing reciprocity as a behavioral feature. This is an important difference from previous models such as [KMM10], in which the structure of a thread does not depend explicitly on the authorship. To the best of our knowledge, the presented model is the first in which the structure and authorship co-evolve jointly. The results on discussion threads from *Menéame* show that our approach not only captures better the distribution of the number of replies and sizes of subthreads, it also reproduces more accurately the temporal evolution of the discussion threads in view of the distribution of the depth of discussion threads.

The model extension includes reciprocity together with the existing behavioral features of popularity, novelty, and root-bias. This allows us to answer our third research question (*RQ3*) which analyzes whether modeling discussion threads can quantify the impact of the conversation view on behavioral features. On the one hand, our results show that the hierarchical view induced more reciprocal behavior, which is consistent with the findings from the regression discontinuity design. On the other hand, we also observe that the transition to threaded discussion makes popular comments to attract more replies and slows down the decay of novelty, i.e., comments take longer to be ignored. This second effect can be explained by the fact that the hierarchical view on *Menéame* does not apply comment folding and, therefore, branches of comments are always fully expanded. With this type of interface, conversation threading gives preference to the first comments and their replies, i.e., branches (and sub-branches) are ordered chronologically. Although it is true that reciprocity increases and online deliberation requires reciprocity [Sch97, Jen03], new contributions with no connection with previous arguments will be less visible to the community. Given that deliberation also requires users gaining knowledge of the perspectives of others [Hab84], additional mechanisms (e.g., comment folding, branch sorting) must receive special attention in the design of online discussion platforms.

Our methodology is based on the structural properties of the discussions and is language-independent. Therefore, it can be easily applied to other platforms. For this reason, modeling approaches like the ones applied here can also be used to assess the impact of other features in online discussion platforms and to compare the model parameters in different environments and communities. Moreover, it might be of interest to extend these models to further explore content-based features from the messages of the discussions. Recent studies have suggested that linguistic indications of reciprocity can measure the chance of success of individual requests in online communities [ADNMJ14]. Also, hierarchical comment threads have been noted to represent a topical hierarchy in online discussions [WZH13]. Therefore, future work might explore whether the transition from a linear to a hierarchical conversation view can also affect the narrative structure and the distribution of topics in online discussion.

Part II

Analysis of Online Petitions

5 The impact of petition ranking

5.1 Introduction

The invention of the Internet has revolutionized our society and, as a result, our forms of governance. Many different approaches have been suggested to leverage information and communication technologies to empower citizens and, therefore, to improve the democratic strength of governments [Nov09]. A clear indicator of this global trend is the increasing popularity of online petition platforms. The reported benefits of these innovative and collaborative practices of policy making are very diverse. For instance, research on the first experiences from local institutions in the UK found out the important role of e-petitioning in providing very valuable feedback [PE12], increasing political engagement [AMJ05], and reinforcing 'civic mindedness' [WRM05].

In spite of the popularity of online petition platforms, critical voices have also been raised. Many of them have accused e-petitioning of promoting the so-called 'slacktivism': a feel-good online activism with no political or social impact [Mor09]. Indeed, these criticisms have become even more thorough by suggesting the term 'technological solutionism' [Mor13]. However, despite these critical considerations of online petitioning, experimental research from human-computer interaction has proven the ability of online activism to influence civic actions [LH13].

Research on data from online petition platforms has covered institutions from many different countries, e.g., the United Kingdom [Wri12, HMY13], Germany [JJ10, LR11], or the United States [DLH$^+$15, MJHY15, YHM17]. In Spain, from 2015 onwards, there has been a trend towards the implementation of civic technologies by local city councils. The platform that has gathered the largest number of participants (hundred of thousands) is *Decide Madrid*[1] which, among other features, allows participants to publish, to discuss, and to sign online petitions. Despite its high level of activity, very few petitions have managed to reach the minimum number of signatures required to be considered by the City Council.

Previous work on e-petitioning concluded that it is essential to understand how these platforms work and hence avoid unrealistic expectations which lead petitioners to be upset at the results [Wri12]. However, as we detail next in the Background section, almost every study about growth and success of online petitions has focused on features extracted from petitions [HMY13, DLH$^+$15, HHU$^+$15, EDK16, CLM17, MGSP17, CLHD17]. In this way, our study has been designed in collaboration with the City Council to characterize the dynamics of petition signing in *Decide Madrid* with the aim of detecting platform features that may be hindering citizen participation.

The analysis of data from petitions and signatures presented in this chapter suggests problems originated by the original design of the home page, which was based on a ranking inspired by *Reddit*, and by having deployed multiple participatory processes on the same platform. Furthermore, our results have motivated a new sorting criterion for that ranking, and we analyze the effect of using one or another strategy to rank online petitions. To the best of our knowledge, this study is the first that analyzes the effect of an intervention

[1] `https://decide.madrid.es`

in the ranking algorithm of an online petition platform. Such intervention has positively affected petition growth in *Decide Madrid*. As a consequence, given that most petition platforms include rankings in their home page, our findings have relevant implications for the design of civic technologies.

5.2 Background

This section presents the background of this study including related work and a detailed description of *Decide Madrid*.

5.2.1 Related work

The increasing popularity of online petition platforms has attracted attention from scholars for many different research purposes, e.g., to characterize sociodemographic characteristics and behavioral patterns of petitioners [LR11, LH13, HSHH15, DLH+15, PBS17, ASTR+18], to detect relevant topics in petitions [HHU+15, CLM17, MGSP17], or to identify the factors which lead online petitions to succeed [Ber17, EDK16, AM17].

Substantial efforts have been devoted in recent years to characterize the growth of online petitions in different platforms. An early study in this field focused on the petition growth and success rates on the *UK government website* [HMY13]. The analysis indicated that petitions grew quickly and revealed that the number of signatures on the first day was the most significant factor in explaining their final number of signatures. Following studies to that work explicitly examined the temporal aspects of petition growth by analyzing the time-series of signatures. The analytical approach in [HMY13] was extended with a formal multiplicative process model framework to explain the rapid rise and decay in petition signing [YHM17]. The model, based on a previous model of novelty and collective attention in *Digg* news stories [WH07], was validated with petitions from the UK government and the US White House websites: petitions grew very rapid in their first two days but the outreach factor decayed very quickly on average. Another analysis of the US White House petition website showed that petitions are more likely to fail when the number of signatures is lower on the second day than on the first day [CLHD17]. Finally, a recent study focused on the inter-signature time distributions of petitions from the *openPetition* platform taking into consideration the final number of signatures [BWMB17]. That analysis concluded that petitions with many signatures are less likely to exhibit bursty signing dynamics.

All of the above works focused on analyzing petition growth based on attributes from petitions. To the best of our knowledge, the only observational study of the impact of a petition platform intervention investigated the introduction of a ranking of trending petitions on the home page of the UK platform [HJMY18]. Results showed that the effect was weak for the complete population of users but strong for a specific group of users, so-called *aimless petitioners* [BC11]. These users usually accessed the platform through the home page rather than starting with a specific petition because they might have a political interest but not a specific purpose. *Aimless petitioners* were numerous enough and affected strongly enough that the social information on trending petitions significantly affected petition signing on the site as a whole. Nevertheless, we must recall that [HJMY18] analyzed the effect of introducing a ranking in the home page, rather than comparing the effect produced by different ranking strategies, as presented later in this chapter.

5.2.2 Decide Madrid

The case study of this work is *Decide Madrid*, a platform launched by the City Council of Madrid to host different participatory processes of direct democracy. The most representative process, named 'Citizen Proposals', consists of online petitions and is active since September 15, 2015. The procedure is as follows: any user can publish a petition and any petition which obtains at least 27,064 signatures[2] in one year will go to a direct voting second phase. Otherwise, the petition is withdrawn. In the second phase, users are able to vote in favor or against the petition. If the majority of votes are positive and the petition fulfills some technical/ethical requirements, the City Council must implement it. So far, only two petitions have reached this threshold.

We should remark that, in order to sign petitions, users must verify their accounts using their personal data from the local city census. This feature is relevant because it restricts participation, but also mitigates the problem of duplicate signatures or false identities, as noted by [Wri12].

Home page design

We show the home page of online petitions in *Decide Madrid* in Figure 5.1. Petitions are presented in a paginated sorted list of 25 petitions per page. Until this study was carried out, the original sorting criterion had been an adapted version of the *Hot score* of *Reddit* to boost trending petitions on the home page and thereby make it easier for newer petitions with fewer total signatures to receive attention[3]. Although users have always been able to change the sorting criterion to display the rankings of *Most Popular Petitions* (sorted by the number of signatures) and *Most Recent Petitions* (sorted by date), the default option for any new user session was the *Hot score*. In addition to this, a yellow banner featuring the three most popular petitions was added in October 2015 for every page of the website[4]. Then, the banner was limited to two petitions in January 2016[5] and removed in September 2016[6].

[2] 1% of the population of Madrid over 16 years old
[3] Commit at `http://bit.ly/commitHot`
[4] Commit at `http://bit.ly/commitAddBanner`
[5] Commit at `http://bit.ly/commitChangeBanner`
[6] Commit at `http://bit.ly/commitRemoveBanner`

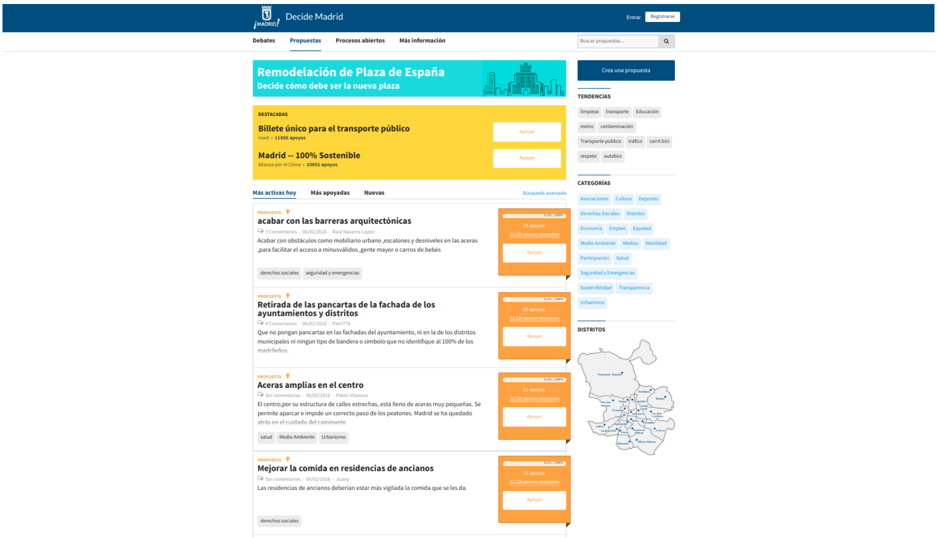


Figure 5.1. Screenshot, taken in 2016, of the home page of online petitions in *Decide Madrid*. Petitions are presented in a paginated list of 25 petitions per page, sorted by an adapted version of *Hot score* from *Reddit*, i.e., recent petitions which are rapidly gathering signatures. However, users are able to explore two alternative rankings: *Most Popular Petitions* and *Most Recent Petitions*. The screenshot also shows the yellow banner featuring the two most popular petitions at that time.

Multiple participatory processes

A very relevant aspect of *Decide Madrid* is that the platform is not only designed for online petitions but also for other massive participatory processes in parallel to the 'Citizen Proposals' process. This is a very particular characteristic with respect to the other platforms analyzed in previous studies, which were exclusively designed for online petitions. The relevance of this difference lies in the fact that these other processes can attract new users to the platform in very specific moments. That is to say, these other participatory processes might drive the arrival of the *aimless petitioners*.

Table 5.1. Participatory budgeting (PB), voting processes and campaigns held at *Decide Madrid* since the platform was launched until January 1, 2018.

Type of process	**Description**	**Start date**	**End date**
PB (1st edition)	Submission of proposals	2016/02/22	2016/03/31
PB (1st edition)	Prioritization of proposals	2016/04/01	2016/05/14
PB (1st edition)	Final voting for proposals	2016/05/15	2016/06/30
PB (2nd edition)	Submission of proposals	2017/01/18	2017/03/08
PB (2nd edition)	Prioritization of proposals	2017/03/11	2017/03/25
PB (2nd edition)	Final voting for proposals	2017/05/15	2017/06/30
Voting	La Gran Votación	2017/02/13	2017/02/19
Voting	Once Plazas	2017/10/08	2017/10/22
Campaign	First advertising	2015/09/15	2015/09/24
Campaign	Second advertising	2015/11/03	2015/11/06
Campaign	Mailing	2016/06/13	2016/06/15

Dataset

To analyze petition signing in *Decide Madrid*, we have collected a dataset from the open data portal of the City Council of Madrid[7,8]. In total, we have obtained 20,131 petitions and 2,564,497 signatures between September 15, 2015 (first day of *Decide Madrid*) and January 1, 2018. We have reviewed the existing participatory processes in the platform during the period of the dataset, i.e., participatory budgeting and voting processes. Also, decision-makers of the City Council informed us that calls for participation were carried out through advertising campaigns. Start and end dates of all the processes and campaigns are presented in Table 5.1.

[7] `https://datos.madrid.es`

[8] We should note that the names of signatories are not public nor included in the dataset.

5.3 Petitions and signatures over time

We first explore the dataset of *Decide Madrid* by examining the distribution of petitions and signatures over time, presented in top and middle graphs in Figure 5.2, respectively. In both graphs, we observe spikes of activity emerging on specific dates. Note that most activity peaks coincide with other processes in *Decide Madrid*, indicated through color bars. The number of signatures on dates when no other process is held is relatively low (avg=1,511; SD=2,101). In contrast, the number of signatures notably increased during the three advertising campaigns (green color bars): first campaign (avg=19,718; SD=10,089), second campaign (avg=19,608; SD=2,347), mailing campaign(avg=22,022; SD=10,099). As observed a the purple color bars, the increase is also relevant for 'La Gran Votación' voting process (avg=23,564; SD=5,572) and, to a lesser extent, 'Once Plazas' voting process (avg=4,020; SD=2,254). Regarding participatory budgeting, the increase is also evident for both the first edition (avg=5,833; SD=5,461) and the second edition (avg=5,375; SD=6,574). However, we find of interest that peaks during participatory budgeting phases often emerge on the first and/or last date of the corresponding phase, which might also be related to calls for participation.

Despite these patterns, note that some activity peaks do not coincide with these processes, e.g., the peak of 21,179 signatures on October 26, 2016. We inspected the petitions signed on that date and found that most signatures targeted a specific subset of petitions published by a human rights organization[9]. Therefore, this should be the result of a campaign orchestrated by such organization.

To better understand how signatures distributed among petitions, the bottom graph in Figure 5.2 shows a scatter plot of the number of signatures on a specific date (horizontal axis) to petitions of a specific age (vertical axis). We observe that much activity is concentrated on the base, that is to say, many signatures are made on the same day

[9] `https://decide.madrid.es/users/180346`

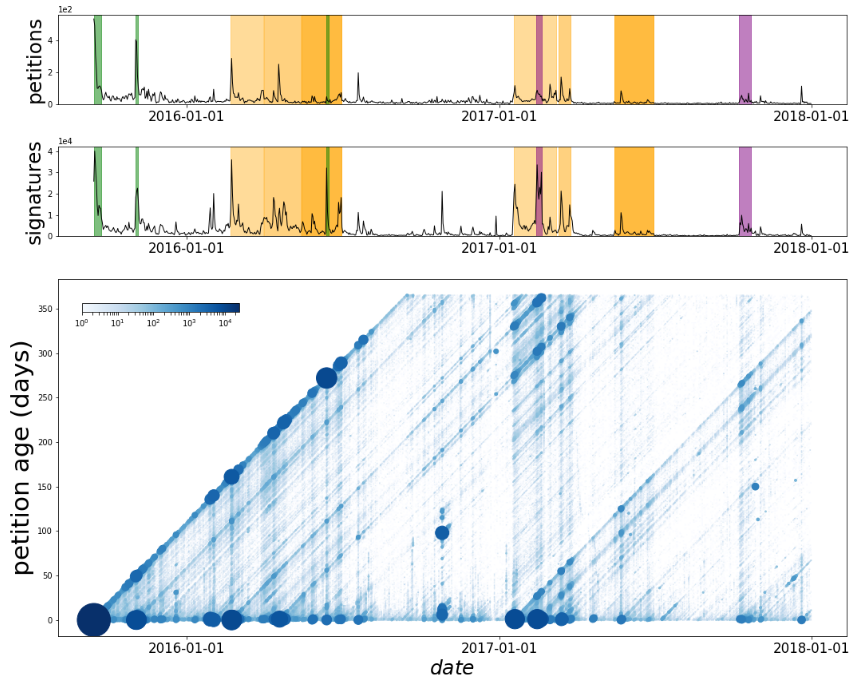


Figure 5.2. Distribution of petitions and signatures by date (top and middle graphs). Color bars indicate when other processes were held: advertising campaigns in green, voting processes in purple, and participatory budgeting phases in a progressive orange scale. Distribution of signatures by signing date and petition age (bottom graph).

the petitions were published. This pattern indicates a strong attractiveness of recent petitions, in agreement with the purpose of the home page ranking petitions based on the *Hot score*. Interestingly, we should notice the existence of diagonal stripes of activity starting on dates with much activity. They correspond to petitions that were created on a day with high activity and kept attracting many signatures over time. They may appeared in the home page due to a large initial number of signatures, and then became highly ranked in the global raking of *Most Popular Petitions*.

The scatter plot also reveals that the peak of signatures to petitions of the human rights organization (October 26, 2016) did not trigger any diagonal stripe. That is to say, despite the notable increase of signatures to petitions from the Human Rights civic organization, none of these petitions gathered the necessary number of signatures to rank high among the *Most Popular Petitions*. Compared to the advantage of being published during a peak of activity, this finding illustrates the little effect of orchestrating a call for signing petitions that are no longer recent or popular.

5.4 Growth patterns of petition signing

We now compare the growth patterns of individual petitions. To do this, we need to define groups of petitions that have similar profiles. We relate petitions by means of the similarity between their associated time-series of signatures according to the Dynamic Time Warping (DTW) metric [BC94][10]. The idea behind Dynamic Time Warping (DTW) is to find the minimal amount of stretching or 'warping' necessary to turn one time-series into another, even if they happen at different time steps. To calculate the DTW distance between two time-series $X=(x_1, x_2, ..., x_n)$ and $Y=(y_1, y_2, ..., y_m)$, it is needed an n-by-m cost matrix where the (i, j) element is the difference between x_i and y_j. A warping path $W=(w_1, ..., w_l)$ is defined as $w_k=(i, j)$ with $k \in [1 : l]$, $\max(m, n) \leq l < m + n - 1$ satisfying three conditions:

- *Boundary:* $w_1 = (1, 1) \wedge w_l = (n, m)$.
- *Monotonicity:* $\forall k \in [1 : l-1], w_k=(i, j), w_{k+1}=(i', j'), i \leq i' \wedge j \leq j'$.
- *One step size:* $\forall k \in [1 : l - 1]$, $w_{k+1} - w_k \in \{(1, 0), (0, 1), (1, 1)\}$.

[10] Code at `https://github.com/shunsukeaihara/pydtw`

While there are exponentially many warping paths over the cost matrix, the DTW distance corresponds to the one minimizing the warping cost: $\mathrm{DTW}(X,Y) = \min\{\sum_{l=1}^{L} w_l\}$.

DTW allows to define groups of petitions that have similar profiles. In particular, we identify similar petitions using hierarchical agglomerative clustering [Mül11] on the matrix of DTW distances between every pair of petitions[11]. To define the resulting clusters, we establish the maximum distance between clusters which cuts the dendrogram by applying the so-called 'elbow' method [Tho53]. To ensure a meaningful separation between clusters, we also validate that the distance between clusters is consistently larger than the avg $\pm$ the standard deviation of the distance between time-series within the same cluster.

5.4.1 Comparing the raw profiles

The results of the DTW-based clustering are presented in Figure 5.3. For clarity, we show the top 100 petitions only. However, we also examined results for petition pools of different size. Interestingly, the lower the pool size, the greater the concentration of popular petitions published on dates coinciding with other processes. As shown in Figure 5.3, we obtain different clusters. The two first ones (m1 and m2) correspond to the only two petitions that reached the signing threshold (dotted horizontal line)[12]. These two petitions exhibit clear abrupt increases when other processes occur in the platform. Note that they were featured in the yellow banner (between October 2015 and September 2016) and benefited from having preferential visibility when many new users, influenced by the other processes in *Decide Madrid*, visited the platform. Therefore, hosting these other processes had a spillover effect on petition signing. Indeed, the third most popular petition followed a similar growth until the banner dis-

[11] Code at `https://docs.scipy.org/doc/scipy/reference`

[12] Proposals at `https://decide.madrid.es/proposals/9` and `https://decide.madrid.es/proposals/199`.

played two petitions instead of three in January 2016. This third petition is part of the cyan cluster (m3) formed by petitions among the first positions of the global *Most Popular Petitions* ranking. Petitions in the other remaining clusters are far from the signing threshold.

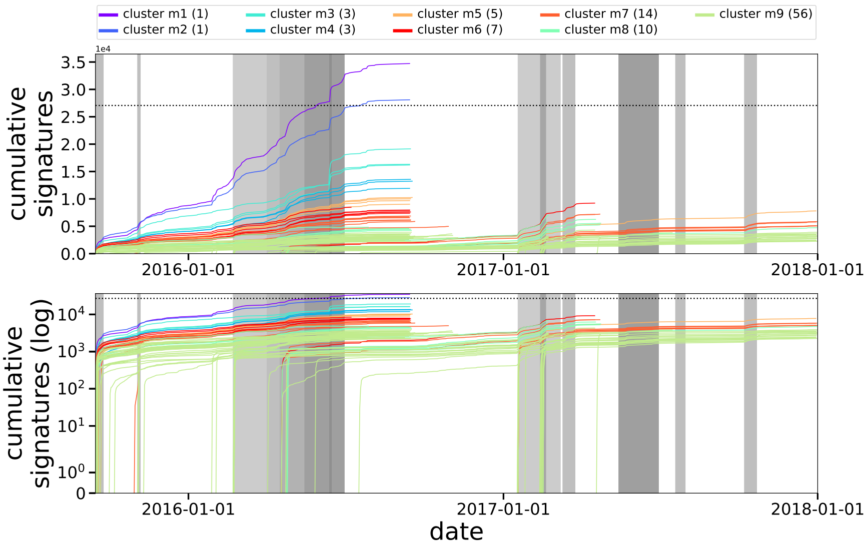


Figure 5.3. Clusters of time-series of signatures for the 100 most signed petitions from *Decide Madrid* with a linear scale (top) and log scale (bottom). Cluster sizes are in parentheses. For clarity, color bars (other processes) are gray-colored. The dotted horizontal line is the 27,064 signatures threshold.

5.4.2 Comparing the normalized profiles

An alternative comparison between the signature profiles can be made after correcting for the effect of the events and normalizing both on the temporal dimension and the number of signatures. This normalization allows to compare the petition profiles in terms of their shape.

For that, we consider as a unit of time any arrival of a signature, instead of a day, and subsequently normalize the temporal axes to one. This effectively stretches the busy peaks and shrinks the periods of low activity, and allows to compare directly the petition profiles. In addition, we also normalize the petitions by the individual signatures, i.e, the total number of signatures each of them gathered individually. We then apply the clustering method on the two normalized (from 0 to 1) dimensions of growth and time.

The left graph in Figure 5.4 shows the major clusters of the petitions started on the first day of *Decide Madrid* that eventually reached 100 or more signatures[13]. Remarkably, we find a clear distinction between two major clusters (dark and light blue), indicating two differentiated dynamic processes for petition growth. The first cluster of 103 petitions corresponds to a rapid growth (dark blue), i.e., from the first signature given to a petition, more than 50% of the signatures mostly occur on the following signatures to any petition. In contrast, the second cluster of 129 petitions (light blue) shows a more constant growth corresponding to a linear trend.

How do petitions from the 'rapid' and the 'constant' clusters differ? To answer this, the right graph in Figure 5.4 shows the complementary cumulative distribution function of petitions from the two clusters by the number of signatures. Petitions from the 'constant' cluster (solid light blue line) are more effective in gathering signatures than petitions from the 'rapid' one (dashed dotted dark blue line). Therefore, this observation suggests an interplay between having many signatures (i.e., petitions easily found in the global *Most Popular Petitions* ranking) and exhibiting high attractiveness over time (constant growth).

[13]We also assessed that results are stable considering different thresholds.

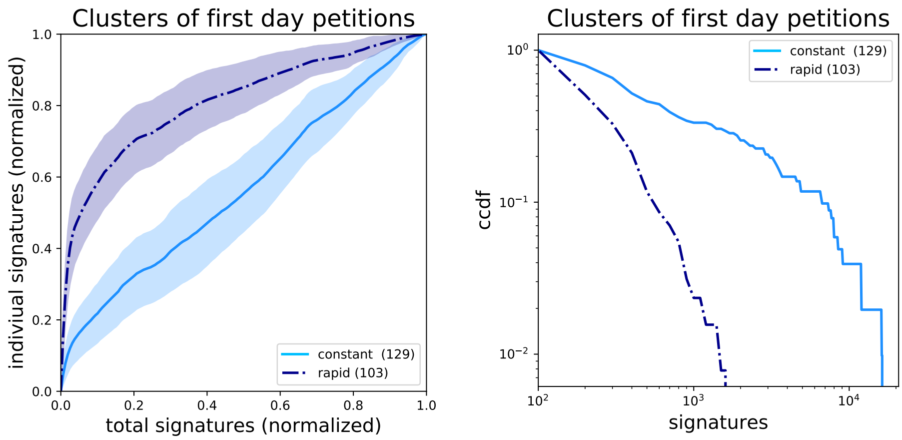


Figure 5.4. Major clusters of growth patterns of petitions published on the first day of *Decide Madrid* (left graph). Cluster sizes are in parentheses. The growth patterns are defined using the arrival of a new signature to any petition as time unit and normalizing both dimensions from 0 to 1. Lines indicate the avg. value while areas the avg. value $\pm$ the standard deviation. Complementary cumulative distribution function of petitions by number of signatures for the two major clusters (right graph).

5.4.3 Comparing topical clusters

So far, we have analyzed the growth of petitions regardless of their content. However, previous studies found that some topics are more effective in gaining attention from signatories [HHU+15, EDK16, CLM17, MGSP17]. For this reason, we finally group petitions based on their topic. To do this, we apply a topic-based clustering process which was proposed in a recent study aiming at grouping petitions precisely from *Decide Madrid* [LP19]. The process is as follows. First, we generate a corpus of documents, one document per petition, with the adjectives, nouns and verbs from the title and body text based on the results of a Spanish POS tagger[14] (we also filter Spanish stop words and lemmatize the remaining ones with a Spanish Wordnet stemmer[15]). Second, we build a topic model based on Latent Dirichlet Allocation (LDA) [BNJ03]. We examine different number of topics and observe than the coherence metric is maximized at 50, which is the value chosen by [LP19]. Third, for each document, we generate a d-dimensional vector ϕ ($d = 50$) with the corresponding document topic probabilities: $\phi = (\phi_1, \phi_2, ..., \phi_d)$. This allows us to compare petitions through their topical similarity. For each pair of petitions p and p', we compute the Jensen-Shannon divergence between their topic vectors ϕ' and ϕ'' as:

$$d_{JS} = \frac{1}{2}\left[\sum_{i=1}^{d} \phi_i' \ln\left(\frac{2\phi_i'}{\phi_i' + \phi_i''}\right) + \sum_{i=1}^{d} \phi_i'' \ln\left(\frac{2\phi_i''}{\phi_i' + \phi_i''}\right)\right].$$

Fourth, we build an undirected weighted full graph comprising a set of nodes (petitions) and a set of edges (similarity between any pair of petitions). To reduce the large number of edges and to only consider edges between very similar petitions, we remove every edge with weight less than 0.01. The resulting network comprises 26,383 nodes and 96,814 edges. Finally, we apply the Louvain method [BGLL08] to detect communities, i.e., topical clusters of petitions (modularity $Q = 0.877$).

Figure 5.5 shows the 10 largest clusters of the network of petitions. To understand the topics of each cluster, we also show tagclouds with the words from petitions of each cluster using the same colors (see Figure 5.6). We examine the words of these clouds and propose the following cluster labels: 'parks and sports facilities' (gold), 'cultural and social facilities' (blue), 'urban mobility' (orange), 'parkings' (green), 'street cleaning/naming' (red), 'public works' (cyan), 'buses & pools' (black), 'bike lanes' (gray), 'subway' (brown), 'public transport & dog poop' (purple).

[14]Code at `https://polyglot.readthedocs.io/en/latest/POS.html`
[15]Code at `https://www.nltk.org/_modules/nltk/stem/wordnet.html`

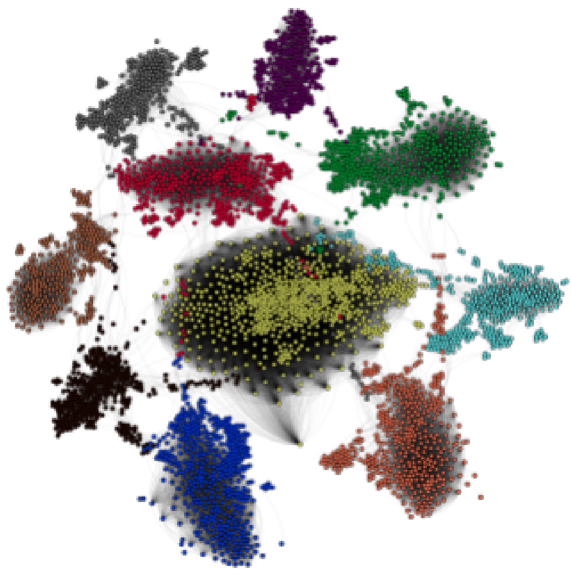


Figure 5.5. 10 largest clusters of the network of petitions linked by edges weighted by their topical similarity. Node color indicates the cluster obtained through the Louvain method. The network is displayed using the Force Atlas2 layout [JVHB14].

The comparison of petition growth across topical clusters is presented in Figure 5.7. For each cluster, we show the distribution of the final number of signatures of petitions using boxplots (top graph). We observe that petitions from clusters like 'public transport & dog poop' or 'bike lanes' usually gather more signatures than petitions from clusters like 'urban mobility'. However, the 75th percentile of daily growth of petitions from each cluster (bottom graph) shows no increase after the first days. Therefore, although some topics might be more appealing to signatories, these differences of attractiveness are present in the first days of petitions, i.e., when petitions are recent and hence easy to find since they are located in the first page of the home page of *Decide Madrid*, based on the *Hot score* ranking.

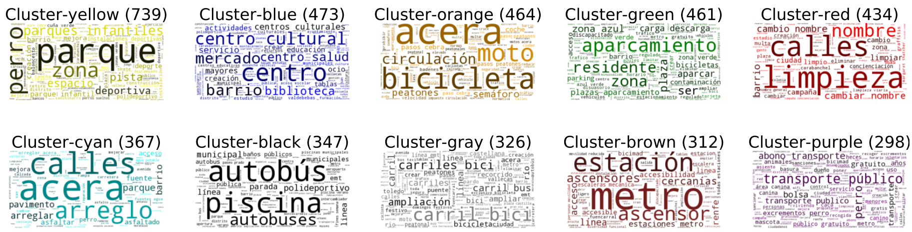


Figure 5.6. Tagclouds of with the words from petitions of the 10 largest clusters. Cluster sizes are in parentheses.

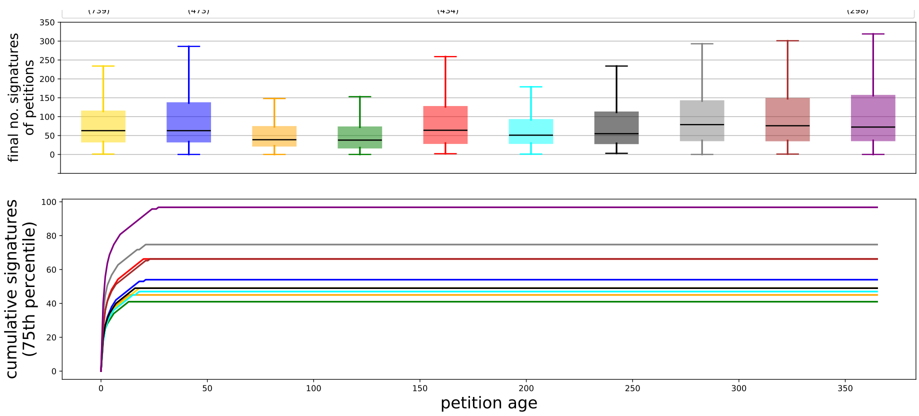


Figure 5.7. Distribution of the final number of signatures of petitions of each cluster with boxplots (top). 75th percentile of daily growth of petitions from each cluster (bottom).

5.5 A new ranking algorithm for online petitions

Our data-driven analysis of *Decide Madrid* has indicated that signing activity largely focuses on recent petitions and few petitions were able to receive signatures over time. These were the two petitions featured in a yellow banner and, to a lesser extent, some petitions that received many signatures on their first days, usually coinciding with another process bringing many *aimless petitioners* to the platform at that time, and hence easily reaching a high position at the global ranking of *Most Popular Petitions*.

We reviewed the open source code of *Decide Madrid* and found that the *Hot score* H_p of a petition p is a trade-off between its activity in terms of comments c_p and signatures s_p, and the recency of that petition t_p, or elapsed days between the petition was originated and the first day of the platform. In particular, it is computed as:

$$H_p \propto \frac{\log_{10}(\max(1, s_p + w \cdot c_p)) \cdot s_p}{\max(1, s_p + w \cdot c_p)} + t_p \tag{5.1}$$

where w is a constant equal to 0.2. When the number of comments is low, which is the case, $\langle c_p \rangle = 3.70$, the previous score can be approximated roughly by $\hat{H}_p \approx \log_{10}(s_p) + t_p$. This means that once a petition was published, it must keep an exponential growth over time. Otherwise, the petition will be replaced in the home page by petitions from subsequent days. We have examined this empirically by inspecting the 6 existing captures of the home page of petitions in *Decide Madrid* at the Internet Archive[16] between 2016 and 2017: 135 of the 150 petitions (25 petitions per home page capture) were at most 3 days old (avg=1.54, SD=1.251).

According to decision-makers in *Decide Madrid*, the reason for adapting the *Hot score* was a past successful experience when deciding to use *Reddit* as the online forum for a Spanish grassroots party [17]. While this score might be valid for sorting news posts under discussion, our results suggested that it is definitely not adequate in this context: petitions have a lifespan of a year in this case, in contrast, the dynamics of reactions to news posts are characterized by fast and short lifespans [LBK09, KGL07]. In the end, the only two petitions reaching the threshold of signatures were promoted during

[16] `https://web.archive.org/web/*/http://decide.madrid.es/proposals`
[17] `https://www.reddit.com/r/podemos/`

months in a banner and received preferential visibility during peaks of attention.

To overcome these drawbacks, our results motivated two changes in the platform that were released on December 11, 2018[18]. First, the yellow banner has been reintroduced to guarantee that users will easily find the three most popular petitions. Second, the new default score for sorting petitions in the home page is computed as:

$$H'_p = \frac{s'_p}{\min(30, t_{now} - t_p)}, \tag{5.2}$$

where s'_p is the number of signatures to petition p *in the last 30 days*. The new ranking was aimed at reinforcing petitions receiving much signing in the last month with less relevance of their publication date.

5.5.1 Characterizing the impact of the new ranking

To assess the impact of the new home page design of *Decide Madrid*, we have extended the dataset with the existing petitions and signatures from the open data portal of the City Council of Madrid until July 1, 2019. Also, we have compiled the dates of all the campaigns and participatory processes in the platform during this extended period, e.g., new editions of participatory budgeting.

We first explore the signing activity once the new ranking was released. Figure 5.8a shows a scatter plot analogous to the graph in Figure 5.2. Once again, peaks in this period largely coincide with other participatory processes in *Decide Madrid*, in particular, phases of the last edition of participatory budgeting. Nevertheless, we observe that the activity is no longer concentrated on the base (i.e., recent petitions) but distributed across different petition age values.

On the one hand, the intervention was aimed at better promoting petition growth by ranking in the first positions not recent petitions,

[18] Pull Request at `http://bit.ly/PRHotv2`

but those that generate interest recently. On the other hand, the global ranking of *Most Popular Petitions* persists as an alternative sorting criterion. Periods coinciding with other participatory processes are characterized by the arrival of *aimless petitioners*. Since these participants do not access *Decide Madrid* to sign a specific petition, they are expected to be particularly susceptible to petitions displayed at the home page at that time.

Does the attractiveness of petitions differ after changing the ranking of the home page? The first phase of the last edition of participatory budgeting started on November 12, 2018, that is to say 29 days before the intervention was performed. This allows us to compare the most signed petitions from that starting date to the intervention, with the most signed petitions from the subsequent period of equal duration. In this way, we are able to compare two equally long intervals characterized by *aimless petitioners* generating much activity. In particular, for each day of each interval we identify the 100 most signed petitions and categorized them as:

- *popular*: petitions in the first page of the ranking of *Most Popular petitions* at that time,
- *recent*: petitions which are at most 3 days old at that time,
- *other*: none of the above.

Results are presented with bloxplots in Figure 5.8b. We clearly observe that, while the attractiveness of popular petitions remains stable, there is a decrease in attractiveness of the recent ones favoring other petitions.

To understand the implications of this shift of attention, we compare petition growth before and after the intervention. Note that, with the original ranking based on the *Hot score* (H_p), for petitions to grow over time, they had to obtain many signatures when they were recent and easy to find, which allowed them to reach a high position at the *Most Popular Petitions* ranking. As a consequence, after releasing the home page with the new score (H'_p), it is expected that petitions that have gathered many signatures in the first day without reaching a high position in the global ranking of *Most Popular Petitions* have grown more constantly and have therefore receive more signatures.

To reduce selection bias while comparing petitions before and after the intervention, we apply propensity score matching [RR83] using a continuous covariate: the number of signatures on the first day; and a categorical covariate: the resulting cluster applying the topic-based clustering process in the extended dataset. The dependent variable of the logistic regression equals 1 if the petition was published after the intervention, otherwise 0. We exclude petitions published in the 365 days prior to the intervention to avoid those that were affected during their lifespan. Also, we should note that the age of petitions published after the intervention is less than a year. As a consequence, for petitions before the change, we consider the number of signatures at the age of the corresponding matched petition on July 1, 2019. In addition, we validate that matched petitions are significantly balanced across the two covariates. For the categorical one, we use a χ^2 test for independence before and after matching, resulting in p=0.99. For the numerical one, we use a Kolmogorov-Smirnov (KS) test on 1,000 permuted samples of petitions obtaining, after matching, p=0.945. Therefore, we verify the failure to reject the null hypothesis in both cases (i.e., $p > 0.05$). As suggested by [BPW+15], we also validate that the standardized mean difference of the numerical covariate decreases after matching (from 0.33 to 0.03).

We present in Figure 5.8c bloxplots comparing the number of signatures to the 849 resulting petition matches stratified into 3 strata according to the number of signatures in the first day: $[0, 24]$, $[25, 49]$, and ≥ 50. We observe that, although petitions with less than 25 signatures on their first day now receive slightly fewer signatures, petitions that collect many signatures on the first day are able to obtain more signatures once the intervention was performed.

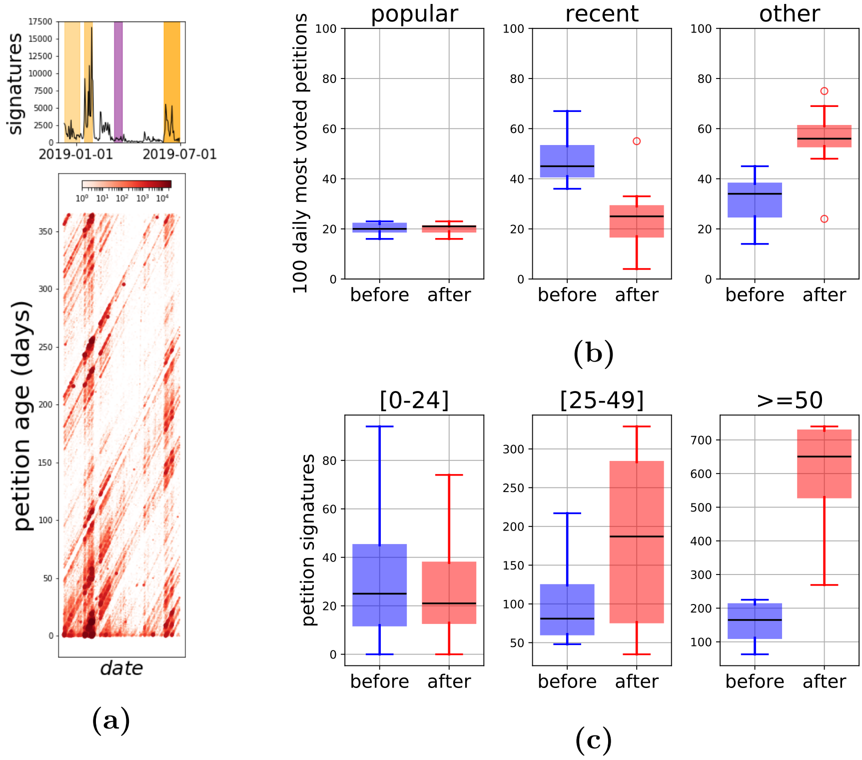


Figure 5.8. a) Distribution of signatures after the intervention by date (top graph), and by signing date and petition age (bottom graph). b) 100 daily most signed petitions before and after the intervention. c) Signatures of petitions, before and after the intervention, matched with the propensity score based on two covariates: the number of signatures in its first day and the topical cluster. Results are stratified into 3 strata according to the number of signatures in the first day: $[0, 24]$, $[25, 49]$, and ≥ 50.

Finally, we compare the growth of petitions from our matching. In particular, we assign each matched petition that gathered at least 10 signatures on their first day to either the 'constant' or the 'rapid' cluster. The assignment is based on the lowest DTW distance from the petition normalized profile to the average growth of each cluster from Figure 5.4. For petitions before the intervention, 41 (33%) correspond to the 'rapid' cluster while 83 (67%) to the 'constant' one. In contrast, every petition after the intervention corresponds to the 'constant' cluster of petition growth. In conclusion, for petitions exhibiting attractiveness among signatories, the new ranking at the home page of *Decide Madrid* favors petitions growing in a more constant way so that more new signatures can be gathered.

5.6 Discussion

The analysis of petition signing in *Decide Madrid* presented in this chapter has revealed important findings with relevant implications for the design of civic technologies.

Previous studies have typically relied on features from petitions [HMY13, DLH$^+$15, HHU$^+$15, EDK16, CLM17, MGSP17]. In contrast, this study has also consider platform features for a more comprehensive characterization. We have observed that both petition creation and petition signing exhibited peaks of activity that were often coupled with other co-existing processes in the platform: advertising campaigns, annual participatory budgeting, and voting processes. Moreover, the most popular petitions in *Decide Madrid* were usually published in these specific peaks. This finding directly relates to previous studies that characterize spillover effects on existing petition platforms. For example, the analysis of the platform of the German Parliament [JJ10, SJ14] found that the attention to popular benefited the less popular ones. In particular, petitions coinciding in time with a successful petition were able to obtain almost twice as many signatures per day. Our findings suggest a similar interpretation which should raise awareness on spillover effects when designing ‘multi-process’ participatory platforms for direct democracy.

The home page of *Decide Madrid* originally presented petitions in a paginated ranking based on an adapted version of the *Hot score* from *Reddit*. Our results have shown that, while this was the platform design, signing focused to a great extent on recent petitions. Furthermore, the clustering of growth patterns has revealed that some petitions in *Decide Madrid* follow the so-called ‘rapid rise and decay’ pattern [YHM17], but that others grew in a ‘constant’ manner. Interestingly, the latter were more effective in gathering many signatures. In addition to this, the topic-based clustering of petitions showed that petitions about some topics tend to receive more signatures, which is consistent with previous studies [HHU+15, EDK16, CLM17, MGSP17]. However, since our analysis not only examines the final number of signatures but also petition growth, results show that these differences mostly occur when petitions were recent.

The rapid rise and decay of petition signing has been defined [MJHY15] as a challenge for the understanding of modern forms of collective action since it does not follow the *S-curve* classically found from economics and political science [Sch78, Gra78]. This phenomenon, validated with petitions from UK and US government platforms [YHM17], has been attributed to the accelerated nature of online environments. Nevertheless, our study shows that the design of the home page plays a crucial role. The intervention performed in *Decide Madrid* in December 2018 significantly reduced the over-exposure of recent petitions by favoring the growth of petitions that are capable of drawing the interest of new signatories.

5.7 Conclusion

The rise of civic technologies to scale up citizen participation comes with great potential for democracy. So far, these systems have been designed favoring transparency and open data standards. However, our analysis has shown how the platform design can originate very different participation dynamics, and suggests some negative consequences of re-using inadequate algorithms. Therefore, our results should motivate more critical studies on the role that civic technologies plays in fairness, accountability, and transparency. We advocate for a more principled approach, for example, by means of causal models or randomized trials to analyze and design current systems. This is a major challenge, mainly because of ethical considerations and the introduction of additional experimental biases. In conclusion, we believe that our work contributes to develop a framework that addresses critical aspects of platform-based democracies.

6 Deliberative Platform Design for Online Petitions

6.1 Introduction

The crisis of representative democracy in the last three decades [RG08, Tor15] has been identified with the crisis of democracy itself [DP13, Kea09]. Some authors have criticized the technocratic tendencies operating in this period as signs of the rise of post-democracy [Cro04] or post-politics [RPB01, Žiž00], while others, more precisely, have used the term "post-representation", to refer to the emptying out (of power and meaning) of representative institutions by dynamics ranging from globalization to growing citizen mistrust [BVR08, Kea09]. Specially in the last years, this political crisis has led to a period of fertile democratic innovation supported by an intensive and creative use of information and communication technologies [Cas09, TCLM+15]. Thus, we are witnessing new forms of participatory and deliberative democracy based on computer mediated communication [Fuc07, HL99].

One of the recent institutional instantiations of this wider democratizing process is *Decidim Barcelona*[1], an online platform developed by the Barcelona City Council for supporting its participatory processes, e.g., the development of the Barcelona's strategic city plan. The strategic city plan defines objectives and actions to be carried out by the local government during the present legislature. The goal of this participatory process was to enroll the citizenry in a two month process of co-production, where citizens could discuss and support the petitions made by the government; and also make, discuss and support their own petitions. In total, more than 40 000 citizens participated in this process.

According to the functional specification of *Decidim Barcelona* [MCLPdL15], different pre-existing tools for participatory democracy were assessed, in particular, *e-Petitions Gov UK* (United Kingdom)[2], *Your Priorities* (Iceland)[3], *Cónsul* (Madrid)[4], and *Open Irekia* (Basque Country)[5]. On the one hand, these four tools share certain commonalities. First, they are web applications based on Free/Libre and Open Source Software (FLOSS). Second, they have been deployed in real environments by city, regional, or national governments. Third, they allow users to make online petitions. On the other hand, there are many differences among these four platforms. An important one is the way petitions are discussed by users. In *e-Petitions Gov UK*, petitions cannot be discussed and, therefore, this tool might be considered as enabling participatory but not deliberative democracy. *Your Priorities* allows users to publish comments either supporting the petition (hereafter *positive comments*) or against it (hereafter *negative comments*). Positive and negative comments are displayed in two columns and sorted by the number of votes they receive to show the best arguments and, ultimately, to facilitate decision making.

[1]`https://www.decidim.barcelona/`
[2]`https://www.gov.uk/petition-government`
[3]`https://www.yrpri.org`
[4]`https://decide.madrid.es/`
[5]`http://www.irekia.euskadi.eus/`

Although this strategy relies on comments, users do not engage in discussions, which might reduce the deliberative capabilities of the platform. In contrast, *Cónsul* corresponds to an opposite scenario given that users are able to discuss any petition with a threaded interface without any visual indication of whether comments are positive or negative. Finally, the approach in *Open Irekia* allows users to indicate whether a comment is positive, negative or neutral. However, neutral, positive and negative comments are presented separately without applying a threaded discussion interface, as done in *Cónsul*. This heterogeneity received special attention in the design specification process of *Decidim Barcelona* [MCLPdL15] resulting in an interface which hybridizes the previous approaches. On the one hand, petitions are discussed in a threaded interface to promote online discussions and, consequently, online deliberation. On the other hand, users are able to establish when posting a first level comment (i.e., a direct comment to a petition) whether is positive, negative or neutral in relation to the petition. In addition, authors of petitions and comments are notified when receiving replies.

Figure 6.1 shows a real petition for the strategic city plan which requested a municipal ice skating rink in Barcelona. The hybrid interface combines both conversation threading and coloring, i.e., positive and negative first level comments include green and red labels, respectively (the interface at that time colored the full text of positive and negative comments). In this way, the discussion page shows two first positive (green) comments with no replies and a third negative (red) comment calling into doubt the adequacy of expending public funding on a winter sport facility in a Mediterranean city. As shown in Figure 6.1, the negative comment triggered a discussion cascade among users.

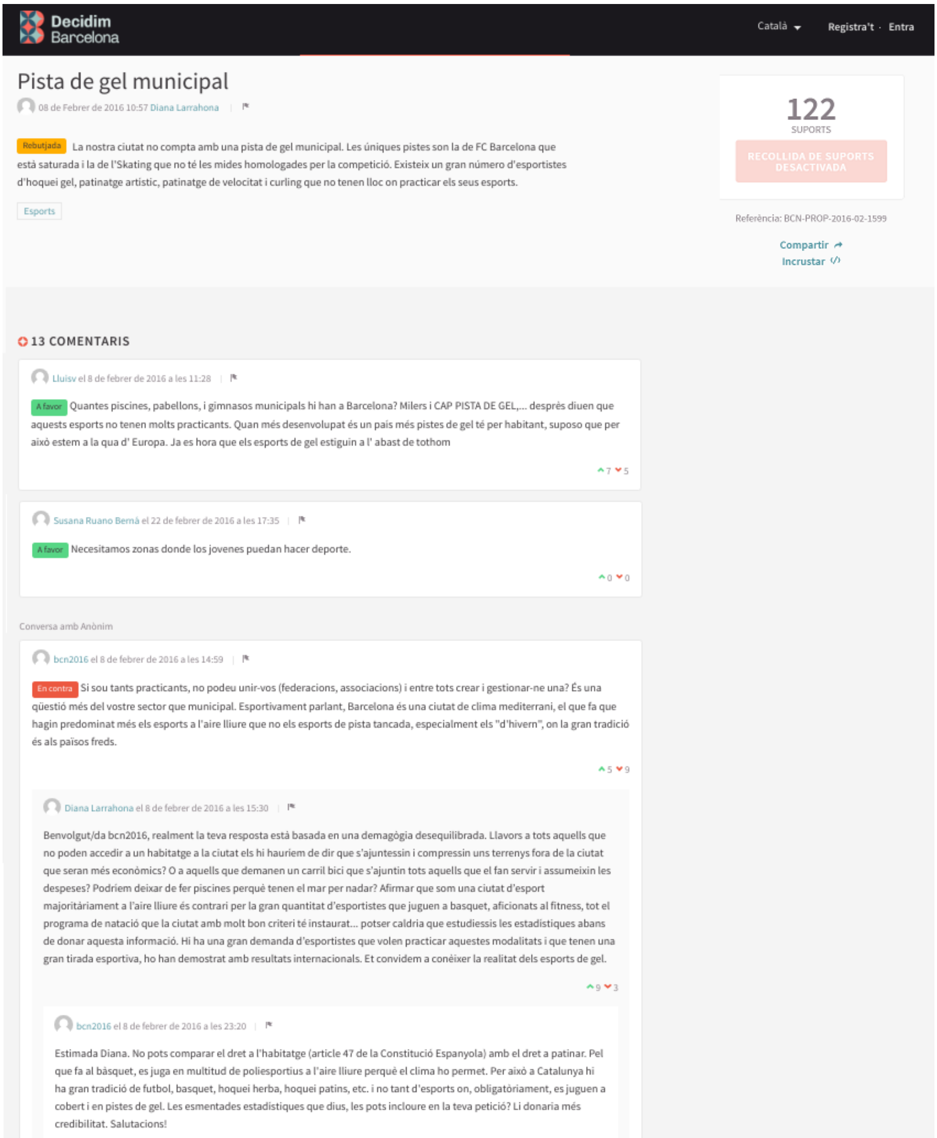


Figure 6.1. Discussion page of a petition in *Decidim Barcelona* for building a municipal ice skating rink.

This petition is an illustrative example of the aim of this hybrid interface: users can engage in online discussion to promote deliberative processes while positive and negative comments are easily distinguishable to facilitate decision making. The combination of both approaches makes *Decidim Barcelona* an interesting case study for multiple reasons. First, we have shown that conversation threading in online discussion platforms promotes the emergence of discussion cascades with higher levels of online deliberation and reciprocity (see Chapters 3 and 4). Second, given that users are able to mark the alignment of comments with the petition (positive, negative and neutral), we can compare the typical network structures originated by the different types of comment alignment. As shown in [GBKB10], these structures can be used as proxies of very basic forms of deliberation.

Given this particular scenario of *Decidim Barcelona*, the research question of this study is as follows:

- *Which are the structural differences of discussion cascades triggered by neutral, positive or negative comments on online petitions?*

As presented in the following section, despite the increasing research work on online petition platforms, how to effectively introduce discussions is an open practical and research challenge [LR09]. We postulate that the combination in *Decidim Barcelona* of both conversation threading and comment alignment (in particular, explicitly negative comments to the petition) should favor cognitive dissonance [Fes62] in users, which would lead to a higher willingness to discuss the petitions, and, therefore, to deliberative practices of decision making.

6.2 Related work

The interest in online petition platforms is reflected by the increasing attention from academia [PE12]. Some of the first studies analysed the platform developed by the German Parliament either to identify different types of users according to the frequency of participation [JJ10] or to characterize the relationship between online participation and offline socio-demographic factors [LR11]. Indeed, much effort has been made to detect which factors affect the signing of online petitions [AGCSM08, HMY13, HSHH15, MJHR15, YHM17].

Previous work has also examined the impact of platform design on the dynamics of online petitioning. A study of the UK government petitions platform showed that introducing trending information on the homepage increased the inequality in the number of signatures across petitions [HJMY18]. In relation to our research question, some papers have precisely assessed the role of the availability and design of discussion features. A study of online petition platforms launched by UK local authorities (Kingston and Bristol) [WRM05] examined the performance of the online forums incorporated in these tools. Results indicated that most users did not visually identify the possibility to discuss petitions and just a few users published comments. Therefore, the study concluded that the discussion section for online petitions needed to be more appealing. A comparative analysis of four online petition systems (the aforementioned platform of the German Parliament and the platforms of the Scottish Parliament, the Parliament of Queensland, and Norwegian municipalities) also examined whether they integrate an online discussion forum [LR09]. Online discussions were available in every platform except for the case of Queensland. The study found little usage of these forums and concluded with the open research question about the function of these discussions and how to channel them into the political decision-making processes.

6.3 Dataset from Decidim Barcelona

Our dataset contains the discussion threads from the petitions in *Decidim Barcelona* for the development of the strategic city plan. To better understand the discussions that originated more activity we present in Table 6.1 the most commented petitions, which are related to controversial topics in Barcelona like housing affordability and mobility.

Data were extracted through the Decidim API[6] to obtain a total of 10,860 petitions and 18,192 comments. 16,217 comments were first level comments (i.e., direct replies to the petition) while 1,975 comments were replies to comments. As mentioned in the introduction, users were able to establish the alignment of first level comments with the petition. Thus, 10,221 comments were marked as neutral (63.03%), 5,198 comments were marked as positive (32.05%), and only 798 comments were marked as negative (4.92%).

6.4 Structural metrics of discussion cascades

Discussion threads are collections of messages posted as replies to either an initial message (the petition) or another message (a comment). For this reason, discussion threads can be represented as a directed rooted tree. We present in Figure 6.2 the petition for a municipal ice skating rink (shown in Figure 6.1) using a radial tree visualization tool (see Appendix B). The black node is the petition (root) and the nodes directly connected to the root are the first level comments (green colored if positive and red colored if negative). This tree structure allows to identify whether a first level comment triggers a discussion cascade, e.g., the red node on the right, which is the negative comment against expending public funding on a winter sport facility in a Mediterranean city, triggers several comments.

[6] `https://www.decidim.barcelona/api/docs`

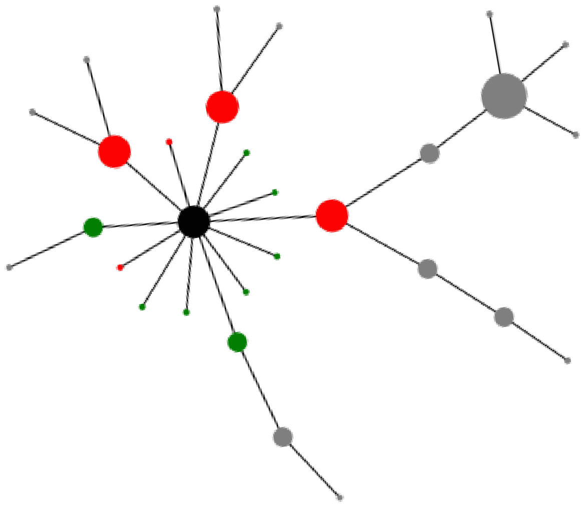


Figure 6.2. Radial tree visualization of the petition presented in Figure 6.1. Black node (root) represents the petition, green nodes are positive and red nodes negative first level comments. Comment nodes are sized by the indegree (number of replies to the comment). The visualization shows a cascade of comments triggered by a negative comment to the petition (red node on the right).

The structure of the discussion cascade of each first level comment can be characterized with typical metrics of tree graphs:

- size: number of nodes,
- width: maximum number of nodes at any level,
- depth: number of levels,
- h-index: maximum level h in which there are, at least, h comments [GKL08].

In the discussion cascade originated by the aforementioned negative comment about public funding (red node on the right in Figure 6.2), size is 9, width is 4, depth is 3, and h-index is 3.

With the exception of the size, which just quantifies the volume of the cascade, these metrics serve to inform about the network topology of a cascade. Morever, the last three metrics have been suggested to quantify the level of deliberation in online discussion threads [GBKB10]. This approach is based on the Madisonian conceptualization of deliberation as the conjugation of two dimensions: representation and argumentation [AF04]. Given that messages at any level often represent users within the discussion, width has been proposed to quantify the extent of representation of the online community in a discussion cascade. Because the exchange of arguments between users commonly occur as exchange of comments, the depth of the discussion cascade (i.e., the largest exchange of comments) has been proposed to capture argumentation. The last structural metric (h-index) both considers width and depth and, therefore, has been proposed to measure online deliberation in a discussion cascade [GBKB10].

6.5 Analysis of discussion cascades

The description of the dataset indicated that most of the first level comments were marked as neutral, an important fraction were marked as positive and just around 5% were marked as negative. To understand the structure of cascades triggered by comments from different alignments, we first examine the distribution of the cascade size depicted in Figure 6.3. We observe a notably lower probability of not triggering a cascade for negative comments. We also observe that, in every alignment, few cascades contain more than five comments.

Figure 6.3 reveals a larger preference for larger cascades triggered by negative comments. However, the size of the cascade is not an informative metric of the structure of the cascade. For this reason, we examine the probability of the alignment of the root (comment) of the cascade with different sizes and different values of the structural metrics (width, depth and h-index). Results are presented in Figure 6.4 using heatmaps, i.e., the darker the more likely. We observe that, if a comment did not trigger any discussion cascade, that comment is probably neutral or positive. In contrast, when comments originated discussion, there is a higher probability that they are negative. Furthermore, the likelihood of negative comments increases when the value of the size and the structural metrics also increase.

These results suggest that discussion cascades occur more frequently due to negative messages and less frequently due to neutral messages. However, to perform a rigorous analysis we need to consider the following observations. First, we found by manual inspection that many neutral comments, despite being clearly positively or negatively aligned with the petition, were not explicitly marked accordingly for some reason, e.g., problems of usability or perhaps a deliberate choice of the user. Second, we have to take into account the class imbalance (5,198 positive vs. 798 negative comments). Because

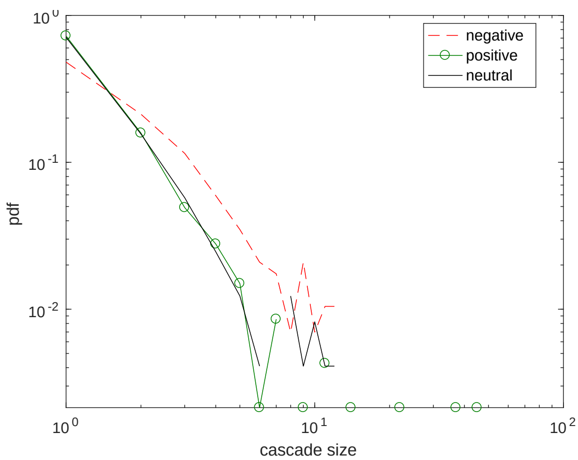


Figure 6.3. Distribution of the cascade size triggered by the first level comments of each alignment (neutral, positive and negative).

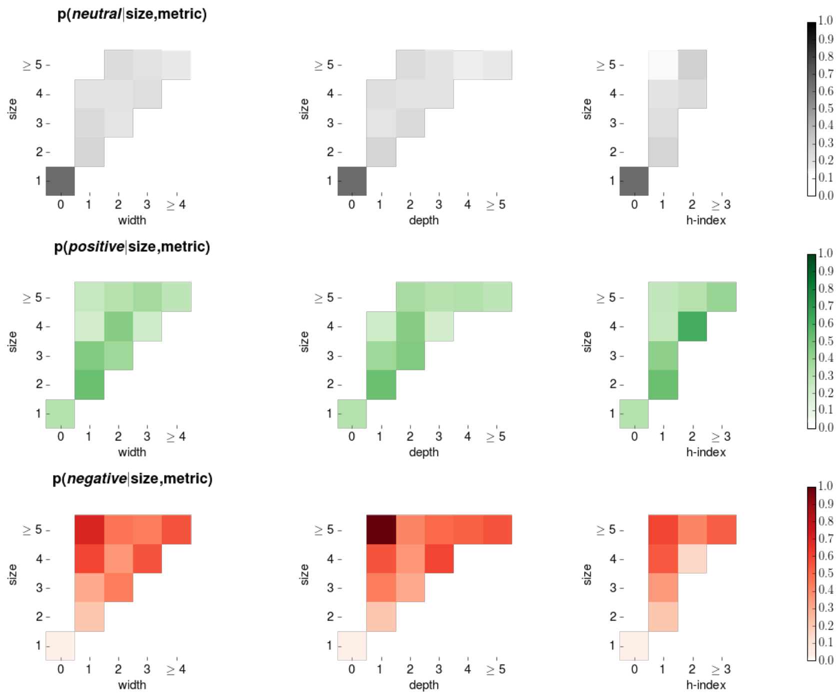


Figure 6.4. Heatmaps of the probability of alignment (gray for neutral, green for positive and red for negative) of a first level comment given size and width, depth, or h-index of the cascade. Large values are aggregated in the top rows and rightmost columns.

of these two reasons, we will restrict our analysis to aligned comments, either positive or negative, which triggered at least one reply. We apply bootstrapping, with 10K evaluations and randomly chosen (with replacement) 10K positive and 10K negative comments. Comments can be chosen more than once. The number of evaluations and threads have been selected, after multiple assessments, to guarantee the significance of the statistical test ($p < 0.05$). Results are presented as heatmaps in Figure 6.5 and confirm that, regarding positive and negative first level comments, when deep and complex cascades are observed, there is a much stronger likelihood to be originated by a negative comment. In conclusion, although we find that positive comments sometimes triggered complex discussion cascades, in general, the deepest and most complex conversations between users in *Decidim Barcelona* were caused by negative comments, i.e., counter-argumentation.

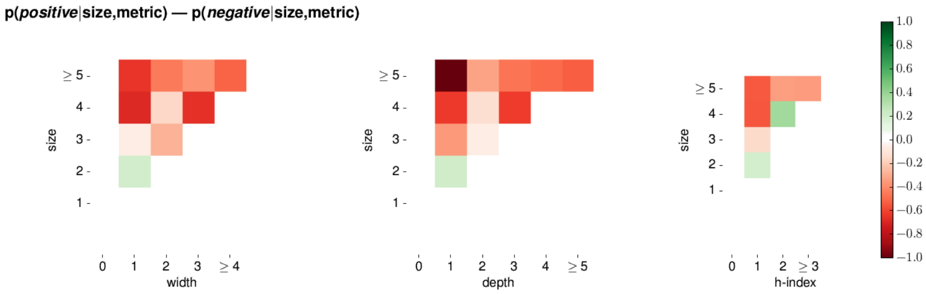


Figure 6.5. Heatmaps of the probability of polar alignment (green for positive and red for negative) of a first level comment given the value of size and structural feature (width, depth, and h-index) of the cascade. Values are obtained with a statistical test of 10K evaluations with 10K random cascades each and shown if significant ($p < 0.05$).

6.6 Discussion

This study has been designed to answer our research question about the structural differences of discussion cascades triggered by neutral, positive and negative comments on online petitions in *Decidim Barcelona*. Our question was motivated by the open research challenge of effectively deploying online discussions in online petition platforms [LR09, WRM05]. The interface in *Decidim Barcelona*, which combines conversation threading and comment alignment, became an innovative case study and an ideal scenario to answer this question. Results are clear: although a low proportion of comments were negative (about 5%), negative comments were more likely to trigger more complex discussion cascades than neutral and positive comments. We should note that users in *Decidim Barcelona* were notified when they received a reply. Therefore, authors of petitions were always aware of negative comments which might also increase their interest in engaging in discussion to advocate for their petitions. This is consistent with the basis of cognitive dissonance [Fes62], i.e., negative comments usually contain new information which contradicts the idea of a given petition and the author and supporters of the petition will be likely to reply to it. We can conclude, thus, when trying to address the open challenge of effectively combining online petitioning and online discussion [LR09, WRM05], the deliberative platform design of *Decidim Barcelona* introduces an innovative solution.

We should remark that our methodology was language-independent. This was a deliberated decision because of the complexity of the bilingual context of *Decidim Barcelona* (Spanish and Catalan), e.g., many natural language processing resources were not available for Catalan. Although this decision allows to easily apply our methodology on any other platform, future work should also focus on the content of messages to compare how linguistic features might also differ in relation to the alignment of comments.

Table 6.1. Top 10 petitions in *Decidim Barcelona* by the number of comments. An English translation is indicated in parentheses.

Title	**N. comments**
Noves llicències per a pisos turístics (New licenses for tourist apartments)	337
Implantar el tramvia a la Diagonal (To build a tramway in Diagonal Avenue)	111
Cubriment de la Ronda de Dalt al seu pas per la Vall d'Hebrón (Roof for Dalt Road in Vall d'Hebrón)	108
Promoció de l'ús de la bicicleta, i millora i ampliació dels carrils bici (Promotion of cycling and improvement and expansion of bike lanes)	80
Regulació del mercat de lloguer (Regulation of the housing rental market)	77
Pla de salut mental (Mental Health Plan)	75
Expropiació i demolició de la parròquia de Santa Maria de Gràcia (Expropriation and demolition of the parish of Santa Maria de Gràcia)	74
Acotar la invasió de gossos en espais públics (To limit the invasion of dogs in public spaces)	68
Pla estratègic de turisme 2016-2020 (Strategic tourism plan 2016-2020)	66
Estratègia per garantir una sanitat pública, universal i de qualitat (Strategy to guarantee public, universal and quality health)	63

7 Conclusions

7.1 Summary of findings

In this dissertation we have characterized online participation in civic technologies, in particular, we have provided readers with practical examples of how technical interventions in real-world platforms were able to influence the behaviour of thousands of participants. These findings have important implications for the design and study of civic technologies.

7.1.1 Detection of platform effects

We examined in Chapter 3 whether it was possible to automatically detect events which affect deliberation in online discussions. We proposed a language-independent methodology based on regression discontinuity design and metrics from network science. Results on the discussions from the social news site *Menéame* showed the influence of online discussion interfaces on the emergence of deliberative network structures. In particular, the change of the conversation view from linear to hierarchical induced deeper discussion threads which are associated with higher argumentation. This was accentuated when the maximal visual depth was increased. Such event was not known while designing the analysis and, therefore, illustrates the flexibility of our methodology for observational studies.

Note that we were able to identify an unknown event because the development process of the open source code of *Menéame* is available at the collaborative platform *GitHub*. In contrast, this was not possible when we detected certain anomalies in a dataset of online petitions from *Avaaz.org* because the source code of that platform is proprietary (see the analysis in Appendix A). Openness, often expressed as the right of citizens to access to relevant files and information, has gradually become a fundamental right in modern democracies [Ale14]. Therefore, given the increasingly prominent role played by civic technologies, the right to study how the program works and understand it (Freedom 1 of the Free Software's Four Freedoms [Sta02]) is essential to guarantee the democratic nature of any political system supported by digital platforms.

7.1.2 The impact of conversation threading

Given the results of Chapter 3, we then quantified the impact of the change of the conversation view from linear to hierarchical in *Menéame*. In particular, we presented in Chapter 4 an analysis of how the reciprocal behaviour of online discussions was affected by this technical intervention. Using interrupted time series analysis and regression discontinuity design, an abrupt and significant increase in social reciprocity was observed after the adoption of a threaded interface.

We furthermore extended the state of the art of generative models of discussion threads [GKLK13] by including reciprocity as a model feature to explain better the typical deep structures of online discussions. Results also confirmed that the adoption of conversation threading induced more reciprocal activity, made popular comments to attract more replies and slowed down the decay of novelty. This finding shows the huge potential of generative modeling approaches to help to assess the inter-dependency between users interaction patterns and platform design elements. In conclusion, such approaches can be exploited to help site owners and community managers to create a positive and constructive environment for large scale online discussions.

7.1.3 The impact of petition ranking

We examined in Chapter 5 data from petitions and signatures detecting platform features that may be hindering citizen participation. First results suggested that the original design of the home page, which was based on a ranking inspired by *Reddit*, was inadequate in this context since petitions are aimed at growing over year, while the dynamics of social news sites are characterized by fast and short lifespans [LBK09, KGL07]. These findings motivated a new sorting criterion for the ranking of the home page which positively affected petition growth in *Decide Madrid*.

Given that most petition platforms include rankings in their home page, our findings have relevant implications for the design of civic technologies. We should note however that the design of some popular non-governmental petition platforms differs from these approaches. On the one hand, petitions in *Change.org* are classified by categories to then show the most popular petition from each category in the home page. On the other hand, the home page of *Avaaz.org* does not present a list of petitions but a list of most recent signatures, regardless of when petitions were published. Given this heterogeneity of home pages, this work should inspire future work involving researchers and civic technologists to reflect on the consequences of the design of home pages for online petition platforms.

7.1.4 Deliberative platform design

In Chapter 6 we focused on the discussion threads of online petitions in *Decidim Barcelona*. Previous works on online petitions found that platforms had little success in engaging participants into discussions [WRM05, LR09]. Petition discussions have traditionally been designed either as comments using an online forum threaded interface, as the one of *Menéame* analyzed in Chapters 3 and 4, or as arguments presented separately according to their alignment for or against the petition. The deliberative platform design of *Decidim Barcelona* combined both strategies and allowed us to characterize the structural differences of discussion cascades triggered by neutral, positive and negative comments. The analysis revealed that few comments were negative but, in comparison to neutral and positive comments, they were more likely to trigger complex discussion cascades. Given that deliberative structures of threads are characterized by complex network structures, the design of petition discussions in *Decidim Barcelona* becomes an interesting approach to favor deliberative practices of decision-making in civic technologies.

7.2 Open research challenges

Based on the findings of this thesis, we propose open research challenges that we have identified and that should be addressed in the future.

7.2.1 Adding content-based features

Most models reviewed in Chapter 2 and the model proposed in Chapter 4 do not include features related to the content of messages within the discussion. The only exception is *Backstrom et al. (2013)* [BKLDNM13] which considers some text-based features like the occurrence of certain terms (e.g. 'comment', 'agree') or the number of question/exclamation marks in a comment. However, this model is rather predictive than generative, i.e., its purpose is not to understand behaviour in discussions but to predict their length. We should note that language-independent approaches for generative models are easily replicable in online discussions of very diverse nature. However, we also believe that only focusing on structural aspects of threads might be limiting when characterizing online discussions, e.g, to fully understand the emergence of online deliberation.

Content-based approaches are mandatory for relevant research topics like modeling antisocial behavior of discussions [CDNML15, CBDNML17, KCL17]. Because modeling online discussion relies on representing discussion threads as information cascades, generative models could be enriched with existing methodologies of emotional cascades of online activity [AGMS15]. In fact, there is evidence supporting that emotional expressions prolong online discussions [CSS+11]. Therefore, understanding collective emotions in online discussion is still a challenging task, requiring generative mechanisms that can bridge individual and collective levels of behavior [SG10]. Besides emotions, the content of messages can also reveal the emergence and evolution of topics in online discussions. For instance, some studies have found strong evidence that hierarchical comment threads represent a topical hierarchy in discussion platforms [WZH13]. Thus, this observation explicitly motivates the inclusion of text-based features (e.g. text similarity between replies) to better characterize the arrival of new comments in a discussion thread. To the best of our knowledge, this approach has never been considered in any generative model of discussion threads and, therefore, adding content-based features is still an open research challenge.

7.2.2 Modeling group behaviour

Homophily and social influence are paradigmatic behavioral phenomena in online interaction and help to explain the emergence of clustered community structures [AMS09]. However, although some generative models of online discussion threads take users into account, none of them considers the existence of groups of users, with common interests or similar opinion about certain topics.

This phenomenon is relevant because many empirical works have found evidence that user groups evolve into echo chambers [Sun09, GBK09, FGR16]. This leads us to reflect on how generative models would better explain online discussion if groups of users were taken into account. In relation to this issue, according to [MSLC01], two of the main research challenges in the topic of homophily are multiplexity, i.e. understanding the role of networks with various layers of interaction types, and analyzing dynamic data in which links appear and disappear over time. Some empirical studies have shown that homophily is stronger in positive interactions rather than in reply interactions [GAS+15]. Thus, the problem of identifying groups of users can be viewed as a community detection problem. This could be handled by the many existing algorithms [For10] on the network of votes among users within the discussion. On the other hand, this could also motivate new methods to detect communities based on interactions (i.e. comments or votes) which only occur between opposing fractions. In fact, the co-evolution of votes and comments is currently receiving increasing attention [CTTJF16]. Therefore, we suggest that the research challenge of modeling online discussions including groups of users (to be inferred from interactions from the voting layer) would provide a better explanation of the behavior of online communities in discussion platforms.

7.2.3 Mitigating information overload

Much of the work in this thesis has focused on understanding the growing process of online discussion threads and online petitions. It is important to note that the growth in online platforms often makes it more difficult for participants to identify messages and ideas under discussion. This phenomenon is even more challenging in civic technologies because discussions are aimed at providing valuable arguments for guiding citizens in participatory decision-making processes.

The knowledge discovery systems described in Appendix B were developed to present information about petitions and discussions from *Decide Madrid* in a concise and well-interpretable way. Other tools like *Wikum*, an experimental technology for crowdsourced recursive summarization [ZVK17], have been tested with citizens and public officers of Madrid for helping them get an overview of the discussions of petitions (see Figure 7.1). In parallel to these prototypes, recent research work has proposed recommender systems to address the exact same problem of information overload in this civic technology [CBCCG17a]. Nevertheless, the real integration and successful assessment of computational tools of this kind remains a pending issue.

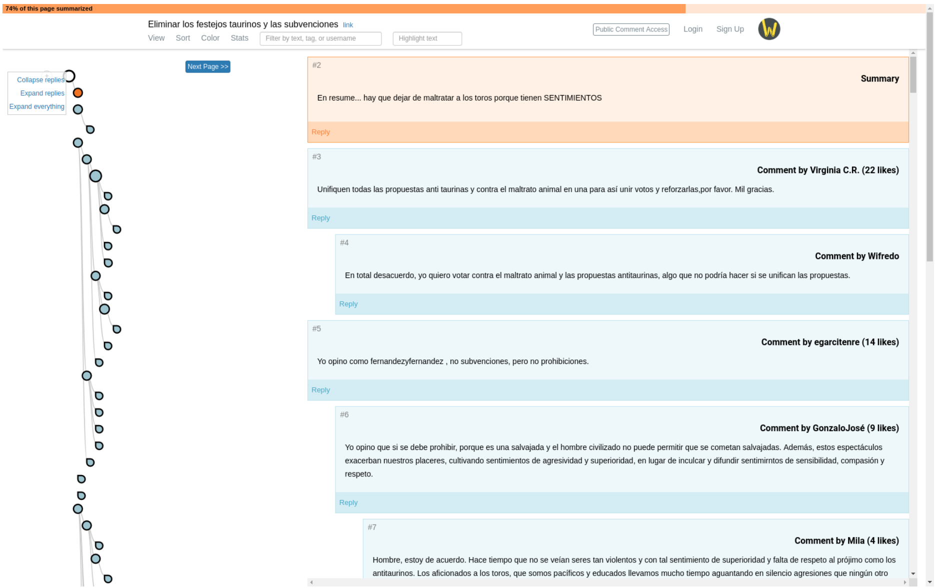


Figure 7.1. Discussion of a petition from *Decide Madrid* using the recursive summarization tool *Wikum* [ZVK17]. `http://wikum.org/visualization_flags?id=386`.

7.3 Final remark

The rise of civic technologies has led to new and innovative ways in which citizens participate democratically. The scope of projects like *Decide Madrid* and *Decidim Barcelona* goes beyond these two cities: Many other municipal governments such as the ones of New York City, Mexico City, Paris or Helsinki have already deployed their own platforms powered by these free open source technologies. Therefore, these civic technologies are already playing a key role in the democratic governance of cities around the world.

Open source allows citizens to audit how technology works. However, we cannot determine which design and which algorithms are best suited for a participatory process only through examining the open source code of a platform. This knowledge requires an exhaustive research process. In fact, most popular online platforms on the Internet are the result of a constant battery of controlled experiments that guides product development and accelerates innovation. In the introduction to this manuscript we indicated the ethical problems associated with experimenting with civic technologies that are being used by real citizens for public policy making. As a consequence, we have advocated for the usage of techniques from causal inference to measure the effect of specific interventions, which were decided by the teams responsible for these platforms and applied to the entire population. Such strategy, while preventing our research from discrimination and manipulation scenarios, is very limiting when it comes to generating relevant public knowledge for a democratic society, increasingly mediated by civic technologies.

There are scenarios to explore between authoritarian experimentation defined by the owners of online platforms and the total absence of experimentation. If the official instances of civic technologies ruled by democratic governments cannot guarantee conditions for ethical experimentation, we need to design and to deploy pre-production environments that can be the object of controlled experimentation and that would allow the generation and transfer of public knowledge.

For example, in recent years, cities around the world have deployed citizen laboratories of many kinds to foster innovation projects, e.g., living labs, hacklabs, medialabs, etc. Actually, *Decide Madrid* and *Decidim Barcelona* projects have been supported by Medialab Prado in Madrid and Laboratori d'Innovació Democràtica in Barcelona. Research and innovation in both spaces with respect to these civic technologies have primarily focused on assessing their performance, discussing key challenges, and exploring similar platforms. However, citizen laboratories would have been ideal pre-production environments for active and citizen experimentation.

Citizen experimentation would open up a participatory and democratic generation of public knowledge for the 'platform society'. This approach must be understood as opposed to the paternalistic approach of nudging, where a privileged group exercises its power by influencing the behavior and decision making of groups or individuals. In the same way that citizen science has shown the potential of the citizenry involved in scientific research, it is essential to reflect on and promote a paradigm of citizen experimentation for democracy.

A

Dataset of Online Petitions from Avaaz.org

A.1 Introduction

Petition signing has long been one of the most popular political activities with a history extending back to at least the Middle Ages [Fox12]. After a decline in the 20th century, petitioning has achieved new prominence through online, e-petition platforms. Online petitions are often disseminated on social media, and their low-costs, low-barriers to entry may bring new people into the political process [MJHY15].

Despite the popularity of petition platforms, this computer mediated form of civic participation has also come under criticism. A review of online petitioning over 10 years in Europe and the United Kingdom concluded that, although these experiences were mostly positive, there was no solid evidence about significant impact [PE12]. Indeed, petition platforms are seen as the essence of the so-called 'slacktivism' or 'clicktivism', where the main effect is not to impact real life but to enhance the feel-good factor for participants [Chr11]. This criticism has fueled a skeptical view of activism through social media [Gla10] and of the promise that the Internet would 'set us free' [Mor11]. Nevertheless, online petitioning has a proven ability to originate public policies of great impact, e.g., the *Unlocking Consumer Choice and Wireless Competition Act* [U.S13].

In the last decade, much research has focused on analyzing data from government platforms in countries like Germany [JJ10, LR11], the United Kingdom [Wri12, HMY13, Wri15], and the United States [DLH+15, MJHY15, YHM17]. As these platforms belong to public institutions, these studies are of interest because of the potential of petitions to influence policy making. However, institutional platforms are restricted by definition to specific territories, and do not take advantage of the global scope offered by the Internet.

At the global level, *Change.org* and *Avaaz.org* are two paradigmatic examples of online petition platforms able to engage millions of people around the world. Nevertheless, there are few empirical studies of these communities. For *Change.org*, launched in 2007 by a for-profit corporation, research has analyzed user behavior [HSHH15], success factors [EDK16] and gender patterns [MGSP17]. These studies relied on data from the API but, since October 2017, it is no longer supported[1]. For *Avaaz.org*, also launched in 2007 but founded by non-profit organizations, the only case study to date focused on whether this community fulfills certain basic democratic dimensions [Hor17]. This study reported serious problems of transparency and accountability because of the lack of information about their activities. In fact, *Avaaz.org* does not provide an open data API, which is a major barrier to research in this field.

To overcome the lack of available data from global petition platforms, we present in this appendix an open dataset of petitions from *Avaaz.org*[2]. In the following section we describe how we obtained the data in line with technical and legal requirements, and the structure of the dataset. We then explore the data to provide some findings of interest. To motivate future work, we conclude by offering example research questions that could be addressed with our dataset.

[1] `https://help.change.org/s/article/Change-org-API`

[2] Available at `https://dataverse.mpi-sws.org/dataverse/icwsm18`

A.2 Data Collection

To generate the dataset of petitions from *Avaaz.org*, we implemented a web crawler based on the incremental nature of their numerical ids. First, for a given petition id, a script sent a request to the AJAX endpoint of *Avaaz.org* and retrieved the corresponding URL. Then, with the petition URL, another script fetched and parsed the HTML to extract and store the corresponding metadata (an example petition page is shown in Figure A.1). The crawling process was done in August 2016 and the petition ids ranged from 1 to 382979 (which was the latest petition at that time). After excluding deleted pages, we obtained a dataset of 366,214 petitions.

It is important to highlight two issues taken into account when the crawler was designed. First, the machine-readable robots.txt file on *Avaaz.org* does not specify any restrictions[3]. Second, every page fetched by the crawler specified a *Creative Commons Attribution 3.0 Unported License* in the footnote. Therefore, our dataset is released under the same terms.

A.2.1 Structure of the dataset

The metadata of the online petitions were processed to produce a standardized and enriched dataset. Besides the id and URL, each petition contains the following fields:

- `title` (string): Title of the petition (limited to 100 characters). Following the official guidelines about how to write a petition title [Ava18b], many of them include the person, organization and/or location it addresses.

- `author` (string): Name of the user who authored the petition. To preserve anonymity, *Avaaz.org* includes only the given name and the first initial of the family name.

[3] `https://secure.avaaz.org/robots.txt`

- `description` (string): Description of the petition.

- `date` (timestamp): Date when the petition was published, from December 2011 to August 2016. Because dates were originally found in different languages including different writing systems (Latin, Arabic, Cyrillic, Kana, Hebrew, and Greek), we standardized them as *yyyy-MM-dd*.

- `country_name` (string): Name of the country of origin of the author. Users are able to customize this field in their profiles. Otherwise, the value is obtained by *Avaaz.org* directly from the user's IP address.

- `country_code` (string): As well as `date`, country names were originally found in different languages and writing systems. Therefore, we standardized countries using ISO 3166-1 alpha-3 codes[4].

- `sign` (integer): Number of signatures at the time the petition was crawled.

- `target` (integer): Number of signatures set as a goal by the platform/creator (this value may change as petitions receive signatures).

- `ratio` (float): Ratio between the two proceeding fields.

- `facebook_count` (integer): Number of shares on *Facebook*.

- `twitter_count` (integer): Number of shares on *Twitter*.

- `whatsapp_count` (integer): Number of shares on *WhatsApp*.

- `email_count` (integer): Number of shares via email.

[4]`https://unstats.un.org/unsd/methodology/m49/`

- `lang_code` (string): Language detected on the concatenation of title and description using a plugin for Apache Nutch project[5] (based on n-grams of over 50 languages). The resulting language is standardized with a ISO-639 2-letter code[6].

- `lang_prob` (float): Probability of success given by the language detection plugin.

- `people` (multivalued string): The names of people found within the concatenation of the title and description fields using the Stanford Named Entity Recognizer (NER) [FGM05]. Results are only provided when the detected language was English, Spanish or German (available languages).

- `organizations` (multivalued string): The names of organizations found using the NER, as done for `people`.

- `locations` (multivalued string): Locations found using the NER, as done for `people`.

- `miscellany` (multivalued string): Other entities found using the NER, as done for `people`.

A.3 Data Exploration

In this section we explore the petitions in our dataset to illustrate its content and relevance for research. We first provide a general overview regarding authors and signatures, and then inspect the link between signatures and shares on social media platforms. Finally, we examine some geographical and multilingual findings about the worldwide community of *Avaaz.org*.

[5] `https://wiki.apache.org/nutch/LanguageIdentifier`
[6] `http://www.loc.gov/standards/iso639-2/php/code_list.php`

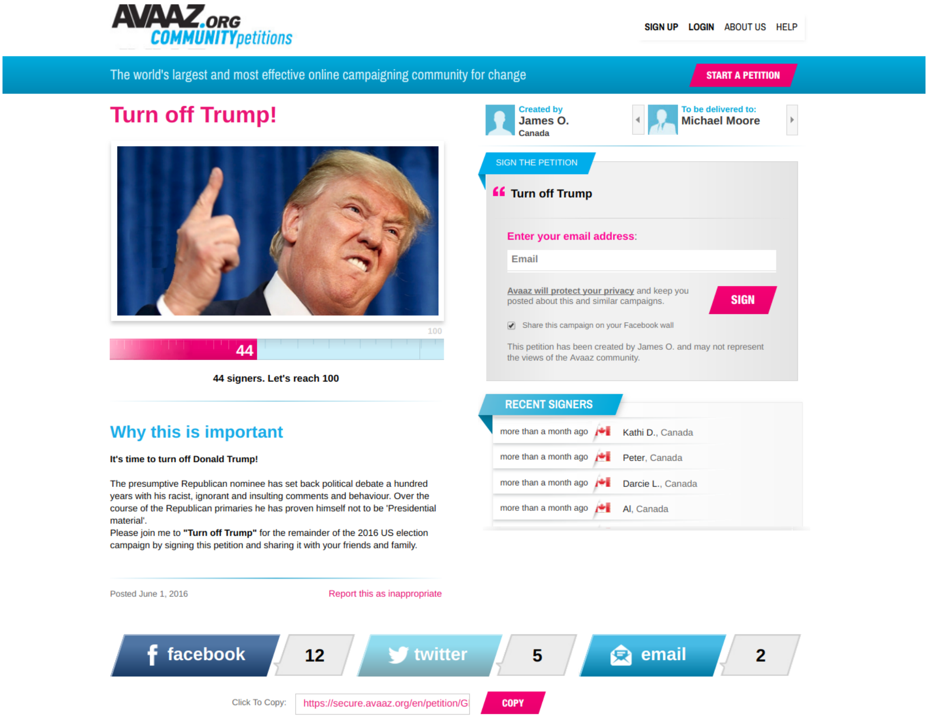


Figure A.1. Web page of a petition in *Avaaz.org* `https://secure.avaaz.org/en/petition/Global_media_consumers_Turn_off_Trump/`.

A.3.1 General overview

Figure A.2 presents general descriptive statistics of petitions in relation to authors and signatures. The plot in Figure A.2a shows the distribution of authors by the number of petitions. As expected, the distribution appears to follow a power law: most authors only published a few petitions. However, there is a small group of authors with over 1,000 petitions each. We examined this group and found that the most prolific user (2,895 petitions) is named *selenium s.*, located in Afghanistan and the United States, and authored peti-

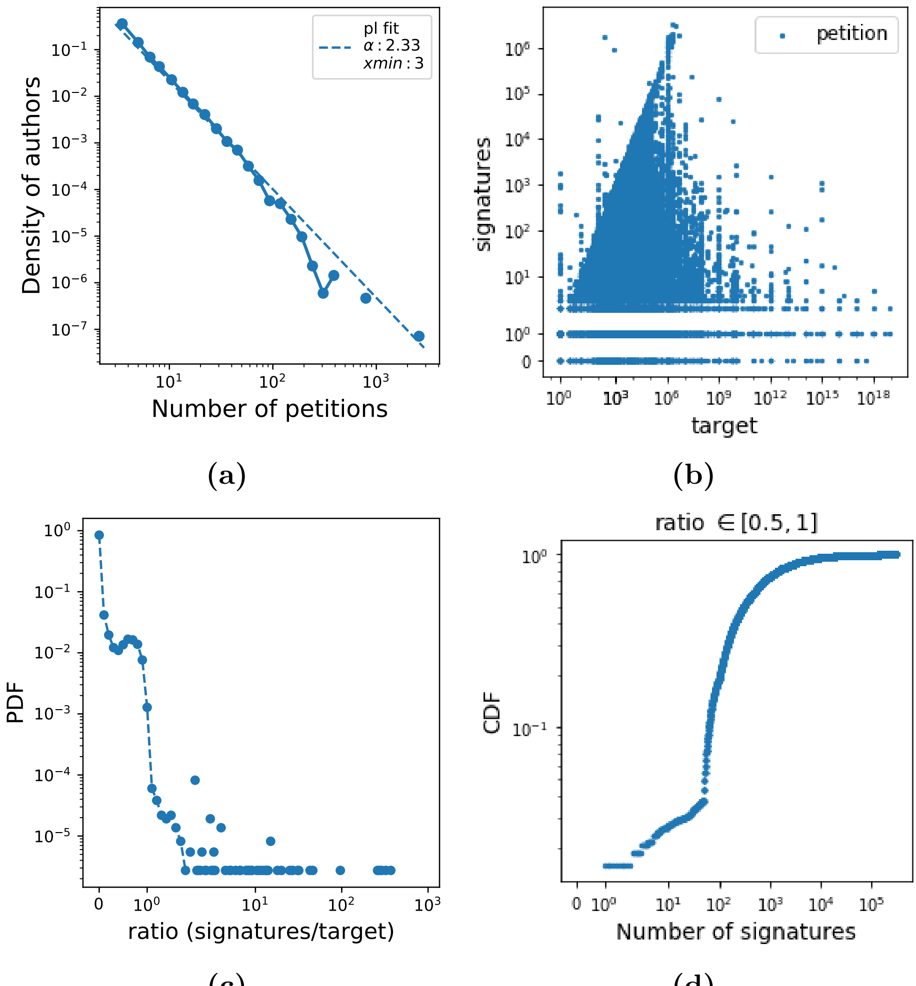


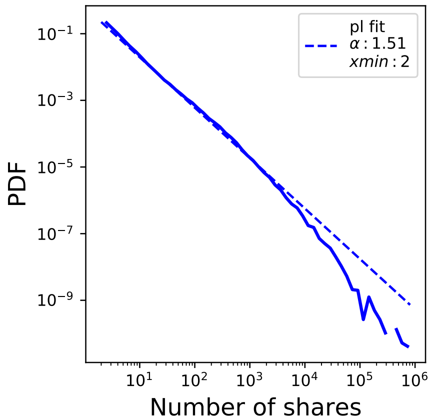
Figure A.2. General descriptive plots of the dataset of petitions: a) Distribution of authors by the number of petitions; b) Target vs signatures; c) Distribution of petitions by ratio; d) Cumulative distribution of petitions by signatures for petitions with a ratio $\in [0.5, 1]$.

tions like *“TEST - AUTOMATED - Who: TEST - AUTOMATED - What”* or *“selenium1440423202: selenium1440423202”*. This might indicate that these petitions were automatically generated with Selenium[7], a popular web browser automation tool[8]. Although much recent research has been devoted to characterize bots in social media platforms [CGWJ10, RCM+11, FVD+16], most studies have focused on Twitter. To the best of our knowledge, this is the first evidence of bots operating in an online petition platform.

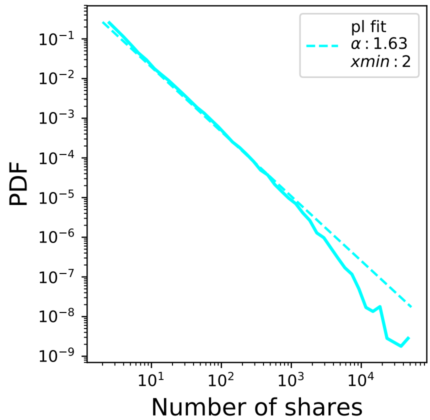
To assess the success of petitions in our dataset, we present a scatter plot of petitions' target number of signatures compared to their actual number of signatures (Figure A.2b). These dimensions are not correlated and the plot reveals that the number of signatures rarely exceeds the targets. This finding is explicit in Figure A.2c which shows the distribution of petitions by ratio, i.e., the number of signatures for each petition divided its target. We should note that, although the fraction of petitions decreases as the ratio increases, this is not the case for petitions with ratio $\in [0.5, 1]$. The observed peak could be the result of targets automatically updating when the number of signatures reaches a specific threshold. This is a relevant design feature in some online petition platforms that could encourage more signatures by magnifying the importance of a new signature to reach the target [MJRH12, FS04]. For this reason, we present in Figure A.2d the cumulative distribution of petitions by their number of signatures for petitions with ratio $\in [0.5, 1]$. The plot shows two trends: one from 1 to 99 signatures and the other above 100 signatures, which is the default initial target on the platform. It appears the target number of signatures is mostly likely to automatically update after a petition has at least 100 signatures.

[7] `http://www.seleniumhq.org`

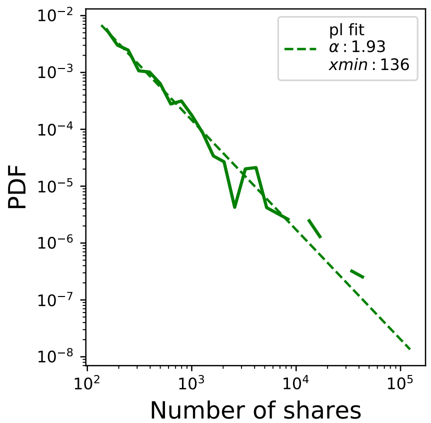
[8] An extended explanation of this user, provided by *Avaaz.org* after submitting the camera-ready copy of the corresponding article, is described in the last section.



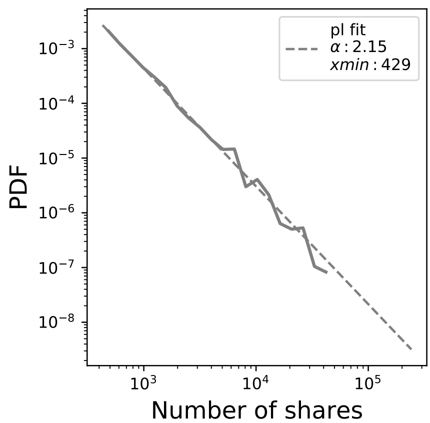
(a) *Facebook.*



(b) *Twitter.*



(c) *WhatsApp.*



(d) Email.

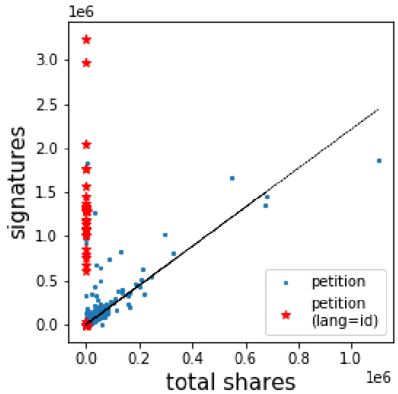
Figure A.3. Distribution of petitions by number of social media shares.

A.3.2 The role of social media

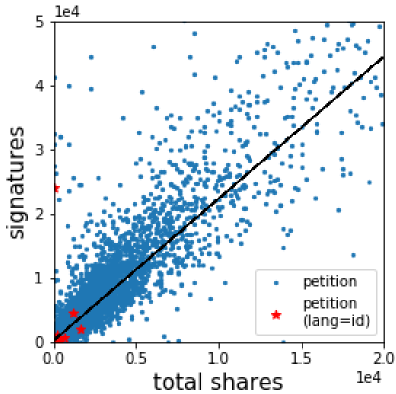
Recent research has found that the growth and success of online petitions is influenced by their popularity in social media [MJHY15, PACM16, PGK+17]. Given that our dataset includes how many times each petition was shared on *Facebook*, *Twitter*, *WhatsApp* and email, we present in Figure A.3 the distribution of petitions by the number of shares in each platform. The plots show heavy-tailed distributions: most petitions are not shared on social platforms while a few are highly shared. This is consistent with previous studies of diffusion on social media [GWG12, HJMY18].

We explicitly examine the link between social media and the success of online petitions with a scatter plot of signatures versus the sum of shares on four social platforms (see Figure A.4a). Although the variables are positively correlated ($r = 0.44, p < 0.001$), as better shown in Figure A.4b, we find of great interest the existence of a group of 25 petitions that received more than 500K signatures but less than 100 shares. We inspected these petitions and found that 24 of them were written in Indonesian by 18 authors not located in Indonesia but in the United Kingdom, the United States, Spain, France, Italy, Costa Rica and the Palestinian territories. Using the `author` field (given name and the initial of the family name), we searched Google using queries of the format: *site:linkedin.com avaaz* ***name***. For each query, we found an employee of *Avaaz.org* (e.g., Campaign Directors, Senior Campaigners) matching with `name` and the geographical location of corresponding petitions. We provide two possible explanations for these results[9]. On the one hand, they could be an indicator of astroturfing, i.e., these petitions could have been massively signed in an artificial manner. We should note that, in addition to the aforementioned problems of transparency and accountability [Hor17], there are specific complaints about the reliability of the number of signatures to petitions on *Avaaz.org* [Haw15]. On the

[9] After submitting the camera-ready copy of this article, the actual cause, described in the appendix, turned out to be neither of these two.



(a) Full dataset.



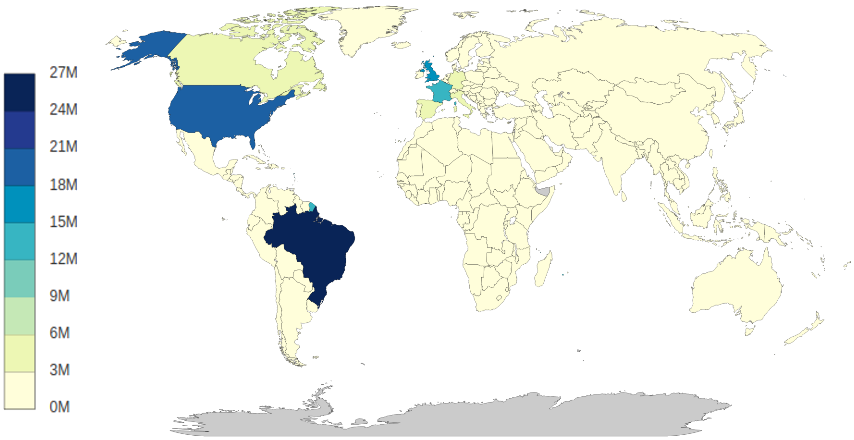
(b) Subset (shares $\leq 2 \cdot 10^4$ and signatures $\leq 5 \cdot 10^4$) covering 99.8% of the full dataset.

Figure A.4. Signatures versus shares in social media. The black line is a linear fit, and petitions written in Indonesian are shown with red star markers.

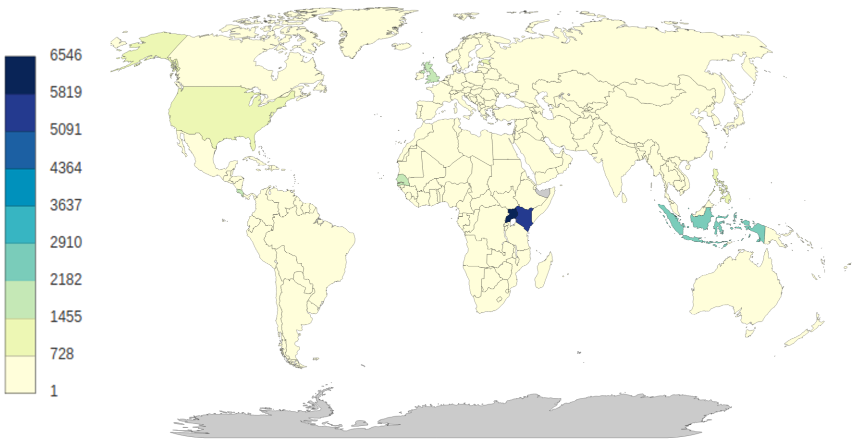
other hand, *Avaaz.org* and other websites have been blocked in Indonesia in recent years [Wik07], and this finding could be the result of very effective campaigns through alternative (even non-digital) diffusion channels.

A.3.3 Geographical and multilingual findings

Because the *Avaaz.org* community is present in over 200 countries, we show in Figure A.5 two choropleth world maps comparing activity across countries. The left map indicates the total number of signatures for all petitions in the dataset and reveals that activity is intense in Brazil, Spain, and countries that are members of the Group of Seven (G7) excluding Japan (Canada, United States, United Kingdom, France, Italy, and Germany). These results are similar to the map of the number of Avaaz 'members' from each country in the



(a) Sum of signatures (in millions).



(b) Average number of signatures per petition.

Figure A.5. Choropleth world maps of signing activity by country.

Avaaz Annual 2016 Poll [Ava16]. In contrast, the right map shows the average number of signatures per petition and depicts a very different distribution with the highest values in two African countries, Uganda and Kenya, followed by Indonesia.

To examine the most popular languages for *Avaaz.org* petitions from very active countries, we present a heatmap in Figure A.6. For better readability, data are normalized by the number of petitions in each country, and countries on the horizontal axis are grouped by linguistic similarity. We observe that, although users tend to use the most spoken language in their countries, English acts as a global language with remarkable use in countries like Afghanistan, Poland, and the Netherlands. The results also allow for the easy identification of strongly multilingual countries, e.g., Canada (French and English), Argelia (French and Arab), Morocco (French and Arab), Ukrania (Russian and Ukranian), Belgium (French, Dutch and English) and Turkey (Turkish, Arab and English).

Finally, we explore named people within the text of petitions. As indicated in the second section, this exploration is limited to petitions written in English, Spanish or German. Table A.1 shows the top 10 people who are mentioned in petitions from the largest number of countries. The table also includes the top 3 countries associated with each person. Because petitions in *Avaaz.org* are aimed at solving global societal challenges, it was expected to find people associated to global leadership (e.g., Barack Obama or Angela Merkel). In addition to politicians, we find of interest the presence of Justin Bieber who appears in many petitions from Latin American countries like Argentina, Uruguay and Mexico. We inspect these petitions and found that many of them were about an indictment from an Argentinian court against him in 2016. This leads us to examine which countries share similar political and social references with a heatmap of the number of people in common between the most active English, Spanish and German speaking countries (see Figure A.7). Besides the overlap among countries very active on *Avaaz.org* (United States, United Kingdom, Germany, Canada and

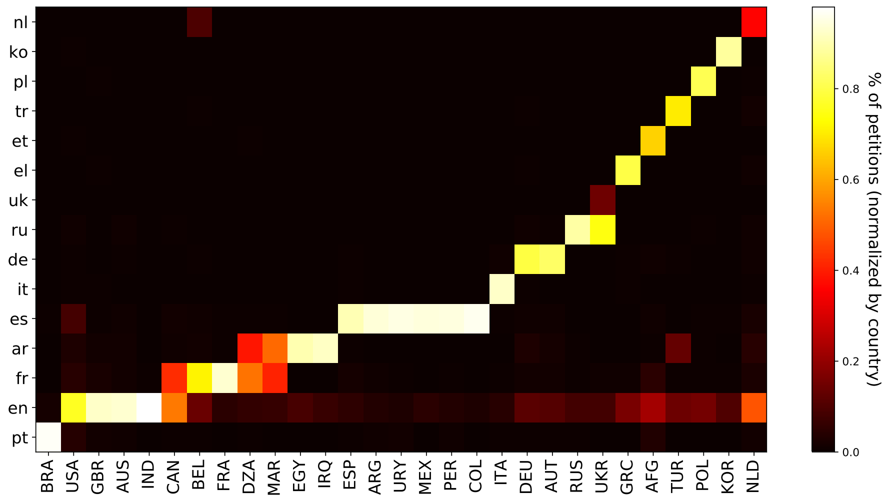


Figure A.6. Heatmap of the percentage of petitions from different countries written in the most popular languages.

Spain), we should note the specific overlap among Spanish speaking countries (Spain, Mexico, Colombia, Argentina, Venezuela, etc.) and between the United States and Mexico.

A.4 Recommendations for future work

In the above section we have presented a preliminary exploration of our dataset of online petitions from *Avaaz.org*. On the basis of these results, we conclude this appendix by proposing different example research questions to reflect potential uses of the dataset.

Bot detection

In the current context of the rise of social bots [FVD^{+}16], we have provided the first evidence of bots on a petition platform by examin-

Table A.1. Top 10 people by the number of countries with English, Spanish or German petitions naming them. The last column indicates the three countries with more petitions naming the corresponding person (values in brackets).

Person	No. of countries	Top 3 countries					
Barack Obama	94	USA	(211)	DEU	(41)	GBR	(40)
Boris Johnson	71	GBR	(60)	USA	(15)	URY	(9)
Angela Merkel	46	DEU	(267)	AUT	(14)	GRC	(10)
Ban Ki-Moon	46	USA	(12)	GBR	(12)	CHE	(11)
David Cameron	41	GBR	(471)	FRA	(7)	USA	(5)
Vladimir Putin	41	USA	(18)	GBR	(14)	RUS	(11)
Edward Snowden	31	DEU	(69)	USA	(18)	GBR	(14)
Justin Bieber	29	ARG	(30)	URY	(19)	MEX	(11)
Donald Trump	26	USA	(17)	GBR	(8)	MEX	(6)
Xi Jinping	26	USA	(11)	GBR	(5)	CAN	(3)

ing the most active users. Nevertheless, *could bot-generated petitions also be detected with other (e.g., text-based) features?*. Indeed, our dataset is not only affected by bots publishing a large number of petitions but also by suspiciously high levels of support to petitions with almost no traction on social media. If astroturfing is the reason behind this pattern, and as done on other platforms like Twitter [RCM+11], *could an automatic classifier identify petitions with artificial support?* Given the nature and potential of online petitions to influence policy makers, the answer to these questions would have clear political and social implications.

Content analysis

Our exploration of the content of online petitions was limited to detected languages and named entities. This has allowed us to obtain a geographic overview of the *Avaaz.org* community, revealing patterns of multilingualism and common political references among countries.

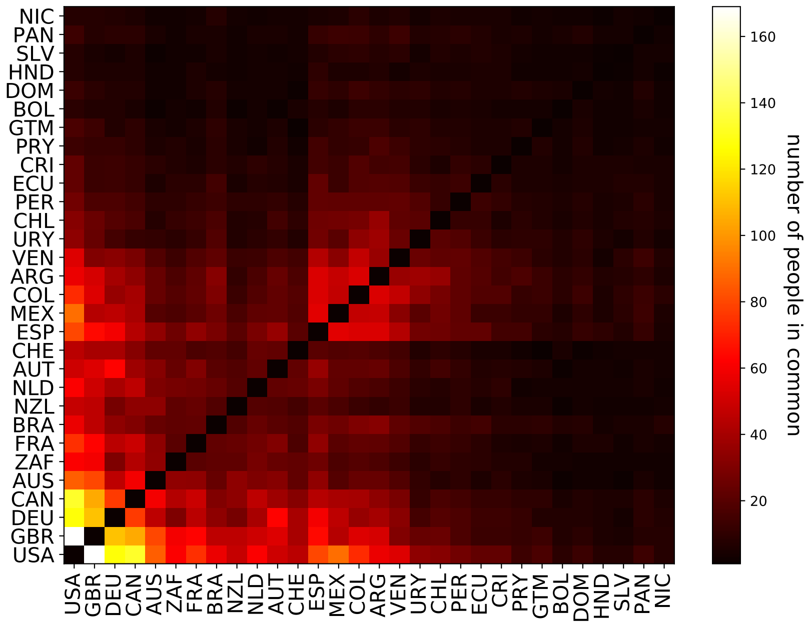


Figure A.7. Heatmap of the number people who are named in petitions from multiple countries. The chosen countries have the largest number of petitions written in English, German or Spanish (the available languages of the Named Entity Recognizer). For better readability, values on the diagonal are set to 0.

However, the textual content of the petitions still has great informative value to be examined. First, future work might focus on sentiment analysis. Given that recent research on *Change.org* found that petitions were more likely to be successful when having positive emotions [EDK16], *is this a specific finding of Change.org or also valid for the community of Avaaz.org?*

Second, topic modeling could also be of great interest. For example, given the topics of online petitions from a multilingual country, *do different linguistic communities care about the same issues?* Our dataset could help to identify socio-political problems of these communities. Furthermore, detected topics along with `date` field values would allow modeling the rise and decay of topics: *how do topics emerge in online petition platforms?* According to *Avaaz.org*, overall priorities for campaigns are set through member polls [Ava18a]. However, the motivations of members in selecting priority issues remain unclear. For the same period of time, a comparison of topics from our dataset to topics from external sources (e.g., news repositories, social media) would shed light on political agenda setting in the era of global activism.

Gender studies

Because `author` in our dataset is essentially the given name of the user who published a petition, future work might also focus on extracting the gender of authors to then identify whether there are significant differences in the topics about which men and women create petitions. A recent study on *Change.org* found that women were more likely to publish petitions about animals and women's rights while men focused on economic justice and (general) human rights [MGSP17]. Besides comparing whether these results are also valid in *Avaaz.org*, the geographical and linguistic information of our dataset would allow for the investigation of a more detailed research question: *is the distribution of relevant topics for men and for women stable over countries or affected by cultural factors?* Given the increasing awareness about gender biases online (e.g., *Wikipedia* [WGGGM16] and *Facebook* [GMKC$^+$18]), our dataset might be a helpful resource to assess whether the democratic purpose of online petitioning is affected by any gender imbalances.

A.5 Comments provided by *Avaaz.org*

After this work was published, we were contacted by Ben Boyd (Chief Technology Officer at *Avaaz.org*) who provided an explanation of the anomalies detected when exploring the dataset. First, some petitions authored by *selenium s.* are likely the result of automated testing by *Avaaz.org* which were not properly removed. However, bot activity is an ongoing challenge that *Avaaz.org* is actively working to tackle. Second, the reason of the petitions with more than 500k signatures with extremely low social share counts is that *Avaaz.org*'s campaigners linked some campaigns internally to the core campaigning platform to boost them. The linking procedure wrongly showed the number of signatures from the source petition of the campaign while not showing the corresponding share counts. This bug is now fixed (since June 2018). We would like to thank Ben Boyd for his insightful feedback and for his interest in this research project.

B Knowledge Discovery Systems for Decide Madrid

B.1 Introduction

Madrid City Council launched in 2015 *Decide Madrid*, a civic technology to foster open democracy and citizen participation through, among other features, online petitions[1]. The functioning of this platform is as follows: First, a user publishes an online petition that proposes a public policy to be implemented in Madrid. Then, other users can comment the petition to favor deliberative practices of decision making, and users verified as citizens of Madrid can also support the petition. After one year, if the petition obtains more than 27,064 supporting votes (1% of the population of Madrid over 16 years old), a voting process, advertised in the main page, is held. If most users vote in favor and the petition fulfills certain technical and ethical requirements, the public policy is implemented by the City Council. In contrast, online petitions which do not reach the support threshold in one year are withdrawn from the platform.

[1] `https://decide.madrid.es/proposals`

To investigate the potential of *Decide Madrid*, the City Council launched in 2016 the Collective Intelligence for Participatory Democracy Lab (Participa LAB)[2]. This space is part of Medialab Prado, a citizen laboratory for exploring the forms of experimentation and collaborative learning that have emerged from digital networks. The specific mission of Participa LAB was to investigate forms of direct and deliberative democracy involving new digital tools, e.g., through the analysis of data from *Decide Madrid*.

The increasing abundance of data on urban activity has motivated the development of interactive data visualization systems (i.e., dashboards) in many cities like Amsterdam[3], Dublin[4], or Boston[5]. While dashboards of this type have typically relied on data from city sensors (e.g. traffic, noise, air pollution), open data portals, or social media platforms (e.g., *Twitter*, *Facebook*, *Instagram*), little effort has been devoted to deploy interactive discovery systems with data from civic participation platforms. One of the few cases is *Civic CrowdAnalytics*, a tool which applies natural language processing with data from a crowdsourced urban planning process in Palo Alto [ACC+16]. The tool served to reveal that the impact of citizens' voices was related to the volume and the tone of their messages, i.e., demands with more messages and more emotive tone led to greater changes in public policies. Another example of interest is *Pol.is*[6], a commenting and survey system which applies principal component analysis to get major variances in opinions for different voters. These variances are then clustered using the k-means method, together with the silhouette score, to identify opinion groups [Yen17]. Although that platform was originally developed for news forums, it rapidly evolved into a civic technology when adopted by the *vTaiwan* com-

[2]`https://www.medialab-prado.es/en/laboratories/participalab`

[3]`http://citydashboard.waag.org`

[4]`http://www.dublindashboard.ie`

[5]`https://boston.opendatasoft.com`

[6]`https://pol.is/home`

munity[7] to detect controversial issues about *Uber/Airbnb* Taiwanese regulations [Meg16].

The work presented below corresponds to two systems developed in collaboration with Participa LAB for the discovery of knowledge through mining and visualizing data from *Decide Madrid*.

B.2 Visualization tool for online discussion threads

The design of online platforms for collective awareness is gaining increasing interest. For example, the European Commission launched in 2012 a research initiative called *Collective Awareness Platforms for Sustainability and Social Innovation* (CAPS) to explore the confluence of knowledge and social networks[8]. For awareness to be leveraged in practice, the way in which people acquire information is crucial [PSB+14]. Deliberative democracy requires citizens to be aware of the argumentative structure of every proposal. Therefore, argument mapping in large discussions becomes essential for the decision-making process. The tool presented in this section is a result of the CAPS principles in *Decide Madrid*. In particular, our tool allows a deeper understanding of the argumentative structure of proposals, which improves collective awareness of citizens and provides useful hints for the refinement of *Decide Madrid*.[9]

We visualize the hierarchical discussion as an interactive radial tree in which the root node corresponds to the proposal and the rest of the nodes correspond to the comments from the users, e.g. [PCK09]. To highlight the arborescence of the discussion and to distinguish the arguments of every branch of the thread, the tool applies a flexible force-directed graph layout that accelerates charge interaction through the Barnes-Hut approximation [PG05]. In addition, to iden-

[7] `https://vtaiwan.tw`

[8] `https://ec.europa.eu/digital-agenda/en/collectiveawareness`

[9] Code available at: `https://github.com/elaragon/decideviz`

tify the messages that receive more attention, the size of the nodes is proportional to the number of votes. Furthermore, the colour of the nodes is determined according to the ratio of positive/negative votes:

- Black: Root (proposal)
- Grey: Comment with no votes
- Green (scale): Comment with majority of positive votes
- Red (scale): Comment with majority of negative votes
- Orange: Comment with no strong preference of positive or negative votes

The visualization also includes an informative panel with the description of the tool and the metadata of a node (author, message, date and number of positive/negative votes) when the user rolls the mouse over it.

To illustrate the potential of this tool, Figure B.1 illustrates the interface using an example of a proposal about banning bullfighting[10]. We observe a large number of orange nodes, reflecting the strong controversy generated by this topic in Madrid. Among the most voted comments (the largest nodes), both positively and negatively voted comments appear. The best rated ones consist of contributions from animal rights organizations and feedback from citizens that suggest merging all the anti-fighting proposals posted in *Decide Madrid*. Indeed, the existence of multiple proposals for the same goal is one of the patterns that has been identified with this tool (recent versions of *Decide Madrid* have addressed this issue). In contrast, the comments that receive a majority of negative votes in this proposal are usually messages that define bullfighting as an artistic discipline in Spain.

[10] https://decide.madrid.es/proposals/105

Figure B.1. Visualization of the discussion of the petition "Eliminar los festejos taurinos y las subvenciones" (Remove bullfighting and its subsidies). Left side shows the discussion as an interactive radial tree. Right side displays the content of a selected comment (see the main text for a detailed description of the interface).

B.3 Interactive discovery system of online petitions

In the first year of *Decide Madrid*, two petitions reached the support threshold: one to unify public transport tickets[11] and the other to implement specific environmental sustainability policies[12].

Despite the success of both petitions, two aspects require special attention. First, both petitions were published on the day *Decide Madrid* was launched. Second, no petition published in subsequent months has reached even half the threshold. According to previous analyses [Bet15], this might be explained by how petitions are presented in the platform. Originally, they were listed using pagination and a sorting criteria based on the number of supporting votes. As one could expect, this strategy embodied the so-called 'Matthew effect' [Mer68]: the first petitions rapidly became popular and covered the first page of the ranking, which made new petitions nearly invisible. To overcome this effect, the City Council replaced the ranking algorithm with the *Hot score* method of Reddit[13]. However, this method was devised to rank news items whose interest usually declines rapidly, while online petitions require longer visibility to engage thousands of citizens and hence reach the threshold[14]. Therefore, the results of this second strategy are still unsatisfactory and illustrate the socio-technical problem of ranking and filtering of information in social systems [SPS14]. Another problem of interest is the generalist approach of the interface, i.e., petitions are shown regardless of the user preferences. Although some other ranking methods with personalized recommendation have been proposed [CBCCG17b], every approach has always relied on the assumption that petitions must be

[11] `https://decide.madrid.es/proposals/9`

[12] `https://decide.madrid.es/proposals/199`

[13] `https://github.com/consul/consul/blob/master/lib/score_calculator.rb`

[14] This observation and the work presented in this section were done before the ranking algorithm of petitions was changed (see details in Chapter 5).

presented as a list and, therefore, requires a ranking algorithm. The system presented here is motivated by the use of alternative graphical interfaces to experiment with other strategies of human-computer interaction in *Decide Madrid*. In particular, our system applies different techniques of data analysis and visualization to facilitate the discovery of topics and petitions[15].

To illustrate how the systems works, we will use an example based on petitions related to trash, a controversial issue in Madrid which has motivated many petitions on *Decide Madrid* proposing different solutions to the existing problems.

Discovery of topics

Figure B.2 shows the current web interface of *Decide Madrid* which displays a list of petitions about any topic using the *Hot score* method as sorting criteria by default. Although it allows users to retrieve petitions matching the term "basuras" (trash), they would have then to scroll through a dozen of pages to find proposals of interest among the 273 existing petitions. In contrast, the alternative web interface, shown in Figure B.3, starts with a query form to force users to explicitly set their interest. Once a user has filled the form, the module of topics retrieves the matching petitions from the API[16] to then groups them into topics with the text clustering method Carrot[2] [OW05]. This is a state of art method for web search results clustering which is based on the the *Suffix Tree Clustering* algorithm (STC) [SW03]. The idea behind this approach is that topics usually correspond to identical sequences of terms (phrases) and, therefore, groups of documents can be identified by such sequences. The algorithm follows two steps:

1. Discovery of groups of petitions with identical phrases.

2. Merge of these groups into larger clusters.

[15] Code available at: `https://github.com/elaragon/decide-topics`
[16] `https://decide.madrid.es/graphiql`

Thus, the phrase which characterizes each final cluster is used by STC as a label to easily identify topics of petitions. Petitions can occur in multiple clusters and those without any assigned cluster can be found at "Other topics".

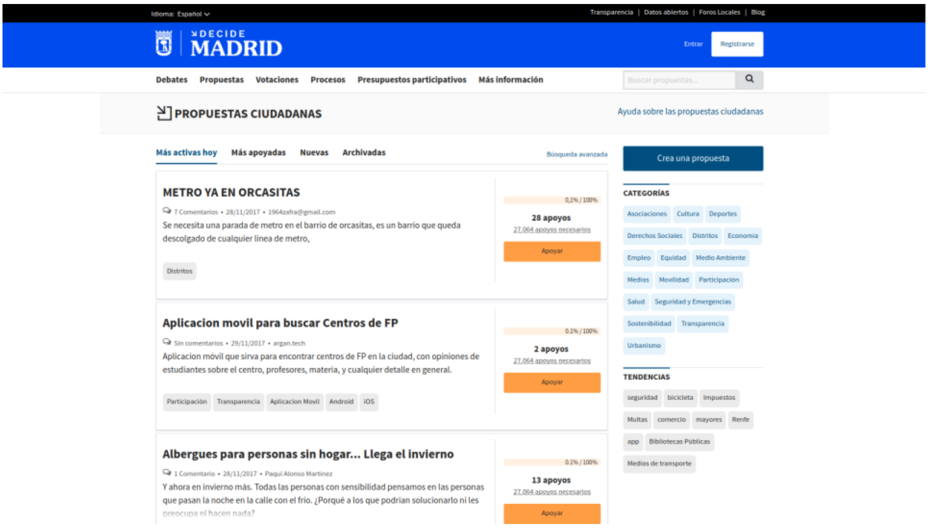


Figure B.2. Current front page of petitions in *Decide Madrid* using a list of items sorted by the Hot Score, the default criteria.

Resulting topics are presented in a mosaic plot with distinctive color for each topic and size according to the sum of supporting votes to the petitions of the corresponding cluster. Metadata of each topic (label and number of petitions and sum of supporting votes) are displayed in a tooltip on mouse over. Figure B.3 shows the topic visualization of the term "basuras" for which the discovery system automatically identifies 44 related topics, e.g., "garbage bins", "garbage tax", "cleaning services". The full list is presented in Table B.1. When the user clicks on a topic (either at the diagram or at the legend) the second module is loaded using the selected topic as input.

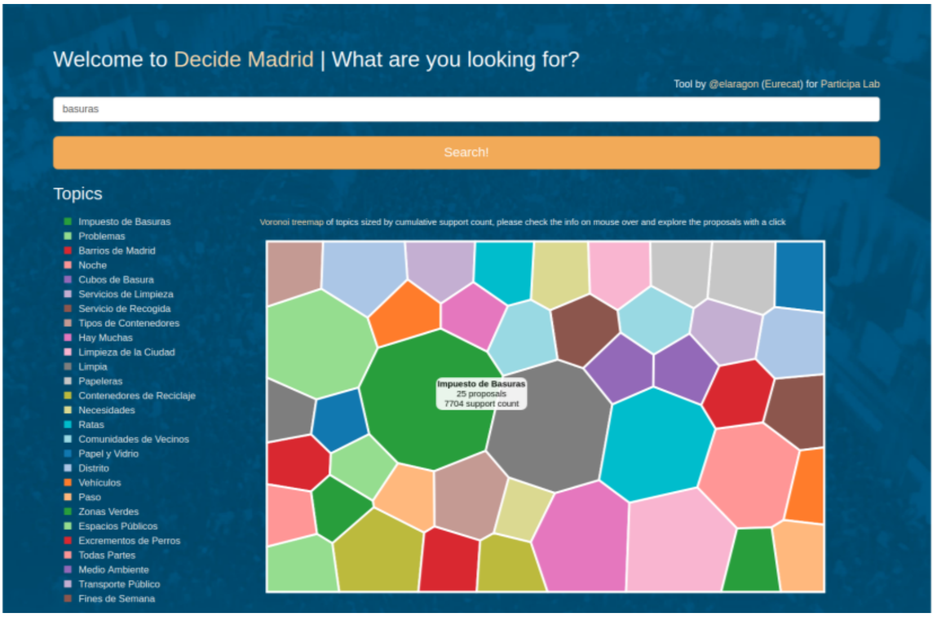


Figure B.3. Interface of the module of topics which groups online petitions into topics and shows them within a mosaic plot (example for the query term "basuras", trash in Spanish).

Discovery of petitions

The web interface of the module of petitions is presented in Figure B.4. To distinguish that items are now petitions of a selected topic, these are shown as circles. The radius of the circle is based on the number of supporting votes to easily identify which are the most popular petitions. However, to avoid a ranking of petitions using lists as done in *Decide Madrid* so far, circles are displayed randomly without overlap by applying a force-directed graph layout [Jak01]. Metadata of each petition (title and number supporting votes) are also shown in a tooltip on mouse over.

The illustrative example in Figure B.4 corresponds to the petitions from the "garbage tax" topic. The cursor is over a popular petition which proposes the implementation of the German recycling system (*Mehrwegpfand*)[17]. When the user clicks on the circle, the petition web page is loaded in a new window (see Figure B.5) to allow the user to review its proposals and to support it.

Figure B.4. The interface shows the corresponding petitions as circles sized by the number of supporting votes (example for the topic "Impuestos de Basuras", garbage tax in Spanish).

[17]`https://decide.madrid.es/proposals/1824`

Figure B.5. Page of a popular petition proposing the implementation of the German recycling system.

B.4 Discussion

In this appendix we have presented how to facilitate knowledge discovery in *Decide Madrid* through visualization tool for argument mapping and an interactive system to browse and topics and petitions.

For the visualization tool of discussion threads, although many tools for collective intelligence through citizen discussion have been released (eg. Deliberatorium[18], , Loomio[19], DemocracyOS[20]), most of them do not present the argumentative structure of the discussion. Tools which include network visualizations (eg. Incoma[21], Edge-

[18] http://deliberatorium.mit.edu/
[19] https://www.loomio.org/
[20] http://democracyos.org/
[21] http://incoma.org/

Sense[22]) are devised to explore online forums with little impact in policy making in comparison to *Decide Madrid*. In this context of binding proposals, we believe that our tool provides a clear picture of the discussion of citizens in order to promote effective deliberative democracy. Furthermore, as shown in Figure B.6, this tool was selected for Big Bang Data - London, an art exhibition held at the Somerset House between December 2015 and March 2016 with the aim of allowing visitors to explore how data is transforming our world[23].

Figure B.6. The visualization tool for online discussion threads from *Decide Madrid* at the Big Bang Data exhibition - London (Somerset House, December 2015 and March 2016). Source: Kaisa Eskola.

[22]`http://wikitalia.github.io/edgesense/`
[23]`http://bigbangdata.somersethouse.org.uk/`

For the interactive discovery system, users interested in petitions about trash would have had to explore the rankings of petitions page by page in the official web interface. Our system is a way to solve this task in a more dynamic manner. Because these paginated rankings of petitions have been identified as a barrier for petition signing, simple but effective data visualization can serve to prevent civic technologies from undesired effects of this kind. In recent years, many similar platforms have been developed to facilitate public participation and, ultimately, collective action [MJHY15]. Nevertheless, research has provided evidence that social systems are very sensitive to how information is ranked and filtered [SPS14]. Given the increasing popularity of civic technologies, our system is an illustrative example of how these platforms might benefit from presenting information in a concise and well-interpretable way.

Table B.1. Full list of discovered topics for the petitions retrieved from *Decide Madrid* with the query term "basuras" (trash in Spanish). For each topic, the table includes an English traslation, the number of petitions, and the sum of supporting votes to the corresponding petitions. "Other topics" contains the petitions which were not assigned to any cluster by the *Suffix Tree Clustering* algorithm.

Topic	English translation	Number of petitions	Sum of supporting votes to petitions
Noche	Night	32	4,519
Calles y Aceras	Streets and sidewalks	31	4,256
Caso	Case	30	4,701
Cubos de Basura	Garbage Bins	30	2,676
Impuesto de Basuras	Garbage Tax	30	8,219
Servicios de Limpieza	Cleaning Services	28	3,869
Horarios de Recogida	Pickup Schedules	26	2,400
También Hay	There are also	25	3,923
Limpieza de la Ciudad	Cleaning the City	23	5,155
Tipos de Contenedores	Types of Containers	23	2,837
Camiones de Basura	Garbage Trucks	22	1,433
Suciedad	Dirt	22	5,236
Comunidades de Vecinos	Neighborhood Communities	21	2,630
Papeleras	Litter bins	21	4,665
Centro de Madrid	Centre of Madrid	20	2,798
Contenedores de Reciclaje	Recycling Containers	19	3,234
Mucha Gente	Many People	19	4,383
Mucho Menos	Much Less	15	2,038
Zonas Verdes	Green Zones	14	2,005
Excrementos de Perros	Dog Droppings	11	1,637
Malos Olores	Bad Smells	10	1,179
Medio Ambiente	Environment	10	1,599
Barrio Limpio	Clean Neighborhood	9	760
Transporte Público	Public Transportation	9	1220
Vía Pública	Public Roads	9	918
Plazas de Aparcamiento	Parking Spaces	8	368
Calidad de Vida	Quality of Life	7	1,492
Descampados	Open fields	7	605
Solares Abandonados	Abandoned Plots	7	1,735
Cantidad de Dinero	Amount of Money	6	3,548
Parques Infantiles	Children's playgrounds	6	362
Residuos Orgánicos	Organic Waste	6	307
Vehículos sin Motor	Non-Motor Vehicles	6	325
Nuestros Hijos	Our Children	5	209
Cualquier Sitio	Any Site	4	552
Problemas de Movilidad	Mobility Problems	4	658
Teniendo en Cuenta	Taking into account	4	186
Debería Hacerse	It Should Be Done	3	573
Distrito más Poblado	Most Populous District	3	340
Control de Plagas	Pest Control	2	291
Necesidad de Buscar	Need to Search	2	81
Pasar por Caja el Producto	Checkout the Product	2	2,256
Redes Sociales	Social Networks	2	250
Vivienda y te Conviertas en un Proscrito	Housing and Become an Outlaw	2	3,183
Other Topics	Other Topics	60	9,388

Bibliography

[ABA+96] Marshall Van Alstyne, Erik Brynjolfsson, Copyright Alstyne, Brian Butler, Rob Fichman, Rob Kling, and David Meyer. Electronic communities: Global village or cyberbalkans. In *Proceedings of the International Conference on Information Systems*, pages 80–98, 1996.

[ABGK18] Pablo Aragón, Yago Bermejo, Vicenç Gómez, and Andreas Kaltenbrunner. Interactive Discovery System for Direct Democracy. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 601–604, 2018.

[ABS+14] Ashley A. Anderson, Dominique Brossard, Dietram A. Scheufele, Michael A. Xenos, and Peter Ladwig. The "Nasty Effect:" Online Incivility and Risk Perceptions of Emerging Technologies. *Journal of Computer-Mediated Communication*, 19(3):373–387, 2014.

[ABS17] Molly Asher, Cristina Leston Bandeira, and Viktoria Spaiser. Assessing the effectiveness of e-petitioning through Twitter conversations. In *Political Studies Association Annual Meeting 2017*, 2017.

[ACC+16] Tanja Aitamurto, Kaiping Chen, Ahmed Cherif, Jorge Saldivar Galli, and Luis Santana. Civic CrowdAnalytics: Making Sense of Crowdsourced Civic Input with Big Data Tools. In *Proceedings of the 20th International Academic Mindtrek Conference*, AcademicMindtrek '16, pages 86–94, New York, NY, USA, 2016. ACM.

[ADNMJ14] Tim Althoff, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky. How to Ask for a Favor: A Case Study on the Success of Altruistic Requests. In *Proceedings of the International AAAI Conference on Web and Social Media*, 2014.

[AE11] Taylor B Arnold and John W Emerson. Nonparametric Goodness-of-Fit Tests for Discrete Null Distributions. *R Journal*, 3(2), 2011.

[AF04] Bruce A Ackerman and James S Fishkin. *Deliberation day.* Yale University Press, 2004.

[AGCSM08] Eva Anduiza, Aina Gallego, Marta Cantijoch, and Josep San Martin. Online resources, political participation and equality. American Political Science Association, 2008.

[AGGK17] Pablo Aragón, Vicenç Gómez, David García, and Andreas Kaltenbrunner. Generative models of online discussion threads: state of the art and research challenges. *Journal of Internet Services and Applications*, 8(1):15, 2017.

[AGK16] Pablo Aragón, Vicenç Gómez, and Andreas Kaltenbrunner. Visualization Tool for Collective Awareness in a Platform of Citizen Proposals. In *Proceedings of the International AAAI Conference on Web and Social Media*, 2016.

[AGK17a] Pablo Aragón, Vicenç Gómez, and Andreas Kaltenbrunner. Detecting Platform Effects in Online Discussions. *Policy & Internet*, 9(4):420–443, 2017.

[AGK17b] Pablo Aragón, Vicenç Gómez, and Andreas Kaltenbrunner. To Thread or Not to Thread: The Impact of Conversation Threading on Online Discussion. In *Proceedings of the International AAAI Conference on Web and Social Media*, 2017.

[AGMS15] Raquel Alvarez, David Garcia, Yamir Moreno, and Frank Schweitzer. Sentiment cascades in the 15M movement. *EPJ Data Science*, 4(1):6, May 2015.

[AKCL+17] Pablo Aragón, Andreas Kaltenbrunner, Antonio Calleja-López, Andrés Pereira, Arnau Monterde, Xabier E. Barandiaran, and Vicenç Gómez. Deliberative Platform Design: The Case Study of the Online Discussions in Decidim Barcelona. In *Social Informatics*, pages 277–287, Cham, 2017. Springer International Publishing.

[Ale14] Alberto Alemanno. Unpacking the principle of openness in EU law: Transparency, participation and democracy. *European Law Review*, 2014.

[AM17] Paolo Antonetti and Danae Manika. The offline spillover of signing online petitions against companies: A dual pathway model. *Information Technology & People*, 30(4):969–990, 2017.

[AMJ05] Nicholas J. Adams, Ann Macintosh, and Jim Johnston. E-Petitioning: Enabling Ground-Up Participation. In Matohisa Funabashi and Adam Grzech, editors, *Challenges of Expanding Internet: E-Commerce, E-Business, and E-Government*, pages 265–279. Springer US, 2005.

[AMS09] Sinan Aral, Lev Muchnik, and Arun Sundararajan. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*, 106(51):21544–21549, 2009.

[Arn69] Sherry R. Arnstein. A Ladder Of Citizen Participation. *Journal of the American Institute of Planners*, 35(4):216–224, 1969.

[ASTR+18] Pablo Aragón, Diego Sáez-Trumper, Miriam Redi, Scott Hale, Vicenç Gómez, and Andreas Kaltenbrunner. Online Petitioning Through Data Exploration and What We Found There: A Dataset of Petitions from Avaaz.org. In *Proceedings of the International AAAI Conference on Web and Social Media*, 2018.

[Ava16] Avaaz. Annual 2016 Poll. `https://secure.avaaz.org/en/poll_results_2016/`, 2016. [Online; accessed 10-June-2019].

[Ava18a] Avaaz. About us. `https://avaaz.org/page/en/about/`, 2018. [Online; accessed 10-June-2019].

[Ava18b] Avaaz. How to write a petition title. `https://secure.avaaz.org/en/petition/how_to_write_a_petition_title/`, 2018. [Online; accessed 10-June-2019].

[AVLK16] Pablo Aragón, Yana Volkovich, David Laniado, and Andreas Kaltenbrunner. When a Movement Becomes a Party: Computational Assessment of New Forms of Political Organization in Social Media. In *Proceedings of the International AAAI Conference on Web and Social Media*, 2016.

[BA99] Albert-László Barabási and Réka Albert. Emergence of Scaling in Random Networks. *Science*, 286(5439):509–512, 1999.

[Bar84] Benjamin R Barber. *Strong democracy: Participatory politics for a new age.* University of California Press, 1984.

[BC+92] Michael W Browne, Robert Cudeck, et al. Alternative ways of assessing model fit. *Sociological Methods & Research*, 21(2):230–258, 1992.

[BC94] Donald J Berndt and James Clifford. Using dynamic time warping to find patterns in time series. In *KDD workshop*, volume 10, pages 359–370, 1994.

[BC11] Rosa Borge and Ana S. Cardenal. Surfing the Net: A Pathway to Participation for the Politically Uninterested? *Policy & Internet*, 3(1):1–29, 2011.

[BEB14] Eytan Bakshy, Dean Eckles, and Michael S. Bernstein. Designing and Deploying Online Field Experiments. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14, pages 283–292, New York, NY, USA, 2014. ACM.

[Ben06] Yochai Benkler. *The wealth of networks: How social production transforms markets and freedom.* Yale University Press, 2006.

[Ber17] Janne Berg. The dark side of e-petitions? Exploring anonymous signatures. *First Monday*, 22(2), 2017.

[Bet15] BetademIC. Informe final grupo de trabajo BetaDemIC. `https://rxnetwalker.gitbooks.io/betadoc/content/`, 2015.

[BEYR12] Eytan Bakshy, Dean Eckles, Rong Yan, and Itamar Rosenn. Social Influence in Social Advertising: Evidence from Field Experiments. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, pages 146–161, New York, NY, USA, 2012. ACM.

[BGLL08] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

[Bim03] Bruce Bimber. *Information and American Democracy: Technology in the Evolution of Political Power.* Communication, Society and Politics. Cambridge University Press, 2003.

[BKLDNM13] Lars Backstrom, Jon Kleinberg, Lillian Lee, and Cristian Danescu-Niculescu-Mizil. Characterizing and Curating Conversation Threads: Expansion, Focus, Volume, Re-entry. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*, WSDM '13, pages 13–22, New York, NY, USA, 2013. ACM.

[BMHH+11] Michael Bernstein, Andrés Monroy-Hernández, Drew Harry, Paul André, Katrina Panovich, and Greg Vargas. 4chan and /b/: An Analysis of Anonymity and Ephemerality in a Large Online Community, 2011.

[BNJ03] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent Dirichlet Allocation. *J. Mach. Learn. Res.*, 3:993–1022, March 2003.

[BPW+15] Seunghee Baek, Seong Ho Park, Eugene Won, Yu Rang Park, and Hwa Jung Kim. Propensity score matching: a conceptual review for radiology researchers. *Korean Journal of Radiology*, 16(2):286–296, 2015.

[BRMA12] Eytan Bakshy, Itamar Rosenn, Cameron Marlow, and Lada Adamic. The Role of Social Networks in Information Diffusion. In *Proceedings of the 21st International Conference on World Wide Web*, WWW '12, pages 519–528, New York, NY, USA, 2012. ACM.

[bro] inferring causal impact using bayesian structural time-series models.

[Bro99] Timothy Brook. *The confusions of pleasure: Commerce and culture in Ming China.* Univ of California Press, 1999.

[BST14] J Balz, C Sunstein, and R Thaler. Choice architecture. *E. Shafir, The behavioral foundations of public policy*, pages 428–439, 2014.

[BT03] Natilene Bowker and Keith Tuffin. Dicing with deception: People with disabilities' strategies for managing safety and identity online. *Journal of Computer-Mediated Communication*, 8(2):0–0, 2003.

[BT15] Anne Bowser and Janice Y. Tsai. Supporting Ethical Web Research: A New Research Ethics Review. In *Proceedings of the 24th International Conference on*

World Wide Web, WWW '15, pages 151–161, Republic and Canton of Geneva, Switzerland, 2015. International World Wide Web Conferences Steering Committee.

[BVR08] Monica Brito Vieira and David Runciman. Representation. *Cambridge: Polity*, 2008.

[BWCD11] Laura W Black, Howard T Welser, Dan Cosley, and Jocelyn M DeGroot. Self-governance through group discussion in Wikipedia: Measuring deliberation in online groups. *Small Group Research*, 42(5):595–634, 2011.

[BWDC12] Young Min Baek, Magdalena Wojcieszak, and Michael X Delli Carpini. Online versus face-to-face deliberation: Who? Why? What? With what effects? *New media & society*, 14(3):363–383, 2012.

[BWMB17] Lucas Böttcher, Olivia Woolley-Meza, and Dirk Brockmann. Temporal dynamics of online petitions. *PLOS ONE*, 12(5):1–12, 05 2017.

[Cam18] Baptiste Campion. Platforms are not neutral. `https://www.eurozine.com/platforms-not-neutral/`, 2018. [Online; accessed 10-June-2019].

[Cas04] Manuel Castells. *The network society A cross-cultural perspective*. Edward Elgar, 2004.

[Cas08] Manuel Castells. The New Public Sphere: Global Civil Society, Communication Networks, and Global Governance. *The ANNALS of the American Academy of Political and Social Science*, 616(1):78–93, 2008.

[Cas09] M Castells. *Communication power*. Oxford University Press, 2009.

[CBCCG17a] Iván Cantador, Alejandro Bellogín, María E. Cortés-Cediel, and Olga Gil. Personalized Recommendations in e-Participation: Offline Experiments for the 'Decide Madrid' Platform. In *Proceedings of the International Workshop on Recommender Systems for Citizens*, CitRec '17, pages 5:1–5:6, New York, NY, USA, 2017. ACM.

[CBCCG17b] Iván Cantador, Alejandro Bellogín, María E. Cortés-Cediel, and Olga Gil. Personalized Recommendations in e-Participation: Offline Experiments for the 'Decide Madrid' Platform. pages 5:1–5:6, 2017.

[CBDNML17] Justin Cheng, Michael Bernstein, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, pages 1217–1230, New York, NY, USA, 2017. ACM.

[CCC06] Roger J. Calantone, Kwong Chan, and Anna S. Cui. Decomposing Product Innovativeness and Its Effects on New Product Success. *Journal of Product Innovation Management*, 23(5):408–421, 2006.

[CCJ04] Michael X. Delli Carpini, Fay Lomax Cook, and Lawrence R. Jacobs. Public deliberation, discursive participation, and citizen engagement: A review of the empirical literature. *Annual Review of Political Science*, 7(1):315–344, 2004.

[CDK+19] Yan Chen, Shuyuan Deng, Dong-Heon Kwak, Ahmed Elnoshokaty, and Jiao Wu. A Multi-Appeal Model of Persuasion for Online Petition Success: A Linguistic

Cue-Based Approach. *Journal of the Association for Information Systems*, 20(2):3, 2019.

[CDNML15] Justin Cheng, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. Antisocial Behavior in Online Discussion Communities. In *Proceedings of the International AAAI Conference on Web and Social Media*, 2015.

[CFHH14] Derrick Coetzee, Armando Fox, Marti A. Hearst, and Björn Hartmann. Should Your MOOC Forum Use a Reputation System? In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '14, pages 1176–1187, New York, NY, USA, 2014. ACM.

[CG01] Stephen Coleman and John Gotze. *Bowling together: Online public engagement in policy deliberation.* Hansard Society London, 2001.

[CGWJ10] Zi Chu, Steven Gianvecchio, Haining Wang, and Sushil Jajodia. Who is Tweeting on Twitter: Human, Bot, or Cyborg? In *Proceedings of the 26th Annual Computer Security Applications Conference*, ACSAC '10, pages 21–30, New York, NY, USA, 2010. ACM.

[Cha03] Simone Chambers. Deliberative democratic theory. *Annual Review of Political Science*, 6(1):307–326, 2003.

[Cha06] Andrew Chadwick. *Internet politics: States, citizens, and new communication technologies.* Oxford University Press, USA, 2006.

[Cha09] Andrew Chadwick. Web 2.0: New challenges for the study of e-democracy in an era of informational exuberance. *Journal of Law and Policy for the Information Society*, 5:9–41, 2009.

[CHH02] Stephen Coleman, Nicola Hall, and Milica Howell. *Hearing voices: the experience of online public consultations and discussions in UK governance.* Hansard Society, 2002.

[Cho13] Daegon Cho. Real Name Verification Law on the Internet: A Poison or Cure for Privacy? In Bruce Schneier, editor, *Economics of Information Security and Privacy III*, pages 239–261, New York, NY, 2013. Springer New York.

[Chr11] Henrik Christensen. Political activities on the Internet: Slacktivism or political participation by other means? *First Monday*, 16(2), 2011.

[Cit12] Living Cities. Field scan of civic technology. *Living Cities. Retrieved from*, 2012.

[CK15] Yan Chen and Joseph Konstan. Online field experiments: a selective survey of methods. *Journal of the Economic Science Association*, 1(1):29–42, Jul 2015.

[CLHD17] Chi Ling Chan, Justin Lai, Bryan Hooi, and Todd Davies. The Message or the Messenger? Inferring Virality and Diffusion Structure from Online Petition Signature Data. In Giovanni Luca Ciampaglia, Afra Mashhadi, and Taha Yasseri, editors, *Social Informatics*, pages 499–517, Cham, 2017. Springer International Publishing.

[CLM17] Stephen Clark, Nik Lomax, and Michelle A. Morris. Classification of Westminster Parliamentary constituencies using e-petition data. *EPJ Data Science*, 6(1):16, Aug 2017.

[CNM83] Stuart K. Card, Allen Newell, and Thomas P. Moran. *The Psychology of Human-Computer Interaction*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1983.

[Cro04] Colin Crouch. *Post-democracy*. Polity Cambridge, 2004.

[CS08] Riley Crane and Didier Sornette. Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences*, 105(41):15649–15653, 2008.

[CSS+11] Anna Chmiel, Pawel Sobkowicz, Julian Sienkiewicz, Georgios Paltoglou, Kevan Buckley, Mike Thelwall, and Janusz A. Hołyst. Negative emotions boost user activity at BBC forum. *Physica A: Statistical Mechanics and its Applications*, 390(16):2936 – 2944, 2011.

[CTTJF16] Alceu Ferraz Costa, Agma Juci Machado Traina, Caetano Traina Jr, and Christos Faloutsos. Vote-and-Comment: Modeling the Coevolution of User Interactions in Social Voting Web Sites. 2016.

[Dah] Lincoln Dahlberg. Re-constructing digital democracy: An outline of four 'positions'.

[Dah01a] Lincoln Dahlberg. Computer-mediated communication and the public sphere: A critical analysis. *Journal of Computer-Mediated Communication*, 7(1):0–0, 2001.

[Dah01b] Lincoln Dahlberg. The Internet and democratic discourse: Exploring the prospects of online deliberative forums extending the public sphere. *Information, Communication & Society*, 4(4):615–633, 2001.

[Dah05] Peter Dahlgren. The Internet, Public Spheres, and Political Communication: Dispersion and Deliberation. *Political Communication*, 22(2):147–162, 2005.

[DFG+16] Pavel Dmitriev, Brian Frasca, Somit Gupta, Ron Kohavi, and Garnet Vaz. Pitfalls of long-term online controlled experiments. In *2016 IEEE International Conference on Big Data (Big Data)*, pages 1367–1376, 2016.

[DG09] Todd Davies and Seeta Peña Gangadharan. *Online deliberation: Design, research, and practice.* University of Chicago Press, 2009. CSLI Lecture Notes.

[DLH+15] Catherine L Dumas, Daniel LaManna, Teresa M Harrison, SS Ravi, Christopher Kotfila, Norman Gervais, Loni Hagen, and Feng Chen. Examining political mobilization of online communities through e-petitioning behavior in We the People. *Big Data & Society*, 2(2):2053951715598170, 2015.

[DP13] Donatella Della Porta. *Can democracy be saved?: participation, deliberation and social movements.* John Wiley & Sons, 2013.

[DVBZ+16] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3):554–559, 2016.

[DVJ03] Daryl J. Daley and David Vere-Jones. *An introduction to the theory of point processes. vol. I, Elementary theory and methods.* Probability and its applications. Springer, New York, Berlin, Paris, 2003.

[DZM+15] Nemanja Djuric, Jing Zhou, Robin Morris, Mihajlo Grbovic, Vladan Radosavljevic, and Narayan Bhamidipati. Hate Speech Detection with Comment Embeddings. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15 Companion, pages 29–30, New York, NY, USA, 2015. ACM.

[EDK16] Ahmed Said Elnoshokaty, Shuyuan Deng, and Dong-Heon Kwak. Success Factors of Online Petitions: Evidence from Change.Org. In *Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS)*, HICSS '16, pages 1979–1985, Washington, DC, USA, 2016. IEEE Computer Society.

[Edw02] Arthur R Edwards. The moderator as an emerging democratic intermediary: The role of the moderator in Internet discussions about public issues. *Information Polity*, 7(1):3–20, 2002.

[EFE17] Katharina Esau, Dennis Friess, and Christiane Eilders. Design Matters! An Empirical Analysis of Online Deliberation on Different News Platforms. volume 9, pages 321–342, 2017.

[Eis05] Eisenberg, Bryan. How to Improve A/B Testing. `https://www.clickz.com/how-to-improve-ab-testing/57741/`, 2005. [Online; accessed 01-June-2019].

[EKB16] Dean Eckles, René F. Kizilcec, and Eytan Bakshy. Estimating peer effects in networks with peer encouragement designs. *Proceedings of the National Academy of Sciences*, 113(27):7316–7322, 2016.

[Els98] Jon Elster. *Deliberative democracy*, volume 1. Cambridge University Press, 1998.

[FBA11] Flavio Figueiredo, Fabrício Benevenuto, and Jussara M Almeida. The tube over time: characterizing popularity growth of youtube videos. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, WSDM '11, pages 745–754, New York, NY, USA, 2011. ACM.

[FE15] Dennis Friess and Christiane Eilders. A systematic review of online deliberation research. *Policy & Internet*, 7(3):319–339, 2015.

[Fes62] Leon Festinger. *A theory of cognitive dissonance*, volume 2. Stanford university press, 1962.

[FGM05] Jenny Rose Finkel, Trond Grenager, and Christopher Manning. Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, ACL '05, pages 363–370, Stroudsburg, PA, USA, 2005. Association for Computational Linguistics.

[FGR16] Seth Flaxman, Sharad Goel, and Justin M. Rao. Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly*, 80(S1):298–320, 03 2016.

[Fis97] James S Fishkin. *The voice of the people: Public opinion and democracy.* Yale University Press, 1997.

[FKH00] Batya Friedman, Peter H. Khan, Jr., and Daniel C. Howe. Trust Online. *Commun. ACM*, 43(12):34–40, December 2000.

[FMN15] Rolf Fredheim, Alfred Moore, and John Naughton. Anonymity and Online Commenting: The Broken

Windows Effect and the End of Drive-by Commenting. In *Proceedings of the ACM Web Science Conference*, WebSci '15, pages 11:1–11:8, New York, NY, USA, 2015. ACM.

[For10] Santo Fortunato. Community detection in graphs. *Physics reports*, 486(3):75–174, 2010.

[Fox12] R. Fox. What Next for E-petitions. Hansard Society, 2012.

[FPDL06] Hugo Fuks, Mariano Pimentel, and Carlos José Pereira De Lucena. RU-Typing-2-Me? Evolving a chat tool to increase understanding in learning activities. *International Journal of Computer-Supported Collaborative Learning*, 1(1):117–142, 2006.

[FR09] Eva María Ferreras Rodríguez. Nuevas herramientas en el ecosistema digital. La promoción social de noticias en castellano. Análisis de Menéame. *Comunicación y pluralismo*, 8:155–167, 2009.

[FR11] Eva María Ferreras Rodríguez. New Proposals for the Creation of an Informative Agenda: The Case of Menéame. *Innovation: Economic, Social, and Cultural Aspects*, page 157, 2011.

[FS04] Bruno S. Frey and Meier Stephan. Social comparisons and pro-social behavior: Testing "conditional co-operation" in a field experiment. *The American Economic Review*, (5):1717–1722, 2004.

[FSW06] Danyel Fisher, Marc Smith, and Howard T Welser. You are who you talk to: Detecting roles in Usenet newsgroups. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*, volume 3, pages 59b–59b. IEEE, 2006.

[Fuc07] Christian Fuchs. *Internet and society: Social theory in the information age.* Routledge, 2007.

[Fuc13] Christian Fuchs. *Internet and surveillance: The challenges of Web 2.0 and social media*, volume 16. Routledge, 2013.

[FVD$^+$16] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. The Rise of Social Bots. *Commun. ACM*, 59(7):96–104, June 2016.

[GAHW15] Sharad Goel, Ashton Anderson, Jake Hofman, and Duncan J Watts. The structural virality of online diffusion. *Management Science*, 62(1):180–196, 2015.

[GAS$^+$15] David Garcia, Adiya Abisheva, Simon Schweighofer, Uwe Serdült, and Frank Schweitzer. Ideological and temporal components of network polarization in online political participatory media. *Policy & Internet*, 7(1):46–79, 2015.

[Gav91] William W. Gaver. Technology Affordances. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '91, pages 79–84, New York, NY, USA, 1991. ACM.

[GAW18] Kristina Gligorić, Ashton Anderson, and Robert West. How Constraints Affect Content: The Case of Twitter's Switch from 140 to 280 Characters, 2018.

[GBK09] E. Gilbert, T. Bergstrom, and K. Karahalios. Blogs are Echo Chambers: Blogs are Echo Chambers. In *2009 42nd Hawaii International Conference on System Sciences*, pages 1–10, Jan 2009.

[GBKB10] Sandra Gonzalez-Bailon, Andreas Kaltenbrunner, and Rafael E Banchs. The structure of political discussion networks: a model for the analysis of online deliberation. *Journal of Information Technology*, 25(2):230–243, Jun 2010.

[GGKST+14] Ruth García-Gavilanes, Andreas Kaltenbrunner, Diego Sáez-Trumper, Ricardo Baeza-Yates, Pablo Aragón, and David Laniado. Who are my audiences? a study of the evolution of target audiences in microblogs. In *International Conference on Social Informatics*, pages 561–572. Springer International Publishing, 2014.

[Gib77] James J Gibson. The theory of affordances. *Hilldale, USA*, 1:2, 1977.

[Gil10] Tarleton Gillespie. The politics of 'platforms'. *New Media & Society*, 12(3):347–364, 2010.

[GKK11a] Vicenç Gómez, Hilbert J. Kappen, and Andreas Kaltenbrunner. Modeling the Structure and Evolution of Discussion Cascades. In *Proceedings of the 22Nd ACM Conference on Hypertext and Hypermedia*, HT '11, pages 181–190, New York, NY, USA, 2011. ACM.

[GKK11b] Vicenç Gómez, Hilbert J. Kappen, and Andreas Kaltenbrunner. Modeling the Structure and Evolution of Discussion Cascades. In *Proceedings of the 22Nd ACM Conference on Hypertext and Hypermedia*, HT '11, pages 181–190, New York, NY, USA, 2011. ACM.

[GKL08] Vicenç Gómez, Andreas Kaltenbrunner, and Vicente López. Statistical Analysis of the Social Network and Discussion Threads in Slashdot. In *Proceedings of the 17th International Conference on World Wide Web*, WWW '08, pages 645–654. ACM, 2008.

[GKLK13] Vicenç Gómez, Hilbert J Kappen, Nelly Litvak, and Andreas Kaltenbrunner. A likelihood-based framework for the analysis of discussion threads. *World Wide Web*, 16(5):645–675, 2013.

[GKSA19] Ashish Goel, Anilesh K. Krishnaswamy, Sukolsak Sakshuwong, and Tanja Aitamurto. Knapsack Voting for Participatory Budgeting. *ACM Trans. Econ. Comput.*, 7(2):8:1–8:27, July 2019.

[GL04] Diego Garlaschelli and Maria I Loffredo. Patterns of link reciprocity in directed networks. *Physical review letters*, 93(26):268701, 2004.

[Gla10] Malcolm Gladwell. Small change. *The New Yorker*, 4(2010):42–49, 2010.

[GMKC+18] David Garcia, Yonas Mitike Kassa, Angel Cuevas, Manuel Cebrian, Esteban Moro, Iyad Rahwan, and Ruben Cuevas. Analyzing gender inequality through large-scale Facebook advertising data. *Proceedings of the National Academy of Sciences*, 115(27):6958–6963, 2018.

[Gra78] Mark Granovetter. Threshold Models of Collective Behavior. *American Journal of Sociology*, 83(6):1420–1443, 1978.

[GS83] Mark Granovetter and Roland Soong. Threshold models of diffusion and collective behavior. *The Journal of Mathematical Sociology*, 9(3):165–179, 1983.

[GSaH] Kimmo Grönlund, Kim Strandberg and, and Staffan Himmelroos. The challenge of deliberative democracy online - a comparison of face-to-face and virtual experiments in citizen deliberation.

[GW03] Todd Graham and Tamara Witschge. In search of online deliberation: Towards a new method for examining the quality of online discussions. *Communications-Sankt Augustin then Berlin*, 28(2):173–204, 2003.

[GWG12] Sharad Goel, Duncan J. Watts, and Daniel G. Goldstein. The Structure of Online Diffusion Networks. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, pages 623–638, New York, NY, USA, 2012. ACM.

[GWN14] Kurt Gray, Adrian F Ward, and Michael I Norton. Paying it forward: Generalized reciprocity and the limits of generosity. *Journal of experimental psychology: General*, 143(1):247, 2014.

[Hab84] Jürgen Habermas. *The theory of communicative action*, volume 2. Beacon press, 1984.

[Hab96] Jürgen Habermas. *Between facts and norms: Contributions to a discourse theory of law and democracy*. John Wiley & Sons, 1996.

[Har93] Linda Marie Harasim. *Global networks: Computers and international communication*. MIT Press, 1993.

[Haw15] Hawkeye. Harriet Sherwood: A Tool of Avaaz Astroturfing Campaign? `https://bitly.com/astroavaaz`, 2015. [Online; accessed 10-June-2019].

[HBKG04] Susan C Herring, S Barab, R Kling, and J Gray. An approach to researching online behavior. *Designing for virtual communities in the service of learning*, 338, 2004.

[HCN10] Khe Foon Hew, Wing Sum Cheung, and Connie Siew Ling Ng. Student contribution in asynchronous online discussion: a review of the research and empirical exploration. *Instructional Science*, 38(6):571–606, Nov 2010.

[HDD+17] Teresa M. Harrison, Catherine Dumas, Nic DePaula, Tim Fake, Will May, Akanksha Atrey, Jooyeon Lee, Lokesh Rishi, and S. S. Ravi. E-Petitioning and Online Media: The Case of #BringBackOurGirls. In *Proceedings of the 18th Annual International Conference on Digital Government Research*, dg.o '17, pages 11–20, New York, NY, USA, 2017. ACM.

[HG13] Daniel Halpern and Jennifer Gibbs. Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. *Computers in Human Behavior*, 29(3):1159–1168, 2013.

[HHM91] Jurgen Habermas, Jürgen Habermas, and Thomas Mccarthy. *The structural transformation of the public sphere: An inquiry into a category of bourgeois society.* MIT press, 1991.

[HHU+15] Loni Hagen, Teresa M. Harrison, Özlem Uzuner, Tim Fake, Dan Lamanna, and Christopher Kotfila. Introducing Textual Analysis Tools for Policy Informatics: A Case Study of e-Petitions. In *Proceedings of the 16th Annual International Conference on Digital Government Research*, dg.o '15, pages 10–19, New York, NY, USA, 2015. ACM.

[HHU+16] Loni Hagen, Teresa M Harrison, Özlem Uzuner, William May, Tim Fake, and Satya Katragadda. E-petition popularity: Do linguistic and semantic factors matter? *Government Information Quarterly*, 33(4):783–795, 2016.

[Hir05] Jorge E Hirsch. An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences of the United States of America*, 102(46):16569–16572, 2005.

[HJMY18] Scott A. Hale, Peter John, Helen Z. Margetts, and Taha Yasseri. How digital design shapes political participation: A natural experiment with social information. *PLOS ONE*, 13(4):1–20, 04 2018.

[HL99] Barry N Hague and Brian Loader. *Digital democracy: Discourse and decision making in the information age.* Psychology Press, 1999.

[HMY13] Scott A. Hale, Helen Z. Margetts, and Taha Yasseri. Petition Growth and Success Rates on the UK No. 10 Downing Street Website. In *Proceedings of the 5th Annual ACM Web Science Conference*, WebSci '13, pages 132–138, New York, NY, USA, 2013. ACM.

[Hor17] Lanka Horstink. Online Participation and the New Global Democracy: Avaaz, a Case Study. *Global Society*, 31(1):101–124, 2017.

[HSHH15] Shih-Wen Huang, Minhyang (Mia) Suh, Benjamin Mako Hill, and Gary Hsieh. How Activists Are Both Born and Made: An Analysis of Users on Change.org. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 211–220, New York, NY, USA, 2015. ACM.

[HTB+05] Hans G. K. Hummel, Colin Tattersall, Daniel Burgos, Francis Brouns, Hub Kurvers, and Rob Koper. Facilitating participation: From the EML web site to the Learning Network for Learning Design. *Interactive Learning Environments*, 13(1-2):55–69, 2005.

[HVD00] Kenneth L Hacker and Jan Van Dijk. *Digital Democracy: Issues of Theory and Practice.* SAGE Publications Ltd, 2000.

[HY09] Andrei Hagiu and David B Yoffie. What's your Google strategy? *Harvard Business Review*, 87(4):74–81, 2009.

[Jac12] Julie A. Jacko. *Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, Third Edition.* CRC Press, Inc., Boca Raton, FL, USA, 3rd edition, 2012.

[Jak01] Thomas Jakobsen. Advanced character physics. In *Game developers conference*, volume 3, 2001.

[JBB+17] Jason J. Jones, Robert M. Bond, Eytan Bakshy, Dean Eckles, and James H. Fowler. Social influence and political mobilization: Further evidence from a randomized experiment in the 2012 U.S. presidential election. *PLOS ONE*, 12(4):1–9, 04 2017.

[Jen03] Jakob Linaa Jensen. Public spheres on the Internet: Anarchic or government-sponsored–A comparison. *Scandinavian political studies*, 26(4):349–374, 2003.

[JJ04] Allan Jeong and Sunyoung Joung. The effects of response constraints and message labels on interaction patterns and argumentation in online discussions.

2003 Annual Proceedings-Anaheim: Volume# 2, page 201, 2004.

[JJ10] Andreas Jungherr and Pascal Jürgens. The Political Click: Political Participation through E-Petitions in Germany. *Policy & Internet*, 2(4):131–165, 2010.

[JK05] Davy Janssen and Raphaël Kies. Online Forums and Deliberative Democracy. *Acta Politica*, 40(3):317–335, Sep 2005.

[JVHB14] Mathieu Jacomy, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian. ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLOS ONE*, 9(6):1–12, 06 2014.

[KBK13] Ruogu Kang, Stephanie Brown, and Sara Kiesler. Why Do People Seek Anonymity on the Internet?: Informing Policy and Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 2657–2666, New York, NY, USA, 2013. ACM.

[KCL17] Srijan Kumar, Justin Cheng, and Jure Leskovec. Antisocial Behavior on the Web: Characterization and Detection. In *Proceedings of the 26th International Conference on World Wide Web Companion*, WWW '17 Companion, pages 947–950, Republic and Canton of Geneva, Switzerland, 2017. International World Wide Web Conferences Steering Committee.

[KDF+12] Ron Kohavi, Alex Deng, Brian Frasca, Roger Longbotham, Toby Walker, and Ya Xu. Trustworthy Online Controlled Experiments: Five Puzzling Outcomes Explained. In *Proceedings of the 18th ACM SIGKDD*

International Conference on Knowledge Discovery and Data Mining, KDD '12, pages 786–794, New York, NY, USA, 2012. ACM.

[KDF+13] Ron Kohavi, Alex Deng, Brian Frasca, Toby Walker, Ya Xu, and Nils Pohlmann. Online Controlled Experiments at Large Scale. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '13, pages 1168–1176, New York, NY, USA, 2013. ACM.

[KDLX14] Ron Kohavi, Alex Deng, Roger Longbotham, and Ya Xu. Seven Rules of Thumb for Web Site Experimenters. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, pages 1857–1866, New York, NY, USA, 2014. ACM.

[Kea09] John Keane. *The life and death of democracy.* Simon and Schuster, 2009.

[KGH14] Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24):8788–8790, 2014.

[KGL07] Andreas Kaltenbrunner, Vicenç Gómez, and Vicente Lopez. Description and prediction of slashdot activity. In *Web Conference, 2007. LA-WEB 2007. Latin American*, pages 57–66. IEEE, 2007.

[KGRDQV11] Andreas Kaltenbrunner, Gustavo Gonzalez, Ricard Ruiz De Querol, and Yana Volkovich. Comparative analysis of articulated and behavioural social networks in a social news sharing website. *New Review of Hypermedia and Multimedia*, 17(3):243–266, 2011.

[KH05] Peter G. Kilner and Christopher M. Hoadley. Anonymity Options and Professional Participation in an Online Community of Practice. In *Proceedings of th 2005 Conference on Computer Support for Collaborative Learning: Learning 2005: The Next 10 Years!*, CSCL '05, pages 272–280. International Society of the Learning Sciences, 2005.

[Kim06] Ji-Young Kim. The impact of Internet use patterns on political engagement: A focus on online deliberation and virtual social capital. *Information Polity*, 11(1):35–49, 2006.

[Kin02] Richard Kingston. The role of e-government and public participation in the planning process. In *XVI Aesop Congress Volos*, 2002.

[KL17] Ron Kohavi and Roger Longbotham. *Online Controlled Experiments and A/B Testing*, pages 922–929. Springer US, Boston, MA, 2017.

[KLSH09] Ron Kohavi, Roger Longbotham, Dan Sommerfield, and Randal M. Henne. Controlled experiments on the web: survey and practical guide. *Data Mining and Knowledge Discovery*, 18(1):140–181, Feb 2009.

[KMM10] Ravi Kumar, Mohammad Mahdian, and Mary McGlohon. Dynamics of conversations. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '10, pages 553–562. ACM, ACM, 2010.

[Kol33] Andrey Kolmogorov. Sulla Determinazione Empirica di una Legge di Distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, 4:83–91, 1933.

[KS02] Peter Kollock and Marc Smith. *Communities in cyberspace*. Routledge, 2002.

[KS19] Emre Kiciman and Amit Sharma. Causal Inference and Counterfactual Reasoning (3Hr Tutorial). In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, WSDM '19, pages 828–829, New York, NY, USA, 2019. ACM.

[KWL16] Srijan Kumar, Robert West, and Jure Leskovec. Disinformation on the Web: Impact, Characteristics, and Detection of Wikipedia Hoaxes. In *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, pages 591–602, Republic and Canton of Geneva, Switzerland, 2016. International World Wide Web Conferences Steering Committee.

[LBK09] Jure Leskovec, Lars Backstrom, and Jon Kleinberg. Meme-tracking and the dynamics of the news cycle. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 497–506. ACM, 2009.

[LFI04] Robert C Luskin, James S Fishkin, and Shanto Iyengar. Considered opinions on US foreign policy: Face-to-face versus online deliberative polling. *International Communication Association, New Orleans, LA*, pages 1369–1387, 2004.

[LGRC12] Janette Lehmann, Bruno Gonçalves, José J Ramasco, and Ciro Cattuto. Dynamical Classes of Collective Attention in Twitter. pages 251–260, 2012.

[LH13] Yu-Hao Lee and Gary Hsieh. Does Slacktivism Hurt Activism?: The Effects of Moral Balancing and Consistency in Online Activism. In *Proceedings of the*

SIGCHI Conference on Human Factors in Computing Systems, CHI '13, pages 811–820, New York, NY, USA, 2013. ACM.

[Lin12] Dennis Linders. From e-government to we-government: Defining a typology for citizen coproduction in the age of social media. *Government Information Quarterly*, 29(4):446 – 454, 2012. Social Media in Government - Selections from the 12th Annual International Conference on Digital Government Research (dg.o2011).

[LL10] David S Lee and Thomas Lemieux. Regression discontinuity designs in economics. Technical Report 2, 2010.

[LP19] Juan De Dios Llorens and David Pérez. Clusterización de propuesta en Decide Madrid mediante NLP. `http://bit.ly/medialab-nlp`, 2019. [Online; accessed 10-September-2019].

[LR09] Ralf Lindner and Ulrich Riehm. Electronic petitions and institutional modernization. International parliamentary e-petition systems in comparative perspective. *JeDEM-eJournal of eDemocracy and Open Government*, 1(1):1–11, 2009.

[LR11] Ralf Lindner and Ulrich Riehm. Broadening Participation Through E-Petitions? An Empirical Study of Petitions to the German Parliament. *Policy & Internet*, 3(1):1–23, 2011.

[Lum16] Alberto Lumbreras. *Automatic role detection in online forums.* PhD thesis, Ecole doctorale de InfoMaths (ED 512) LYON, 2016.

[Mar15] Alba Marín. El usuario prosumidor en la promoción social de las noticias: estudio etnográfico en Menéame. In *Desafíos éticos de la comunicación en la era digital: III Congreso Internacional de Ética de la Comunicación, libro de actas. Facultad de Comunicación de la Universidad de Sevilla, días 24, 25 y 26 de marzo de 2015. Luis Alfonso Guadarrama Rico, Juan Carlos Suárez Villegas y María del Mar González Jiménez (eds.).(pp. 500-513). Sevilla: Universidad de Sevilla*, 2015.

[Mar17] Helen Z. Margetts. Political behaviour and the acoustics of social media. *Nature Human Behaviour*, 1(4):s41562–017, 2017.

[Mat19a] J. Nathan Matias. Preventing harassment and increasing group participation through social norms in 2,190 online science discussions. *Proceedings of the National Academy of Sciences*, 2019.

[Mat19b] J. Nathan Matias. The Civic Labor of Volunteer Moderators Online. *Social Media + Society*, 5(2):2056305119836778, 2019.

[MC04] Elizabeth Murphy and Elizabeth Coleman. Graduate Students' Experiences of Challenges in Online Asynchronous Discussions. *Canadian Journal of Learning and Technology / La revue canadienne de l'apprentissage et de la technologie*, 30(2), 2004.

[McF12] Colin McFarland. *Experiment!: Website conversion rate optimization with A/B and multivariate testing.* New Riders, 2012.

[McG] Michael McGuire. Collaborative public management: Assessing what we know and how we know it.

[MCLPdL15] Arnau Monterde, Antonio Calleja-López, and Andrés Pereira de Lucena. Disseny del procés de participació digital del programa d'actuació municipal (PAM) i de districtes (PAD) 2016-2019. page 93, 2015.

[McV07] J Gregory McVerry. Forums and functions of threaded discussions. *New England Reading Association Journal*, 43(1):79, 2007.

[MD13] Helen Margetts and Patrick Dunleavy. The second wave of digital-era governance: a quasi-paradigm for government on the Web. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1987):20120382, 2013.

[Meg16] Colin Megill. pol.is in Taiwan. `https://blog.pol.is/pol-is-in-taiwan-da7570d372b5`, 2016.

[Men11] José Candón Mena. La dimensión híbrida del movimiento 15M: entre lo físico y lo virtual. In *Actas del V Congreso Online del Observatorio para la Cibersociedad "Hybrid Days"*, volume 16, 2011.

[Mer68] Robert K. Merton. The Matthew Effect in Science. *Science*, 159(3810):56–63, 1968.

[MGSP17] Jonathan Mellon, Hollie Russon Gilman, Fredrik M. Sjoberg, and Tiago. Peixoto. Gender and Political Mobilization Online: Participation and Policy Success on a Global Petitioning Platform. In Tony Saich, editor, *Ash Center Occasional Papers*. Harvard Kennedy School, 2017.

[Min] Seong-Jae Min. Online vs. face-to-face deliberation: Effects on civic engagement.

[Mit04] Michael Mitzenmacher. A brief history of generative models for power law and lognormal distributions. *Internet mathematics*, 1(2):226–251, 2004.

[MJHR15] Helen Z. Margetts, Peter John, Scott A Hale, and Stéphane Reissfelder. Leadership without leaders? Starters and followers in online collective action. *Political Studies*, 63(2):278–299, 2015.

[MJHY15] Helen Z. Margetts, Peter John, Scott Hale, and Taha Yasseri. *Political turbulence: How social media shape collective action*. Princeton University Press, 2015.

[MJRH12] Helen Z. Margetts, Peter John, Stephane Reissfelder, and Scott A. Hale. Social Influence and Collective Action: An Experiment Investigating the Effects of Visibility and Social Information Moderated by Personality. 2012.

[MM18] J. Nathan Matias and Merry Mou. CivilServant: Community-Led Experiments in Platform Governance. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 9:1–9:13, New York, NY, USA, 2018. ACM.

[MMLMB19] Momin M Malik, Katja Mayer, Hemank Lamba, and Claudia Müller-Birn. Workshop on Critical Data Science. 2019.

[MO+09] Enrique Orduña Malea, José Antonio Ontalba, et al. Propuesta de indicadores métricos para gestores sociales de noticias: análisis de la prensa digital española en Menéame. *Information research*, 14(3):2, 2009.

[Mor09] Evgeny Morozov. The brave new world of slacktivism. *Foreign policy*, 19(05), 2009.

[Mor11] Evgeny Morozov. *The Net Delusion: The Dark Side of Internet Freedom*. Perseus Books, Cambridge, MA, USA, 2011.

[Mor13] Evgeny Morozov. *To save everything, click here: The folly of technological solutionism*. Public Affairs, 2013.

[MP16] Momin Malik and Jürgen Pfeffer. Identifying Platform Effects in Social Media Data. In *Proceedings of the International AAAI Conference on Web and Social Media*, 2016.

[MR00] Michael Margolis and David Resnick. *Politics as usual: The Cyberspace 'Revolution'*, volume 6. SAGE Publications Ltd, 2000.

[MSLC01] Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, pages 415–444, 2001.

[Mül11] Daniel Müllner. Modern hierarchical, agglomerative clustering algorithms. *arXiv preprint arXiv:1109.2378*, 2011.

[Mur12] Kevin P Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012.

[NBG11] Mor Naaman, Hila Becker, and Luis Gravano. Hip and trendy: Characterizing emerging trends on Twitter. *Journal of the American Society for Information Science and Technology*, 62(5):902–918, 2011.

[Nov09] Beth Simone Noveck. *Wiki government: how technology can make government better, democracy stronger, and citizens more powerful*. Brookings Institution Press, 2009.

[NTO+16] Ryosuke Nishi, Taro Takaguchi, Keigo Oka, Takanori Maehara, Masashi Toyoda, Ken-ichi Kawarabayashi, and Naoki Masuda. Reply trees in Twitter: data analysis and branching process models. *Social Network Analysis and Mining*, 6(1):26, 2016.

[Ogi08] Paul Ogilvie. Modeling blog post comment counts. `http://livewebir.com/blog/2008/07/modeling-blog-post-comment-counts`, 2008.

[OKC18] Alexandra Olteanu, Emre Kiciman, and Carlos Castillo. A Critical Review of Online Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, WSDM '18, pages 785–786, New York, NY, USA, 2018. ACM.

[O'R05] Tim O'Reilly. What is Web 2.0: Design patterns and business models for the next generation of software. *Communications & strategies*, (1):17, 2005.

[O'R11] Tim O'Reilly. Government as a Platform. *Innovations: Technology, Governance, Globalization*, 6(1):13–40, 2011.

[OS12] Teresa Piñeiro Otero and Carmen Costa Sánchez. Ciberactivismo y redes sociales. El uso de facebook por uno de los colectivos impulsores de la 'spanish revolution', Democracia Real Ya (DRY). *Observatorio (OBS*)*, 6(3), 2012.

[OW05] Stanisław Osiński and Dawid Weiss. Carrot2: Design of a flexible and efficient web information retrieval framework. In *International Atlantic Web Intelligence Conference*, pages 439–444. Springer, 2005.

[PACM16] Julia Proskurnia, Karl Aberer, and Philippe Cudré-Mauroux. Please Sign to Save...: How Online Environmental Petitions Succeed. In *Proceedings of the Workshop on Social Web for Environmental and Ecological Monitoring*, 2016.

[Pap] Zizi Papacharissi. The virtual sphere: The internet as a public sphere.

[Pap04] Zizi Papacharissi. Democracy online: civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, 6(2):259–283, 2004.

[PB12] Etienne Pelaprat and Barry Brown. Reciprocity: Understanding online social relations. *First Monday*, 17(10), 2012.

[PBS17] Cornelius Puschmann, Marco Toledo Bastos, and Jan-Hinrik Schmidt. Birds of a feather petition together? Characterizing e-petitioning through the lens of platform data. *Information, Communication & Society*, 20(2):203–220, 2017.

[PCK09] Victor Pascual-Cid and Andreas Kaltenbrunner. Exploring Asynchronous Online Discussions through Hierarchical Visualisation. In *Proceedings of the 2009 13th International Conference Information Visualisation*, pages 191–196, Washington, DC, USA, 2009. IEEE Computer Society.

[PCW07] Gabriele Plickert, Rochelle R Cote, and Barry Wellman. It's not who you know, it's how you know them: Who exchanges what with whom? *Social networks*, 29(3):405–429, 2007.

[PE12] Panagiotis Panagiotopoulos and Tony Elliman. *Online Engagement from the Grassroots: Reflecting on over a Decade of ePetitioning Experience in Europe and the UK*, pages 79–94. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[PFdL03] Mariano Gomes Pimentel, Hugo Fuks, and Carlos José Pereira de Lucena. Co-text loss in textual chat tools. In *Modeling and Using Context*, pages 483–490. Springer, 2003.

[PG05] Susanne Pfalzner and Paul Gibbon. *Many-body tree methods in physics*. Cambridge University Press, 2005.

[PGK+17] Julia Proskurnia, Przemyslaw Grabowicz, Ryota Kobayashi, Carlos Castillo, Philippe Cudré-Mauroux, and Karl Aberer. Predicting the Success of Online Petitions Leveraging Multidimensional Time-Series. In *Proceedings of the 26th International Conference on World Wide Web*, WWW '17, pages 755–764, 2017.

[Poo05] Nathaniel Poor. Mechanisms of an Online Public Sphere: The Website Slashdot. *Journal of Computer-Mediated Communication*, 10(2):00–00, 2005.

[Por18] Alexander Porshnev. Success Factors of Electronic Petitions at Russian Public Initiative Project: The Role of Informativeness, Topic and Lexical Information. In Steffen Staab, Olessia Koltsova, and Dmitry I. Ignatov, editors, *Social Informatics*, pages 243–250, Cham, 2018. Springer International Publishing.

[Pos14] John Postill. Democracy in an age of viral reality: A media epidemiography of Spain's indignados movement. *Ethnography*, 15(1):51–69, 2014.

[PSB+14] Antonella Passani, Francesca Spagnoli, Francesco Bellini, Alessandra Prampolini, and Katja Firus. *Collective Awareness Platform for Sustainability and Social Innovation: An Introduction*. 2014.

[PSGH13] Mayur Patel, Jon Sotsky, Sean Gourley, and Daniel Houghton. The emergence of civic tech: Investments in a growing field. *The Knight Foundation*, 2013.

[RCM+11] Jacob Ratkiewicz, Michael Conover, Mark Meiss, Bruno Gonçalves, Alessandro Flammini, and Filippo Menczer. Detecting and Tracking Political Abuse in Social Media. In *Proc. 5th International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2011.

[RFM10] Jacob Ratkiewicz, Alessandro Flammini, and Filippo Menczer. Traffic in social media I: paths through information networks. In *Social Computing (SocialCom), 2010 IEEE Second International Conference on*, pages 452–458. IEEE, 2010.

[RG08] Pierre Rosanvallon and Arthur Goldhammer. *Counter-democracy: Politics in an Age of Distrust*, volume 7. Cambridge University Press, 2008.

[Ros17] Aliza Rosen. Giving you more characters to express yourself, 2017.

[RPB01] Jacques Rancière, Davide Panagia, and Rachel Bowlby. Ten theses on politics. *Theory & event*, 5(3), 2001.

[RR83] Paul R Rosenbaum and Donald B Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

[RS97] Sheizaf Rafaeli and Fay Sudweeks. Networked interactivity. *Journal of Computer-Mediated Communication*, 2(4):0–0, 1997.

[SCB00] Marc Smith, Jonathan J Cadiz, and Byron Burkhalter. Conversation trees and threaded chats. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, CSCW '00, pages 97–105, New York, NY, USA, 2000. ACM.

[Sch78] Thomas C Schelling. *Micromotives and macrobehavior*. WW Norton & Company, 1978.

[Sch97] Steven Michael Schneider. *Expanding the public sphere through computer-mediated communication : political discussion about abortion in a Usenet newsgroup*. PhD thesis, Massachusetts Institute of Technology, 1997.

[SDCC19] Jorge Saldivar, Florian Daniel, Luca Cernuzzi, and Fabio Casati. Online Idea Management for Civic Engagement: A Study on the Benefits of Integration with Social Networking. *Trans. Soc. Comput.*, 2(1):1:1–1:29, January 2019.

[Sea10] Paul Seabright. *The company of strangers: A natural history of economic life*. Princeton University Press, 2010.

[Sel11] Joel Seligstein. Trends in Facebook Messages, 2011.

[SG07] Jennifer Stromer-Galley. Measuring deliberation's content: A coding scheme. *Journal of public deliberation*, 3(1), 2007.

[SG10] Schweitzer, Frank and Garcia, David. An agent-based model of collective emotions in online communities. *Eur. Phys. J. B*, 77(4):533–545, 2010.

[Sim55] Herbert A Simon. On a class of skew distribution functions. *Biometrika*, 42(3-4):425–440, 1955.

[SJ14] Jan-Hinrik Schmidt and Katharina Johnsen. On the use of the e-petition platform of the German Bundestag. 2014.

[SJSN09] Graham Smith, Peter John, Patrick Sturgis, and Hisako Nomura. Deliberation and internet engagement: initial findings from a randomised controlled trial evaluating the impact of facilitated internet forums. In *ECPR General Conference*, 2009.

[SK86] Lee Sproull and Sara Kiesler. Reducing Social Context Cues: Electronic Mail in Organizational Communication. *Management Science*, 32(11):1492–1512, 1986.

[SL15] Matthew J. Salganik and Karen E. C. Levy. Wiki Surveys: Open and Quantifiable Social Data Collection. *PLOS ONE*, 10(5):1–17, 05 2015.

[Smi48] Nickolay Smirnov. Table for estimating the goodness of fit of empirical distributions. 19(2):279–281, 1948.

[SPA+19] Jorge Saldivar, Cristhian Parra, Marcelo Alcaraz, Rebeca Arteta, and Luca Cernuzzi. Civic Technology for Social Innovation. *Computer Supported Cooperative Work (CSCW)*, 28(1):169–207, Apr 2019.

[SPRG13] Tiziano Squartini, Francesco Picciolo, Franco Ruzzenenti, and Diego Garlaschelli. Reciprocity of weighted networks. *Scientific reports*, 3, 2013.

[SPS14] Ingo Scholtes, René Pfitzner, and Frank Schweitzer. The Social Dimension of Information Ranking: A Discussion of Research Challenges and Approaches. In

Katharina Zweig, Wolfgang Neuser, Volkmar Pipek, Markus Rohde, and Ingo Scholtes, editors, *Socioinformatics - The Social Impact of Interactions between Humans and IT*, pages 45–61, Cham, 2014. Springer.

[SS98] Peter Sprent and Peter Sprent. *Data driven statistical methods*. Number 04; QA276. 4, S6. 1998.

[Sta02] Richard M Stallman. What is free software. *Free Society: Selected Essays of*, 23, 2002.

[Sum] Summit. Propensity Score Matching. `https://www.summitllc.us/propensity-score-matching`. [Online; accessed 01-June-2019].

[Sun01] Cass R. Sunstein. *Republic.Com 2.0*. Princeton University Press, Princeton, NJ, USA, 2001.

[Sun02] Cass R. Sunstein. The Law of Group Polarization. *Journal of Political Philosophy*, 10(2):175–195, 2002.

[Sun09] Cass R Sunstein. Republic. com 2.0, 2009.

[SW03] Jerzy Stefanowski and Dawid Weiss. Carrot2 and Language Properties in Web Search Results Clustering. In *Proceedings of the 1st International Atlantic Web Intelligence Conference on Advances in Web Intelligence*, AWIC'03, pages 240–249, Berlin, Heidelberg, 2003. Springer-Verlag.

[TC60] Donald L Thistlethwaite and Donald T Campbell. Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, 51(6):309, 1960.

[TCLM+15] Javier Toret, Antonio Calleja-López, Óscar Marín, Pablo Aragón, Miguel Aguilera, Xabier Barandiaran, and Arnau Monterde. *Tecnopolítica y 15M: La potencia de las multitudes conectadas.* Editorial UOC, 2015.

[Tho53] Robert L. Thorndike. Who belongs in the family? *Psychometrika*, 18(4):267–276, Dec 1953.

[Tor07] Lars Hasselblad Torres. Citizen sourcing in the public interest. *Knowledge Management for Development Journal*, 3(1):134–145, 2007.

[Tor15] Simon Tormey. *The end of representative politics.* John Wiley & Sons, 2015.

[Tré04] Matthias Trénel. Measuring the deliberativeness of online discussions. Coding scheme 2.4. *report, Berlin: Social Science Research Centrex*, 2004.

[Tri10] Ana I Bernal Triviño. Presencia de las redes sociales en los cibermedios españoles. *Textual & Visual Media: revista de la Sociedad Española de Periodística*, 3:25–42, 2010.

[TS09] Richard H Thaler and Cass R Sunstein. *Nudge: Improving decisions about health, wealth, and happiness.* Penguin, 2009.

[TSH17] Nathan TeBlunthuis, Aaron Shaw, and Benjamin Mako Hill. Density Dependence Without Resource Partitioning: Population Ecology on Change.org. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17 Companion, pages 323–326, New York, NY, USA, 2017. ACM.

[Tuf15] Zeynep Tufekci. Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency. *Colo. Tech. LJ*, 13:203, 2015.

[TWdR09] Manos Tsagkias, Wouter Weerkamp, and Maarten de Rijke. Predicting the Volume of Comments on Online News Stories. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, CIKM '09, pages 1765–1768, New York, NY, USA, 2009. ACM.

[U.S13] U.S. Congress. H.R.1123 - Unlocking Consumer Choice and Wireless Competition Act. `https://www.congress.gov/bill/113th-congress/house-bill/1123`, 2013. [Online; accessed 01-June-2019].

[VD99] Jan Van Dijk. *The network society*. Sage Publications, 1999.

[VD13] José Van Dijck. *The culture of connectivity: A critical history of social media*. Oxford University Press, 2013.

[VDPDW18] José Van Dijck, Thomas Poell, and Martijn De Waal. *The platform society: Public values in a connective world*. Oxford University Press, 2018.

[vdRAWF16] Arnout van de Rijt, Idil Afife Akin, Robb Willer, and Matthew Feinberg. Success-Breeds-Success in Collective Political Behavior: Evidence from a Field Experiment. *Sociological Science*, 3(41):940–950, 2016.

[vdRKRP14] Arnout van de Rijt, Soong Moon Kang, Michael Restivo, and Akshay Patil. Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences*, 111(19):6934–6939, 2014.

[VN03] Gina Danielle Venolia and Carman Neustaedter. Understanding sequence and reply relationships within email conversations: a mixed-model visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 361–368, New York, NY, USA, 2003. ACM.

[VTA+15] Gabriela Gonzalez Vaillant, Juhi Tyagi, Idil Afife Akin, Fernanda Page Poma, Michael Schwartz, and Arnout van de Rijt. A Field-Experimental Study of Emergent Mobilization in Online Collective Action. *Mobilization: An International Quarterly*, 20(3):281–303, 2015.

[WG75] Henry William Watson and Francis Galton. On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4:138–144, 1875.

[WG99] Barry Wellman and Milena Gulia. Net surfers don't ride alone: Virtual communities as communities. *Networks in the global village*, pages 331–366, 1999.

[WGGGM16] Claudia Wagner, Eduardo Graells-Garrido, David Garcia, and Filippo Menczer. Women through the glass ceiling: gender asymmetries in Wikipedia. *EPJ Data Science*, 5(1):5, Mar 2016.

[WH07] Fang Wu and Bernardo A. Huberman. Novelty and collective attention. *Proceedings of the National Academy of Sciences*, 104(45):17599–17601, 2007.

[Wig96] Jerry S Wiggins. *The five-factor model of personality: Theoretical perspectives*. Guilford Press, 1996.

[Wik07] Wikipedia. Internet in Indonesia. `https://en.wikipedia.org/wiki/Internet_in_Indonesia`, 2007. [Online; accessed 10-June-2019].

[Wil98] Anthony G Wilhelm. Virtual sounding boards: How deliberative is on-line political discussion? *Information, Communication & Society*, 1(3):313–338, 1998.

[Wil00] Anthony G. Wilhelm. *Democracy in the Digital Age: Challenges to Political Life in Cyberspace.* Routledge, New York, NY, 10001, 2000.

[Wri09] Scott Wright. The role of the moderator: problems and possibilities for government-run online discussion forums. *Online deliberation: Design, research, and practice*, pages 233–242, 2009.

[Wri12] Scott Wright. Assessing (e-)Democratic Innovations: "Democratic Goods" and Downing Street E-Petitions. *Journal of Information Technology & Politics*, 9(4):453–470, 2012.

[Wri15] Scott Wright. Populism and Downing Street E-petitions: Connective Action, Hybridity, and the Changing Nature of Organizing. *Political Communication*, 32(3):414–433, 2015.

[WRM05] Angus Whyte, Alistair Renton, and Ann Macintosh. e-Petitioning in Kingston and Bristol. *Edinburgh, UK: International Teledemocracy Centre, Napier University*, 2005.

[WS07] Scott Wright and John Street. Democracy, deliberation and design: the case of online discussion forums. *New Media & Society*, 9(5):849–869, 2007.

[WYH12] Chunyan Wang, Mao Ye, and Bernardo A Huberman. From user comments to on-line conversations. In *Proceedings of the 18th ACM SIGKDD International*

Conference on Knowledge Discovery and Data Mining, KDD '12, pages 244–252, New York, NY, USA, 2012. ACM, ACM.

[WZH13] Tim Weninger, Xihao Avi Zhu, and Jiawei Han. An exploration of discussion threads in social news sites: A case study of the Reddit community. In *Advances in Social Networks Analysis and Mining (ASONAM), 2013 IEEE/ACM International Conference on*, pages 579–583. IEEE, 2013.

[XCF+15] Ya Xu, Nanyu Chen, Addrian Fernandez, Omar Sinno, and Anmol Bhasin. From Infrastructure to Culture: A/B Testing Challenges in Large Scale Social Networks. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, pages 2227–2236, New York, NY, USA, 2015. ACM.

[Yen17] Tzu-Chi Yen. Understanding Online Survey Data as a Bipartite Network: A Case Study on pol.is. In *XXXVII Sunbelt Conference of the International Network for Social Network Analysis (insna)*, 2017.

[YHM17] Taha Yasseri, Scott A. Hale, and Helen Z. Margetts. Rapid rise and decay in petition signing. *EPJ Data Science*, 6(1), 2017.

[ZCP08] Xiang Zhou, Yuen-Ying Chan, and Zhen-Mei Peng. Deliberativeness of online political discussion: A content analysis of the Guangzhou Daily website. *Journalism Studies*, 9(5):759–770, 2008.

[ZGR+14] Matteo Zignani, Sabrina Gaito, Gian Paolo Rossi, Xiaohan Zhao, Haitao Zheng, and Ben Zhao. Link and Triadic Closure Delay: Temporal Metrics for Social Network Dynamics, 2014.

[Žiž00] Slavoj Žižek. *The ticklish subject: The absent centre of political ontology.* Verso, 2000.

[Zub19] Shoshana Zuboff. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* PublicAffairs, 1st edition, 2019.

[Zuc14] Ethan Zuckerman. New Media, New Civics? *Policy & Internet*, 6(2):151–168, 2014.

[ZVK17] Amy X. Zhang, Lea Verou, and David Karger. Wikum: Bridging Discussion Forums and Wikis Using Recursive Summarization. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, pages 2082–2096, New York, NY, USA, 2017. ACM.