



EJ: A Free Software Platform for Social Participation

Fábio Macêdo Mendes^{1(✉)}, Ricardo Poppi^{1,2}, Henrique Parra²,
and Bruna Moreira¹

¹ Universidade de Brasilia (UnB), Brasília, Brazil
{fabiomendes, brunamoreira}@unb.br

² Instituto Cidade Democrática, São Paulo, Brazil
{ricardo, henrique}@cidadedemocratica.org.br

Abstract. As the Internet grows on importance as a forum for political activity, it is necessary to occupy it with proper tools for democratic discussion, dialogue and deliberation. Currently, a substantial part of political debate is conducted on social media inside proprietary networks. Those solutions are flagrantly inadequate to build consensus seeking understandings and to mediate the interaction between the government and the citizenry. This work present EJ, a platform for crowd-sourced social participation which uses machine learning based intelligence and gamification techniques to increase engagement and counteract the formation of opinion bubbles and the “echo chamber” effect of social networks.

Keywords: Social participation · E-gouvernement · Web · Free software

1 Introduction

As a growing proportion of the world’s population gains access to the Internet, so grows its importance as a place for public discourse and social organization. An optimistic stance on this fact points to a networked democracy in which information flows horizontally and the common citizen can exert real decision making power and produce solutions to societal problems. On the other hand, the same Internet has allowed mechanisms that support privacy violation, surveillance [1] and spread of misinformation [6] in a scale never seen before. For better or for worse, it is unquestionable that the Internet plays an important role in shaping societies and political systems of the present and of the future [3].

This contradiction of potentialities is perhaps better demonstrated in today’s largest social network, Facebook, with its more than 2 billion users¹. The rise of modern social networks has indeed delivered novel forms of horizontal communication and has given voice to actors that would otherwise be excluded from traditional processes. At the same time, Facebook’s business model is centered

¹ <https://newsroom.fb.com/company-info/>.

around keeping users connected to the platform to deliver targeted marketing content, while gathering any information useful for profiling. This model has been frequently abused to obtain audience and political gain via questionable practices such as inflammatory and sensationalist discourse, disinformation, hate speech², etc; all very detrimental to further democratic values, albeit frequently treated as a tolerable source of revenue by the social network itself³.

We believe that a platform for social participation must be explicitly designed to be resilient to those forms of manipulation rather than arise as an accessory feature of a larger social network. Moreover, participation technologies (digital and otherwise) should be seen as social goods and thus must employ greater levels of transparency and public oversight than for-profit products. An acceptable level of public control of software products is perhaps only attained by using open source tools and technologies.

This report introduces a platform for social participation named EJ (Empurrando Juntas) that aims to fill a specific niche in this universe. We present EJ both at technical level and contextualize it in a more conceptual and perhaps sociological discussion. Technical aspects are considered in Sects. 3 and 4, and technically oriented readers are invited to go there to understand the rationale behind the technology and the most important architectural decisions. The rest of the paper is dedicated to analyzing the role of the platform in promoting participation and the challenges we perceived to creating the networked democracy hinted in the first paragraph.

2 Platforms for Social Participation

Compared to more traditional processes for consultation and public debate, the Internet enables a much faster flow of information which helps making large scale decentralized organizations viable. Those features have certainly sparked many initiatives towards building tools for a networked democracy that aims to bring prominence to the common citizen in shaping collective decisions. Those initiatives can usually be categorized in 4 broader areas (with many possible intersections): transparency⁴, petitions and public awareness⁵, debating⁶ and deliberation⁷. We can also classify initiatives with regard to its relationship with the State, in which some take a decisively anti-government stance (e.g., Wikileaks⁸), while others are directly conducted or sponsored by the State.

While we acknowledge that all of those areas are similarly important, we decided to focus our efforts on debating and consultation. We took this path out

² <https://www.theguardian.com/technology/2018/apr/06/myanmar-facebook-criticise-mark-zuckerberg-response-hate-speech-spread>.

³ <https://www.propublica.org/article/facebook-enforcement-hate-speech-rules-mistakes>.

⁴ <https://www.opengovpartnership.org/>, <https://serenata.ai/>.

⁵ <https://avaaz.org>, <https://www.change.org/>.

⁶ <https://debatemap.live/>, <https://www.artikulate.in/>.

⁷ <https://decidim.org/>, <https://arxiv.org/abs/1707.06526>.

⁸ <https://wikileaks.org/>.

of the belief that the two major features of a platform for social participation should be the possibility of probing opinions and listening to novel inputs that users might add to the debate. The archetypal tools for handling each of those cases are the discussion forum and the poll. The first is useful to construct and refine positions and ideas in an unstructured fashion, in open dialogue; the second probes perceptions of the crowd from a predefined set of options.

Internet forums have (an often deserved) reputation for leading to toxic behaviors. Even when discounting for this possibility, participation in those environments demand users to invest time and knowledge, presenting a substantial barrier for engagement. There are of course many examples of groups of engaged and uniquely qualified individuals that produce rich discussions and communities, but we believe that the discussion forum model is severely limited when applicable to wide audiences.

If only a minority of interested users decide to participate, discussions may feature an unfair representation. This fragility undermines the political weight of those initiatives by making it easier to portray conclusions as instances of group-think and undemocratic. Excluding the rare cases in which a hard consensus is obtained, it is hard to elicit opinions in a structured form. How then should the result of open discussions be used to support a collective decision?

While platforms for open ended discussion produce barriers to participation and engagement, polls suffer from the opposite problem: the cost of participation is very low, but more qualified users are unable to enrich the debate or introduce novel ideas even when willing to do so. Polling make it relatively easy to educe the preferences of the majority since data is created in a quantitative form. However, users might feel alienated from the decision making process if they perceive all options as unappealing or, worst yet, as being designed to fulfill some specific agenda.

2.1 Crowdsourcing

In order to bridge those two approaches, we believe that the design of collective deliberation must consider the pedagogical aspect of interaction. A good example of digital tools with accessible interfaces that take elements from both ends are the so called “crowdsourced discussion” architectures⁹ [4,15]. As a crude approximation, those platforms are based on a polling mechanism in which the list of alternatives can be continuously expanded with user interaction.

Participants of those platforms can control the energy dedicated to each discussion by switching from “voting” to “making new contributions” modes. Since users are constantly evaluating contributions from other users, it should be possible to automate data analysis to understand and identify different opinion profiles.

A good example of this kind of platform is the open source application Pol.is. In its current state, it provides a simple interaction in which administrators propose a theme for discussion and users can express their positions either by

⁹ <https://www.allourideas.org/>.

publishing a simple paragraph of text or by evaluating other user opinions. Pol.is then uses machine learning algorithms to identify affinity groups and present this information to all participants. While this is a promising approach, we believe that there are improvements to be made on how to use this information to promote some kind of effective collective action. We use Pol.is as the main inspiration of our approach, but add an additional layer of gamification in order to put some incentives for user to engage and create prolific debates and deliberation.

2.2 Social Networks

The enormous user base of social network websites (Facebook alone comprises almost a third of human population) guarantees a diverse audience. In spite of that, each user may experience very low diversity of ideas: algorithms that perform content selection end up placing users into “echo chambers” that recirculate a small set of ideas without challenging them with input from different social groups [12]. Diversity of viewpoints are rarely appreciated and debates tend to be contentious. This dynamics shifts the goal of political struggle from seeking democratic compromises and consensus to establishing (real or apparent) numeric majorities to subject the opposition to the rule of the crowd.

Those proprietary networks use their position as keepers of giant volumes of data to control the flow of information and influence user behaviours. One plausible response to those threats is to advocate for decentralization. This crucial aspect can be handled in many different ways, the most simple one is distribution of a software product as open source. This gives the possibility of users and organizations to manage their own deploys and set their own rules of governance.

Beyond that, some projects advocate for protocols and formats that enable different instances to communicate and share data in a federated fashion. Notable projects in this area include PeerTube¹⁰ and Mastodon¹¹. More radical approaches like Tim Berners-Lee’s Solid [16] and the Dat network [11] even calls for an overhaul of the Internet structure to favor local computation, granular access controls and P2P communication. It is conceivable those projects they might experience short term success, but it is unlikely to be in a scale that significantly changes the current panorama of social networks in spite of the very interesting technical achievements.

2.3 The Role of the State in Participation Platforms

It seems to be a clear desire of governmental and civil society organizations to have tools for democratic dialogue and direct communication. There are many examples of governmental entities that sponsored initiative based on collective intelligence, but most of those experiences suffered from difficulties in scaling or generating the desired impact.

¹⁰ <https://joinpeertube.org/>.

¹¹ <https://joinmastodon.org/>.

A notable experience took place in Brazil during the consultations for the Marco Civil da Internet (Brazil Rights Framework for the Internet), in which an existing Open Source Social Network developed in Brazil called Nosfero¹² was used. This led to number of other public consultations using similar technologies for collective deliberation. State’s collaboration with those open source communities to build shared digital goods preserves sovereignty, while ensuring that citizens have autonomy to develop civic technologies independent of a direct custody from a State actor.

3 Conception and Architecture

EJ was conceived to work both as a standalone product and as an embedded component in some existing application. In both cases, we expect that most of our user base will access the platform using mobile devices since that number is growing worldwide and already accounts to most of web traffic, nowadays¹³. Mobile access implies that we have the opportunity to engage our users in a continuous manner, either by expecting constant availability, or by using more active approaches such as push notifications and web apps.

3.1 User Experience

EJ organizes topics of discussion around “conversations”. Each conversation is introduced with a single paragraph of text that should try to frame how users should contribute to the discussion, like in Fig. 1. Users are then able to insert comments associated with a conversation or vote if they agree or disagree with other user’s propositions.

As users interact with some conversation (either proposing new comments or by voting on propositions), it gradually become possible to recognize the different opinion profiles. As soon as the platform is able to classify users into distinct opinion groups, it display results to all participants. That way, not only owners of each conversation can extract useful metrics to guide policy and decision making, but they can also justify their decisions from the shared information available on the platform. Moreover, each participant is able to access how their own opinions, or those from their network of peers and compare with the whole. We believe that this form of transparency towards information is an important factor to counteract the “echo chamber” effect of traditional social networks.

Only minimal personal information is requested from new users and none of it is exposed publicly like a social network. A potential threat of this model is that it

¹² <http://noosfero.org/>.

¹³ StatCounter, a firm specialized on web traffic analysis, shows worldwide mobile usage slightly above desktop at 52.98% vs 47.02% (<http://gs.statcounter.com/platform-market-share/desktop-mobile-tablet/worldwide>). A more detailed account (<https://www.stonetemple.com/mobile-vs-desktop-usage-study/>) shows even large gaps depending on the selected metric (e.g., page views, unique users, time on site, etc).



Fig. 1. EJ user interface showing a conversation

provides little safeguards against bots and other malicious users since it lowers the barrier for creating new accounts. We decided not to tackle this difficult problem, but rather to make it possible to exclusively use external authentication via OAuth and rely on an identity provider trusted by deployment.

3.2 Machine Learning

Machine learning enriches the platform in the detection of opinion groups and understanding how each user relates to each group. In Machine Learning jargon, detection of opinion groups is a non-supervised classification problem. We tried a few classical methods (K-means and SVM) with moderate success. More reliable outcomes are obtained when making allowance for prior information. Our approach to classification requires conversations to be created in association with some stereotypical opinion groups. Stereotypes are controlled by the conversation author and behave similarly to regular users when interacting with comments. Those users define their initial centroids in the clusterization algorithm and serve as labels that more reliably identify clusters across different classifications.

EJ encodes votes as +1 (agree), -1 (disagree) and 0 (skip). The set of user interactions is thus a vector of components in that range. In a typical situation, this data set comprises mostly of missing data entries, and requires a good imputation strategy. We used a mean-value based imputation for simplicity and

computational efficiency and better results were obtained by using collaborative filtering [17], but at a significant efficiency cost.

Traditional K-means consists of 4 steps: (1) centroid initialization, (2) labeling elements to the closest centroid, (3) re-calculation of each centroid, and (4) go to 2 until convergence. Our modification uses stereotypes as the initial centroids in step (1) and forces that each stereotype will stay in its own cluster by never re-labeling them in step (2). This simple modification not only is able to introduce some form of prior information, but has the positive side-effect of generating predictable labels for each run of the classification algorithm.

The default visualization display cluster sizes, similarities and diversity in a stylized manner. We also employ classic machine learning algorithms for dimension reduction aimed at other visualization options. For instance, a scatter plot of user inputs displays the opinion of each user together with the stereotypes. Our empirical evaluation selected principal component analysis (PCA) as opposed to more sophisticated alternatives such as t-SNE [9] or MDS [5] for its superior reliability and efficiency. PCA is a rotation based method, which preserves some intuitive geometric relationships between elements, making it more accessible to non-experts.

3.3 Web Platform

As a web native platform, EJ subscribes to a few broad principles that guides all architectural decisions.

Progressive enhancement: Web applications are created using HTML, CSS and JavaScript. Progressive enhancement postulates that the website must be usable in its most basic HTML-only form and each new layer (CSS, JavaScript) should progressively enhance the user experience instead of being a hard requirement.

Mobile first design: UX (user experience) and UI (user interaction) is designed primarily for mobile, and desktop experience is enhanced with CSS media queries. It also implies that the application should be usable from poor connections on underpowered devices. We prioritize small JavaScript size, since loading, parsing and executing implies non-trivial CPU costs which blocks the main application thread.

Progressive web app: EJ only uses functionality available in the web platform. We developed it as Progressive Web App (PWA) in order to avoid the costs of a native implementation. PWAs blur the distinction of native mobile applications and web sites. Users can install an icon on their home screens and it spawns a background process that allows push notifications and offline interaction. Those are not intrusive technologies and we were able to develop most of the platform largely ignoring any specific needs for “PWAs”.

API as a product: Since we expect EJ to be easily embedded on user applications, the REST API is designed as an independent product and should be able to easily expose any information or interaction to third parties. The API also implements business logic regarding authentication and authorization policies in the platform.

3.4 Technology Stack

EJ was initially conceived as fork of the open source project Pol.is¹⁴, which is based on Node.js and Clojure (for a in-house implementation of Machine Learning algorithms). This initial experience was afflicted with a bad deployment story and poor communication with Pol.is core development team. This lead to the decision of abandoning Pol.is and change the stack in favor of Python.

EJ uses the Django Framework¹⁵ for web development and the Pydata stack (Numpy [20], Pandas [10], Scikit-Learn [13]) for building the data analysis pipelines. On the frontend side, it uses Sass to manage CSS and adopts a somewhat traditional web architecture in which pages are rendered server-side rather than by the client. This approach is sometimes touted as ROCA¹⁶ (Resource Oriented Client Architecture) to confront with the recent SPA (Single Page Application) architectural trend, and is perhaps nothing more than adoption of what used to be considered the best practices for web development a few years ago. Following the principles of ROCA, we used a library for unobtrusive JavaScript named Unpoly¹⁷ and decided to avoid client-side frameworks completely.

By following the principle of progressive enhancement, it was possible to create a product that is accessible by many clients, ranging from command line web-browsers, web spiders, screen readers to the evergreen versions of major Browsers. While we expect most users to lie in the last category, progressive enhancement also benefit this audience by providing a workable (but degraded) website even when connection is poor and downloads fail.

4 Development and Results

EJ is still in active development and has released only a single stable version. We do not have data about its real use performance. This section rather discuss how development was organized and which results were obtained.

4.1 Team

EJ has gone through two different development phases, each one carried by a different team. In the first phase, it was developed by HackLab/São Paulo¹⁸, which is a socially oriented Startup that employs and develops free software technologies. This initial phase was focused on a product for the Perseu Abramo Foundation, which is a Brazilian think tank associated with the Partido dos Trabalhadores (Workers' Party, of ex-president Lula da Silva fame). In the second phase, development moved to the University, in which part of the original team

¹⁴ <https://pol.is/>.

¹⁵ <https://www.djangoproject.com/>.

¹⁶ <https://roca-style.org>.

¹⁷ <https://unpoly.com/>.

¹⁸ <https://hacklab.com.br/>.

was replaced by students. This report describes the second phase of development in which all authors took an active role and was a collaboration with the Federal Government to create a participation platform for teenagers via the National Secretary for the Rights of Children and Adolescents.

The team was divided roughly into 4 parts: design/UX, frontend, backend and operations. Since EJ was developed as a University project in a Software Engineering department, we took into account pedagogical outcomes when allocating team members. Every two weeks, students could then negotiate the allocation in each of the last 3 teams in which we encouraged a moderate level of rotation so everyone could experience different aspects of development (sometimes at expense of raw productivity).

4.2 Continuous Delivery and Deployment

Continuous deployment was an important part of the development culture in EJ. After some adjusting, we finally ended up organized our Git repository¹⁹ into two main branches: a master branch which only receives tagged release commits and a development branch that is used to integrate contributions from all developers. Any commit in those branches that pass an automated test suite automatically trigger a deploy script pointing to some specific instance of the platform.

The system is also integrated with Codeclimate²⁰, a static analysis tool that tracks the evolution of code quality metrics, test coverage and can alert to the introduction of defects and regressions. Since most development takes place in feature-specific branches, those warnings can prevent some defects from entering into the main development branch. A negative side effect of requiring manual merging of feature branches with explicit code review is that it burdens team members assigned to those tasks. We decided that changes deemed low risk (e.g., updates to translation strings, documentation, small CSS fixes or improvements to the test suite) should be developed directly into the development branch without any external review step before merging.

Generally speaking, productivity increases when good tools and automations are available [2]. Since we made the initial decision of hosting the code on Github, the choice of tools was largely guided by the options that platform already provides (e.g., issue tracker, project management Kanban board) and have a good story of integration.

Since a large part of the development team is comprised of students with high rotation rate and considerable learning curve, we decided to avoid emphasizing too much the role of processes and culture and tried to rely heavily on automation. The objective is to gain better efficiency and shorter learning curve due to better tooling, at the same time we guarantee all contributions to the development process are crystallized in the form of code. Creation of automation scripts is thus treated as an important part of the development as is writing application code. This includes code responsible for deployment and testing, but

¹⁹ <https://github.com/ejplatform/ej-server/>.

²⁰ <https://codeclimate.com/>.

also automation of everyday chores such as creating new environments, building assets and test databases, managing and executing Docker images, etc.

5 Conclusions

Building effective and constructive social participation mechanisms is a chief problem of contemporary democracies. Citizens are increasingly reporting a sense of misrepresentation and frustration with democracy worldwide, and the perceived efficacy of the government is diminishing [19]. Some scholars have pointed that this growing scepticism on the capacity of democracies to solve social issues may lead us to an unprecedented state of “a recession” [7], in the number of democracies in the world. All this in a time in which humanity faces unique and very pressing challenges in politics [8], economy [14], environment preservation [18], health, education and many others aspects of life.

Of course, none of those problems have simple solutions. We, as technologists, may offer tools that improve the collective capacity for finding answers to those dilemmas. This report introduces EJ, a simple system of participation that can be used to improve the dialogue between the government and autonomous organizations with the population. While the technology has not been adopted in any large deployment yet, we believe that it was built taking into consideration important aspects of collective participation.

Our two fundamental working hypothesis is that EJ is able to promote a higher level of engagement than existing platforms for social participation, and that its mechanisms to counteract polarization are effective for building civil dialogue and discussion. While engagement is easy to measure, polarization requires more indirect observation. Existence of clearly separated opinion groups is not necessarily an indication of polarization, but might rather stem from valid and rational disagreements between different perspectives. Nevertheless, it is still possible to probe some indirect effects of polarization. Polarized discussions arise from unfair and distorted perceptions about the other group’s motivations and points of view and generally produces a mismatch in familiarity with basic facts about a subject and acceptance of fake and incorrect information.

EJ was built on the notion that socially relevant technologies should be created as common goods: making it free (as in freedom) for anyone to use, to study, and to adapt. This provides a good level of control and oversight and protects collectivity from vested interests from corporate owners of proprietary solutions.

Acknowledgements. The authors would like to thank Fundação Perseu Abramo and the former Ministry of Human Rights (now transformed into a secretary by the current Brazilian government) for recognizing the importance of direct social participation in shaping public policy and for financial support.

References

1. Albrechtslund, A.: Online social networking as participatory surveillance. *First Monday* **13**(3) (2008). <https://doi.org/10.5210/fm.v13i3.2142>, <http://journals.uic.edu/ojs/index.php/fm/article/view/2142>

2. Barry, E.J., Kemerer, C.F., Slaughter, S.A.: How software process automation affects software evolution: a longitudinal empirical analysis. *J. Softw. Maint. Evol.: Res. Pract.* **19**(1), 1–31 (2007). <https://doi.org/10.1002/smr.342>
3. Best, M.L., Wade, K.W.: The internet and democracy: global catalyst or democratic dud? *Bull. Sci. Technol. Soc.* **29**(4), 255–271 (2009). <https://doi.org/10.1177/0270467609336304>
4. Brabham, D.C.: Crowdsourcing the public participation process for planning projects. *Plan. Theory* **8**(3), 242–262 (2009). <https://doi.org/10.1177/1473095209104824>
5. Cox, T.F., Cox, M.A.A.: Multidimensional scaling. In: No. 88 in *Monographs on statistics and applied probability*, 2nd edn. Chapman & Hall/CRC, Boca Raton (2001)
6. Del Vicario, M., et al.: The spreading of misinformation online. *Proc. Natl. Acad. Sci.* **113**(3), 554 (2016). <https://doi.org/10.1073/pnas.1517441113>
7. Diamond, L.: Facing up to the democratic recession. *J. Democr.* **26**(1), 141–155 (2015). <https://doi.org/10.1353/jod.2015.0009>
8. Levitsky, S., Way, L.: The rise of competitive authoritarianism. *J. Democr.* **13**(2), 51–65 (2002). <https://doi.org/10.1353/jod.2002.0026>
9. van der Maaten, L., Hinton, G.: Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**(Nov), 2579–2605 (2008)
10. McKinney, W.: Data structures for statistical computing in python. In: *Proceedings of the 9th Python in Science Conference (SCIPY 2010)*, pp. 51–56 (2010). [https://doi.org/10.1016/S0168-0102\(02\)00204-3](https://doi.org/10.1016/S0168-0102(02)00204-3)
11. Ogden, M., Mc Kelvey, K., Madsen, M.B., et al.: Dat-distributed dataset synchronization and versioning. *Open Sci. Fram.* **10** (2017)
12. Pariser, E.: *The Filter Bubble: What The Internet Is Hiding From You*. Penguin Books Limited (2011). <https://books.google.com.br/books?id=-FWO0puw3nYC>
13. Pedregosa, F., et al.: *Scikit-learn: machine learning in python* (2012). <https://doi.org/10.1007/s13398-014-0173-7.2>. <http://arxiv.org/abs/1201.0490>
14. Piketty, T., Goldhammer, A.: *Capital in the twenty-first century* (2017). oCLC: 1063105013
15. Salganik, M.J., Levy, K.E.C.: Wiki surveys: open and quantifiable social data collection. *PLOS ONE* **10**(5), e0123483 (2015). <https://doi.org/10.1371/journal.pone.0123483>, <https://dx.plos.org/10.1371/journal.pone.0123483>
16. Sambra, A.V., et al.: *Solid: a platform for decentralized social applications based on linked data* (2016)
17. Sarwar, B., Karypis, G., Konstan, J., Riedl, J.: Item-based collaborative filtering recommendation algorithms. In: *Proceedings of the 10th International Conference on World Wide Web, WWW 2001*, pp. 285–295. ACM, New York (2001). <https://doi.org/10.1145/371920.372071>
18. Masson-Delmotte, V., et al.: IPCC. Global warming of 1.5 $\hat{\text{A}}^{\circ}\text{C}$. An IPCC special report on the impacts of global warming of 1.5 $\hat{\text{A}}^{\circ}\text{C}$ above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate c, p. 792 (2018). <https://doi.org/10.1017/CBO9781107415324>
19. Van Reybrouck, D.: *Against Elections* (2013)
20. van der Walt, S., Colbert, S.C., Varoquaux, G.: The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* **13**(2), 22–30 (2011). <https://doi.org/10.1109/MCSE.2011.37>