

FILTERING DISSENT

Social Media and Land Struggles in Brazil

PUBLIC DISCOURSE IS increasingly mediated by proprietary software systems owned by a handful of major corporations. Google, Facebook, Twitter and YouTube claim billions of active users for their social media platforms, which automatically run filtering algorithms to determine what information is displayed to those users on their feeds. A feed is typically organized as an ordered list of items. Filtering algorithms select which items to include and how to order them. Far from being neutral or objective, these algorithms are powerful intermediaries that prioritize certain voices over others. An algorithm that controls what information rises to the top and what gets suppressed is a kind of gatekeeper that manages the flow of data according to whatever values are written into its code. In the vast majority of cases, platforms do not inform users about the filtering logics they employ—still less offer them control over those filters. As a ubiquitous, automated, powerful, and yet largely secret and unexamined form of information control, this filtering process deserves more critical attention.¹ Its implications for euphoric predictions about political mobilization in a new information age—exemplified by talk of ‘Facebook revolutions’ in the Arab world—have yet to be fully explored.

There is little doubt that filtering algorithms can serve political purposes effectively. In 2013, Facebook researchers conducted experiments to test whether manipulations of its algorithm could change user moods and voting behaviour, varying the number of posts containing positive or negative emotional words in the feeds of 689,003 users.² They claimed to have found evidence of ‘massive-scale emotional contagion’; that is, people who saw posts with either more positive or more negative words

were more likely to write posts with the same emotional bias. In another experiment, during the 2010 US congressional elections, Facebook inserted an item into the feeds of 60 million users that encouraged them to vote.³ Its researchers then cross-referenced the names of users with actual voting records and concluded that users with manipulated feeds were more likely to vote: they even claimed that the manipulation had increased turnout by 340,000. If such manipulation was directed towards specific social and political groups, which is already possible through the paid sponsorship of filtering, it could determine the outcome of an election.⁴ Significant attempts to sway elections in several Latin American countries through more straightforwardly criminal abuses of social media have already been documented.⁵

Some platforms employ a combination of algorithmic filters and human curators. The latter are typically low-wage contractors, whose involvement recently became the subject of a major controversy: in May 2016, former Facebook ‘news curators’—young American journalists subcontracted through Accenture—anononymously accused the platform of routinely suppressing right-wing content in its ‘trending topics’, which

¹ Frank Pasquale (in *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge, MA 2015) and Nicholas Diakopoulos (in *Algorithmic Accountability Reporting: On the Investigation of Black Boxes*, New York 2014) consider the algorithmic management of information in contexts as varied as financial markets and search engines, but do not examine the algorithmic filtering of feeds in depth or detail. Others, such as Cass Sunstein (in *Republic.com 2.0*, Princeton 2009) and Eli Pariser (in *The Filter Bubble: What the Internet Is Hiding from You*, New York 2011), focus on the ‘echo chamber’ effect that results from ‘personalization’ and that limits exposure to divergent viewpoints, as we discuss below.

² Adam Kramer et al., ‘Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks’, *Proceedings of the National Academy of Sciences*, vol. 111, no. 24, 17 June 2014.

³ Robert Bond et al., ‘A 61-Million-Person Experiment in Social Influence and Political Mobilization’, *Nature*, vol. 489, no. 7415, 13 September 2012. The item was a direct exhortation to vote, displaying photos of ‘friends’ who had voted, a link to find the user’s polling station, and a button to tell others that ‘I voted’.

⁴ For a discussion of this Facebook experiment and the lack of legal regulation on feed manipulation, see Jonathan Zittrain, ‘Engineering an Election’, *Harvard Law Review Forum*, vol. 127, 20 June 2014.

⁵ The cracker Andrés Sepúlveda was reportedly paid large sums for services including manipulation of social media during election campaigns in Mexico, Colombia and Venezuela; for example, using software that could direct an army of fake social media accounts to generate false ‘trends’. See Jordan Robertson et al., ‘How to Hack an Election’, *Bloomberg*, 31 March 2016.

appear as a list of news items separate from the main feed, and which supposedly prioritize the most ‘popular’ news topics of the day.⁶ American conservatives jumped on the accusations, claiming that Facebook has a liberal bias, and prompting an inquiry from the Republican chair of the Senate Commerce Committee. In its defence, Facebook responded that the human curators merely ‘review’ stories that are ‘surfaced by an algorithm’—as if algorithmic filtering automatically assures neutrality—while claiming to stand ‘for a global community . . . giving all people a voice, for a free flow of ideas and culture across nations’.⁷

There are deep flaws with both the conservatives’ charge and Facebook’s response. The official list of ‘1,000 trusted sources’ for trending topics actually includes many right-wing news outlets, but very few on the left.⁸ Moreover, there have been more serious and better-documented cases of censorship by Facebook ‘content moderators’ that have been largely neglected by the mainstream press. In 2012, for example, a former moderator leaked Facebook’s list of abuse standards, whose ‘international compliance’ section prohibited any content critical of the Turkish government or Kemal Atatürk, or in support of the Kurdistan Workers’ Party.⁹ This censorship occurred not in the small box of trending topics, but in the main feed. Thus, in comparison with the suppression of leftist dissent, the conservatives’ charge is weak in both substance and evidence. It is also striking that the issue of Facebook’s non-neutrality in the selection of news topics was raised by the revelation that humans are involved in the editorial process; the implication throughout the controversy has often been that the use of filtering algorithms is unbiased and

⁶ The accusations surfaced after the website *Gizmodo* exposed details of Facebook’s secretive ‘trending news project’ and published the allegations of news suppression. See Michael Nunez, ‘Want to Know What Facebook Really Thinks of Journalists? Here’s What Happened When It Hired Some’, 3 May 2016; and Nunez, ‘Former Facebook Workers: We Routinely Suppressed Conservative News’, 9 May 2016.

⁷ See the following Facebook press releases: Justin Osofsky, ‘Information About Trending Topics’, 12 May 2016; and Colin Stretch, ‘Response to Chairman John Thune’s Letter on Trending Topics’, 23 May 2016.

⁸ Sam Thielman, ‘Facebook news selection is in hands of editors not algorithms, documents show’, *Guardian*, 12 May 2016. Facebook’s May 23 response to the US Senate inquiry states that it ‘will no longer rely on lists of external websites and news outlets’ for trending topics. However, it does not name any substitutes and thus leaves its news selection even less transparent.

⁹ Rosa Gilbert and Alfredo Mazzamauro, ‘Silencing dissent: Turkey’s crackdown on press freedom’, *Roar Magazine*, 17 March 2016. See *Roar’s* sources for the list of abuse standards.

objective. As we will demonstrate, algorithmic filtering routinely suppresses some political perspectives and promotes others, independently of human ‘editorial’ intervention.

In other words, overt censorship of the internet—for example, server takedown, seizure of domain names, denial of service and editorial manipulation—is not necessary to control the flow of information for political purposes. Algorithmic filtering can accomplish the same end implicitly and continuously through its logics of promotion and suppression.¹⁰ In the algorithmic control of information, there are no clearly identifiable censors or explicit acts of censorship: the filtering is automated and inconspicuous, with a tangled chain of actors (computer scientists, lines of code, private corporations and user preferences). This complex process systematically limits the diversity of voices online and in many cases suppresses certain kinds of speech. Although the outcome may be viewed as tantamount to censorship, we need to broaden our conceptual framework to take account of the specific logics that are built into the selection, distribution and display of information online.

In what follows, we will describe how filtering algorithms work on the leading social media platforms, before going on to explain why those platforms have adopted particular filtering logics, and how those logics structure a political economy of information control based primarily on advertising and selling consumer products. Political activists regularly use such platforms for outreach and mobilization. What are the consequences of relying on commercial logics to manage political speech? We show the impact of algorithmic filtering on a contemporary social conflict, the land disputes between agribusiness and the Guarani and Kaiowá peoples in Mato Grosso do Sul, Brazil. The predominant filtering logics result in various forms of information promotion and suppression that negatively affect indigenous activists and benefit the agribusiness lobby—but we also show how activists can sometimes

¹⁰ We distinguish algorithmic from internet filtering, of which there is extensive documentation, showing how intermediaries such as service providers deny access to data: Jonathan Zittrain and John Palfrey, ‘Internet Filtering’, in Ronald Deibert et al., eds, *Access Denied*, Cambridge, MA 2008. There are many mechanisms for internet filtering: see Steven Murdoch and Ross Anderson, ‘Tools and Technology of Internet Filtering’, in Deibert et al., eds, *Access Denied*. Usually, private companies impose internet filtering to comply with state-issued directives. Most countries in Latin America, North America and Europe do not have systematic regimes of technical internet filtering, but algorithmic filtering is ubiquitous worldwide.

strengthen their voices by circumventing those logics in creative ways. In conclusion, we will propose a number of strategies to subvert the predominant logics of information control and to nurture alternatives that would enable a more democratic circulation of information online. Given the overwhelming importance of online mediation for social and political life, this is an urgent task.

Filtering logics

How does algorithmic filtering work? What are its predominant logics today?¹¹ Filtering algorithms typically determine a selection and order of items in a feed by calculating numerical scores for each item in a database based on user actions. If an item has a high score, its position will be higher and therefore more visible. A recent Facebook study demonstrates that items in top positions are more likely to be clicked on.¹² Platforms gather data for the calculation of feed positions from the surveillance of user actions. The constant tracking of clicks, browsing histories and communication patterns provides the data on which algorithms operate. Some data consist of direct user input such as clicks of buttons, including ‘likes’ on Facebook and ‘retweets’ on Twitter. Other data involve sophisticated tracking of involuntary input, such as how much time a user spends viewing each item before scrolling down. In some cases, surveillance reaches beyond the platform itself. Installed on many websites as promotional tools, Facebook’s ‘like’ and Twitter’s ‘tweet’ buttons also run background operations to track all visitors to those sites. Both companies use this surreptitiously obtained information for profiling, advertising, filtering and other purposes. Algorithmic filtering makes such surveillance profitable.

Algorithms are usually not disclosed, much less explained, to users. Platforms claim that they select content that is ‘most interesting’ (Twitter), ‘top’ (Facebook), or ‘best’ (Reddit) for the user. Some, like Google Plus, do not describe their selection criteria at all. We have identified some of the actual filtering logics that these platforms employ

¹¹ By logic, we mean both the structure of operations that are written in code and the encoded values; in other words, a conceptual framework that bundles together specific calculations and variables. For example, if sorting the feed by popularity (however defined) is the value, then an algorithm processes a structured set of operations to determine which items in the database are presented as most popular.

¹² See Figure S5a in the supplement to Eytan Bakshy et al., ‘Exposure to Ideologically Diverse News and Opinion on Facebook’, *Science*, 7 May 2015.

(Table 1), using three methods of analysis. First, we tested the platforms, observing how our actions affected our feeds and reading what other users said about their own. Second, we compared the feeds of different user accounts, including some that people use regularly and others that we created solely for the purposes of our experiment. Third, we analyzed patent documents, corporate reports and research papers about filtering algorithms to decipher their technical operations and trace their historical developments. We combined these three methods to reveal the structure of operations of filtering algorithms and to generate a conceptual index of their logics.¹³ Based on our synthesis, we can identify the following logics:

1. *Popularity*. Items that a platform considers more popular (however determined) are more visible.
2. *Similarity*. Users deemed similar are more likely to see items that each other posts or interacts with (e.g., by recommending, ‘liking’ and commenting on).
3. *Social ties*. Users with strong ties (e.g., ‘friends’) are more likely to see items that each other posts or interacts with.
4. *Paid sponsorship*. An item becomes more visible if a sponsor pays to push it to the top of the feed.
5. *Subscription*. Users are more likely to see items related to keywords, tags, other users and groups that they select to follow.
6. *Time*. Newer items are more visible to users.

Algorithmic definitions of popularity vary by platform, but they typically incorporate measures of user approval, attention and activity. The popularity of an item may be based on approval as registered, for example, by a count of ‘likes’ on Facebook. On Reddit, users can vote submissions ‘up’ to promote, or ‘down’ to suppress; their scores are the differences between the counts of up and down votes. It is often said that filtering by popularity makes a platform democratic, a claim based on a limited concept of democracy in which divergent views are unlikely to appear because they do not receive enough up-votes.¹⁴ Popularity may also include measures of user activity and attention, such as the sum total of

¹³ We cannot confirm that our index is comprehensive because other logics may be undisclosed and undetected.

¹⁴ Reddit moderators recently made that claim in a *New York Times* op-ed: Brian Lynch and Courtnie Swearingen, ‘Why We Shut Down Reddit’s “Ask Me Anything” Forum’, *NYT*, 8 July 2015.

TABLE 1: *Filtering logics of selected social media platforms*

<i>Platform</i>	<i>Monthly Active Users*</i>	<i>Institutional Control</i>	<i>Software</i>	<i>Filtering Logics</i>
Facebook	1.45 billion (March 2015)	for-profit corporation	proprietary	popularity, similarity, social ties, paid sponsorship, subscription, time
YouTube	1 billion (March 2013)	for-profit corporation	proprietary	popularity, similarity, paid sponsorship, subscription, time
Google Plus	300–540 million, contested (October 2013)	for-profit corporation	proprietary	popularity, similarity, social ties, paid sponsorship, subscription, time
Twitter	288 million (December 2014)	for-profit corporation	proprietary	paid sponsorship, subscription, time
Reddit	159 million (February 2015)	for-profit corporation	free, open-source	popularity, paid sponsorship, subscription, time
Diaspora	34,380, partial (February 2015)	non-profit foundation, decentralized servers	free, open-source	subscription, time

* This is an industry standard measure, but as neither ‘active’ nor ‘user’ is precisely defined, it is not fully consistent. Corporations that own the platforms report the counts. For Diaspora, partial counts were collected with the opt-in pod statistics feature on 12 February 2015. Google Plus statistics are contested because profiles are automatically created and only a fraction of registered users actually visits the platform.

comments, views and viewing durations. In these cases, the logic of popularity has a cumulative effect: items with more views appear in more feeds, thus garnering even more views.

The best-known system for ranking items by popularity is PageRank, the first version of Google’s search engine algorithm, which company founders Larry Page and Sergey Brin developed as a research project at Stanford in 1996. Although they describe the PageRank score of a web page as an objective measure of its importance, it is more accurately seen as a measure of popularity, since it is based upon how many other pages link to the page, increasing the weight of those pages that are themselves linked from many others.¹⁵ In essence, PageRank is a more sophisticated version of an academic citation count, with the main difference being that citation counts do not consider whether citing articles are themselves popular. Although PageRank determines the popularity of web pages specifically for the purposes of a search engine,

¹⁵ Page, Brin et al., *The PageRank Citation Ranking: Bringing Order to the Web* (Technical Report), Stanford InfoLab, 1999.

it was an influence on EdgeRank, the first version of Facebook’s feed algorithm. Facebook stopped using the name EdgeRank in 2011, when it updated its algorithm to incorporate machine-learning techniques. It now considers thousands of variables and automatically ‘learns’ how to make use of them to maximize user attention and advertising profit.

Similarity-based filtering draws on statistical correlations between the actions and characteristics of users, with the goal of displaying items that match users’ tastes, opinions and preferences. A general method emerged in the early 1990s, which researchers at MIT and Xerox PARC termed ‘collaborative filtering’ and ‘social information filtering’.¹⁶ However, since the technique merely tracks the actions of individual users to infer their similarity, we prefer to call it similarity filtering. Traditionally, similarity-filtering algorithms determine which users have similar tastes via statistical correlations between their ratings of items. A table of users and items is initially filled with ratings that each user gives to each item, which may be explicit or implicit (a review or a purchase), precise or minimal (a ten-point scale or the ‘like’ button). It is not necessary for every user to rate every item, since empty cells can be filled with predictions based on correlations between users’ previous ratings. If there is a high correlation between users X and Y—both like the same jazz artists—they are considered to have similar tastes, and the algorithm predicts that items rated highly by X will probably be rated similarly by Y. More recent techniques incorporate not only ratings but a variety of user actions and characteristics, such as gender, age and geographical location. Although advanced techniques do not necessarily store ratings in a table format, they maintain the same basis.

The researchers at MIT and Xerox PARC in the early 1990s were primarily interested in finding efficient ways to discover songs that people might like, sort email messages, and distribute technical reports. However, their work soon attracted the interest of the advertising industry, which has played a decisive role in the development of similarity filtering since. Since the late 90s, this industry has provided most of the research funding and defined the authoritative performance benchmarks. The most

¹⁶ David Goldberg et al., ‘Using Collaborative Filtering to Weave an Information Tapestry’, *Communications of the ACM*, vol. 35, no. 12, December 1992; and Upendra Shardanand and Pattie Maes, ‘Social Information Filtering: Algorithms for Automating “Word of Mouth”’, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York 1995, pp. 210–7.

commonly cited empirical analysis of similarity-filtering methods is a 1998 partnership between Microsoft and Nielsen, a market research corporation. This study is based on datasets of viewing patterns of movies, television programmes and the Microsoft corporate website.¹⁷

A few years later, Amazon developed ‘item-to-item’ similarity filtering to recommend products to its customers. This was destined to become one of the most widely employed methods; the 2003 paper by Amazon researchers introducing the item-to-item method has been cited by more than four thousand publications.¹⁸ Then in 2006, the video-streaming and rental service Netflix announced a US\$1m prize for the similarity-filtering algorithm that would most accurately predict the movies individual members would like. The prize attracted prominent research groups in both industry and academia, and has been referred to in journals and conferences thousands of times. Amazon, Netflix and many other online retailers employ similarity filtering because it significantly improves sales and advertising performance, but many social media platforms also employ the logic of similarity to filter all of their communications, from ads to political speech.¹⁹

Paid sponsorship constitutes another filter, enabling sponsors to inflate the visibility of their items by paying the platform operator. Here, advertising is directly integrated into algorithmic filtering: an ad is an item like any other, but its visibility is manipulated, targeted and promoted. In the US, the Federal Trade Commission regulates paid sponsorship under the framework of consumer protection, requiring that sponsored items be identified as such ‘clearly and conspicuously’, though most platforms add only a subtle label, such as ‘sponsored’ or ‘promoted’.²⁰

Most of the revenues of Google, Facebook and Twitter come from such paid sponsorship. For the third quarter of 2015, Google reported revenues of \$18.7bn in its filings to the US Securities and Exchange

¹⁷ John Breese et al., *Empirical Analysis of Predictive Algorithms for Collaborative Filtering* (Technical Report MSR-TR-98-12), Microsoft Research, March 1998.

¹⁸ Greg Linden et al., ‘Amazon.com Recommendations: Item-to-Item Collaborative Filtering’, *IEEE Internet Computing*, vol. 7, no. 1, January–February 2003.

¹⁹ Google and Facebook own multiple patents for similarity-filtering techniques, for example ‘Collaborative filtering’ (US8135718) and ‘Compatibility scoring of users in a social network’ (US8775324).

²⁰ Federal Trade Commission, *.com Disclosures: How to Make Effective Disclosures in Digital Advertising*, Washington, DC 2013, p. ii.

Commission. About 90 per cent—\$16.8bn—came from advertising; 70 per cent from paid sponsorship on Google’s platforms, and 20 per cent from ads on the websites of Google’s partners. More than 95 per cent of Facebook’s reported revenues for the same quarter (\$4.3bn out of \$4.5bn) came from advertising, mainly through paid sponsorship in feeds. Twitter reported revenues of \$569m for the quarter, of which 90 per cent came from advertising.²¹ Although Twitter’s revenues are modest in comparison to Google’s and Facebook’s, they have quintupled since the first quarter of 2013. This increase follows Twitter’s progressive adoption of paid sponsorship. In addition to sponsored ‘tweets’ in the platform’s main feed (‘Timeline’), Twitter also sells placements in its list of ‘trending topics’: in the US, a 24-hour placement costs about \$200,000. Most Twitter users do not know that placements in the list have been purchasable since 2010, and this list is still often cited by journalists and researchers as an accurate measure of popularity.

Targeting users

Some platforms enable sponsors to target specific groups of users based on race, gender, age, location, income, net worth, employer, political affiliation, interests, consumer behaviour, and so on. The data used to generate such classifications derive from the surveillance techniques both of the platforms themselves and of third-party ‘data brokers’. For example, to estimate the net worth of Facebook users, Acxiom Corporation collects data on age, income, presence of children in the household, occupation, property, vehicles and investments. On Facebook, gender, age and employment are self-reported; location is either self-reported or tracked by IP address or GPS coordinates. Interests too can be self-reported, or guessed from other interests. Facebook also guesses the race of its users from statistical correlations between their actions, dividing American users, for example, into African-American, Asian-American and Hispanic—the latter with subcategories for bilingual, English-dominant and Spanish-dominant. Though the resulting classifications are invisible to normal users, they appear as ‘ethnic affinity’ on the control panel for advertisers.²²

²¹ See Google, Facebook and Twitter FY15-Q3 Form 10-Q filings for the period ending 30 September 2015.

²² Interestingly, there is no category for whites: only people of colour are algorithmically categorized and targeted by race. Modern technologies of surveillance and filtering work on the assumption that people of colour are subjects with

Some platforms infer the strength of social ties. Facebook deduces what is sometimes called the ‘affinity’ between each pair of users. Such calculations can combine many factors, including the frequency of interactions, whether the users are ‘friends’, and their similarity. If two users have a strong tie, their posts are much more likely to appear in each other’s feeds, whereas users with weak or absent ties—acquaintances and strangers—rarely see each other’s posts.²³ In 2012, Facebook researchers conducted large-scale field experiments and concluded that a ‘social cue’ alongside an ad ‘causes significant increases in ad performance’.²⁴ This cue might be a line of text in Tom’s friend’s feed saying ‘Tom likes Coca-Cola’, alongside a sponsored post promoting Coca-Cola. The effect is more pronounced if the user mentioned in the cue has a strong tie with the ad viewer. The logic of social ties is fundamental to maximizing the profits of Facebook and other platforms whose business models rely on advertising. One consequence of such calculations is to segregate the communication of users deemed to have weak or absent ties—as we will show in the Brazilian case, indigenous from non-indigenous.

Users do have a degree of agency, especially through subscription. On most platforms they can select specific keywords, tags, other users and groups to follow; on Facebook, Google Plus and Twitter, users may follow one another. There are also ways to block or restrict information from specific users and groups. In July 2015, Facebook introduced a feature that allows users to select which sources to follow more closely (‘see first’), though users have only two main options: ‘top stories’ (the default setting) and ‘most recent’. The latter sorts the feed in chronological order, but still employs undisclosed logics to select which posts to include and which to exclude. On Google Plus and Twitter there are no options at all. This kind of agency is thus clearly constrained, and users have almost no control over basic filtering logics.

homogeneous and categorizable behaviours—‘ethnic affinities’—while white people are not. Concerns have been raised recently that Facebook’s ethnic affinity classifications may amount to racial profiling. See Alex Hern, ‘Facebook’s “ethnic affinity” advertising sparks concerns of racial profiling’, *Guardian*, 22 March 2016.

²³ For further discussion, see Motahhare Eslami et al., “‘I Always Assumed That I Wasn’t Really That Close to [Her]”: Reasoning About Invisible Algorithms in News Feeds’, in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, New York 2015.

²⁴ Eytan Bakshy et al., ‘Social Influence in Social Advertising: Evidence from Field Experiments’, in *Proceedings of the 13th ACM Conference on Electronic Commerce*, New York 2012.

Some researchers question the extent to which algorithmic filtering reduces the diversity of viewpoints available to users; others attribute such effects to different factors. The legal scholar Cass Sunstein suggests that the ability of consumers to filter their own information by choice leads to ideological ‘echo chambers’ where they mostly communicate with like-minded people and rarely encounter alternative viewpoints.²⁵ But this argument is misleading: while individual preferences may contribute to the problem, their automated extrapolation by algorithmic filters is far more important. Our analysis shows that the root cause of reduced diversity is not the ability of users to control their information environments as Sunstein contends, but in large measure their inability to do so. Activist Eli Pariser uses the term ‘filter bubble’ when discussing the narrowing of perspectives on social media, which he attributes to the ‘personalization algorithms’ imposed by platforms. Although he does not provide a succinct definition, for Pariser, ‘personalization’ appears to mean that platforms display different information to different users according to algorithms over which those users have no control. But he assigns too much importance to personalization as the key factor giving rise to ‘a narrow, overfiltered world’, in contrast with ‘the old, unpersonalized web’, which ‘offered an environment of unparalleled richness and diversity’.²⁶

The absence of personalization is neither necessary nor sufficient to limit the suppression of diverse viewpoints. Television and newspapers are both non-personalized media, yet they often dramatically limit the range of views available. Moreover, it is quite possible to imagine personalized filters that would allow for a greater diversity of views, as we will discuss later. The cause of ‘filter bubbles’ on social media is not personalization itself, but rather a combination of specific filtering logics that have become predominant—especially those of similarity and social ties, which reduce diversity by design.²⁷ Since all data processed by computers is inevitably placed in a certain order, with a determinate

²⁵ Sunstein, *Republic.com* 2.0.

²⁶ Pariser, *Filter Bubble*, pp. 66, 102.

²⁷ A controversial 2015 article by three Facebook employees purports to demonstrate that individual choices play a greater part in limiting exposure to ‘attitude-challenging content’ than the company’s algorithms: Bakshy et al., ‘Exposure to Ideologically Diverse News and Opinion on Facebook’. The authors quantify the extent to which 10.1 million users who self-identified their political orientation encountered politically diverse content, classifying users and posts as either liberal or conservative and measuring the amount of ‘cross-cutting content’ shared by users’ friends, displayed in feeds and clicked. But this self-identifying sample is not representative

structure, there can be no platform without ‘filtering’ of some kind. The key question is how particular logics work to promote and exclude different viewpoints in line with commercial priorities, and whether alternative logics are conceivable that would not have the same effects.

Filtering social movements

How do these logics affect the visibility of social movements and activists? We will now address this question directly in a case study of the land disputes between agribusiness and approximately forty-five thousand Guarani and Kaiowá people who live mainly in Mato Grosso do Sul, a western Brazilian state located near Bolivia and Paraguay.²⁸ All sides use Facebook to mobilize. For centuries, Mato Grosso do Sul has been the focus of intense and violent disputes between native peoples and various agents of colonization. At present, conflicts rage throughout the state not just between agribusiness and the Guarani and Kaiowá peoples, but also the Terena and other groups such as *quilombolas* and *ribeirinhos*. Broadly, these may all be termed movements for decolonization, with struggles for the right to land and for the deployment of intercultural knowledges in education, health, law, media and government. Most of these people have been forced into state reservations as a result of aggressive settler colonialism and land enclosures.²⁹

of all users. Moreover, the authors’ own data show unequivocally that the greatest ideological narrowing on Facebook occurs because of the company’s policy of only displaying items that have already been shared by the user’s ‘friends’. The authors do not recognize this policy of selection by social ties as a type of algorithmic ranking, allowing them to exclude it from consideration. And in spite of the controversy, Facebook has not allowed other researchers to reproduce the experiment.

²⁸ Our analysis of social media is principally based on field research conducted by Rodrigo Ochigame in Mato Grosso do Sul in May–June 2014. For a *longue durée* view of land conflict, land law, and labour regimes in Brazil, see James Holston, *Insurgent Citizenship: Disjunctions of Democracy and Modernity in Brazil*, Princeton 2008, especially pp. 112–45. For more comprehensive studies of Guarani and Kaiowá movements for decolonization and land recovery, see the work of Kaiowá anthropologist and activist Tónico Benites, especially his 2014 doctoral dissertation, ‘Rojeroky hina ha roike jevy tekohape (Rezando e lutando): o movimento histórico do Aty Guasu dos Ava Kaiowa e dos Ava Guarani pela recuperação dos seus tekoha’, Universidade Federal do Rio de Janeiro.

²⁹ Guarani and Kaiowá people mostly use the Guarani term *tekoha* instead of the Portuguese term *terra* to refer to their native lands. Tónico Benites defines *tekoha* as the physical space where it is possible to practice the Guarani and Kaiowá ‘way of being and living’ (*teko*).

In the colonial period, Portuguese and Spanish settlers enslaved indigenous people in the region under the Asunción-based *encomienda* regime, and tried to catechize them. After the War of the Triple Alliance (1864–70), the Brazilian state conceded Guarani and Kaiowá lands to Companhia Matte Larangeira, a *yerba mate* extraction business based on indigenous labour. In the twentieth century, the settler state intensified the expulsion of Guarani and Kaiowá peoples from their native lands, confined them to eight reservations totalling a mere eighteen thousand hectares, sold their lands for agriculture, and promoted deforestation and monoculture. The demarcations of the reservations, established by the Indian Protection Service from 1915 to 1928, have not changed since then. Today, Guarani and Kaiowá communities face alarming rates of suicide and malnutrition.³⁰ Meanwhile, the usurped lands are used for agribusiness, which is rapidly attracting foreign capital and expanding its area. In 2013, Mato Grosso do Sul exported US\$4.76bn in agricultural products—24.8 per cent more than the previous year—especially soy, beef and cellulose.³¹

Since 1979, Guarani and Kaiowá communities have organized themselves into democratic forums called *Aty Guasu* ('the great assembly' in Guarani) to discuss issues of common concern. Participants identify mutual needs, share tactics, make collective decisions, and prepare documents for public dissemination. The recovery of land by reoccupation has been a major concern. If a community decides to proceed with reoccupation, they will march towards their land, set up encampments, and pursue legal negotiations. Article 231 of Brazil's 1988 constitution declares that indigenous peoples have a right to become exclusive usufruct holders of 'traditionally occupied lands' once the state has recognized them as such, but inefficiency and corruption plague both executive and judicial aspects of this process, and the lands of many Guarani and Kaiowá communities remain unrecognized.³² Although the reoccupations are generally nonviolent, many non-indigenous farmers respond by hiring private militias—

³⁰ From 2000 to 2013, there were at least 659 suicides of Guarani and Kaiowá people in Mato Grosso do Sul; in 2013 alone, there were at least 72 cases: Conselho Indigenista Missionário, *Violência contra os Povos Indígenas no Brasil: Dados de 2013*, Brasília 2014, pp. 79–85.

³¹ Statistics from Agrostat, a database maintained by the Brazilian Ministry of Agriculture, Livestock and Food Supply.

³² Furthermore, the pro-agribusiness bloc in Brazil's national congress has proposed many constitutional amendments to stall the process, most notably 215/2000, which would transfer the duty of recognizing traditionally occupied lands from the executive to the legislative branch.

known locally as *pistoleiros* or ‘hired guns’—to attack the encampments. A militia named Gaspem, which was registered as a private security firm owned and operated by a former military police officer, was—according to federal prosecutors—responsible for at least two murders and eight attacks against indigenous communities from 2005 to 2013. In 2013, there were 31 murders of Guarani and Kaiowá people in Mato Grosso do Sul, most of which were not investigated sufficiently to determine cause. Among the few that were, three were attributed to land disputes.³³ But there is considerable overlap between land and media ownership, so these murders rarely appear on television or in newspapers. One family alone controls the largest media conglomerate in the region—a Rede Globo affiliate with over half of the local television audience—in addition to cattle farms and soybean oil mills. While mainstream media often interview representatives from agribusiness organizations, they hardly ever quote indigenous activists.

If voices from agribusiness dominate old media, are new media any different? Some Aty Guasu participants have organized a team to publish reports on social media, hoping to reach a wider audience. Despite the lack of internet access on reservations, the Aty Guasu team has managed to publish urgent reports on Facebook—their primary platform—and their own now defunct blog, often within a day. How does Facebook’s algorithmic filtering determine the visibility of these reports, and to what effect? We can begin by asking who sees them. Perceptions of online visibility differ, because feeds are personalized and private. However, when we compute visibility statistics for the Facebook pages of the various actors involved in land disputes—indigenous activists, Catholic missionaries, agribusiness organizations, NGOs, land reform movements and local newspapers—the results are striking (Table 2, overleaf). The most representative voice of Guarani and Kaiowá activists, Aty Guasu, had by far the least visible page of all the actors we examined. Even among those in favour of land recovery, Aty Guasu had much less visibility than national and international NGOs such as Survival International.³⁴ We

³³ Conselho Indigenista Missionário, *Violência*, pp. 49–55.

³⁴ Ochigame conducted informal queries among people in Campo Grande, the state capital, to find out whether they knew about the indigenous land movements: those questioned included strangers in the main public square, family friends, acquaintances, and non-indigenous activists. Some said they had seen stories, particularly about ‘indigenous invasions of farms’, in the print and broadcast media; others had witnessed indigenous street protests themselves. But only the non-indigenous activists had seen posts by indigenous activists on social media.

TABLE 2: *Visibility statistics of Facebook pages of actors in land disputes*

	Type of Actor	Page Likes	Posts	Likes per Post	Shares per Post
Aty Guasu	Guarani and Kaiowá assemblies	5,154	320	0.9	1.1
Conselho Indigenista Missionário (CIMI)	Catholic missionary organization	20,278	4,858	19.6	24.5
Federação de Agricultura e Pecuária de MS (Famasul)	Agribusiness organization	49,194	6,436	9.3	3.9
Instituto Socioambiental (ISA)	National NGO (society and environment)	69,128	2,004	113.4	116.4
Survival International	International NGO (indigenous rights)	147,957	1,605	229.4	156.3
Movimento dos Trabalhadores Sem Terra (MST)	Social movement (land reform)	178,213	7,517	176.1	123.6
Correio do Estado	Local newspaper	343,352	18,053	141.9	35.5

Source: Facebook Graph API, 19 April 2015. Aty Guasu, Famasul and *Correio do Estado* are based in and focused on Mato Grosso do Sul. The other organizations are national or international in scope.

will now examine in detail how the filtering logics of paid sponsorship, popularity, social ties, and similarity affect the visibility and voice of the various actors in the land disputes.

1. *Paid sponsorship and privileged visibility.* The Federation of Agriculture and Livestock of Mato Grosso do Sul (Famasul), a private agribusiness organization, has routinely paid Facebook to inflate the visibility of its posts, many of which concern land disputes with indigenous communities. When Famasul published multiple posts with almost identical content, the sponsored ones were displayed much more often. One sponsored post, describing indigenous people as invaders, had more ‘likes’ and comments than the organization’s next dozen non-sponsored posts combined. There are also informal schemes of paid sponsorship. In April 2013, for example, Famasul announced a raffle for a tablet computer: in order to participate, people had to ‘like’ the organization’s Facebook page, causing it to become more visible. By contrast, indigenous activists lack the funds to pay for visibility, either

formally or informally. The logic of paid sponsorship thus privileges well-funded organizations.

2. *Popularity and the suppression of dissent.* Some algorithmic definitions of popularity tend to exclude important kinds of political speech. For example, many posts on Facebook by Aty Guasu include videos of shootings by farmer-controlled militias and photos of the funerals of murdered indigenous activists. The popularity of a post is largely defined by its count of ‘likes’, as registered by clicking either the word or the ‘thumbs-up’ icon next to it. In Facebook’s Brazilian Portuguese version, the word ‘like’ is translated as *curtir*, which is closer to the English word ‘enjoy’. This linguistic choice undermines posts that express dissent and denounce violence: how many people will ‘enjoy’ scenes of violence? Unsurprisingly then, most of these Aty Guasu posts had zero ‘likes’. Despite their importance and urgency, the filtering algorithm considers them unpopular and reduces their visibility; they are much less likely to be high in the feed, and may well be excluded altogether.

In contrast to Aty Guasu, agribusiness organizations tend to publish more ‘likable’ content, such as market growth statistics and photos of smiling children with crops. Famasul uses a friendly, inviting, seemingly generous tone even in posts about land disputes. In June 2013, they published an image juxtaposing a white farmer and an indigenous person to represent a dichotomy between productive and unproductive land use, with the following text: ‘Where there is justice, there is space for everybody. It’s time to take action for productivity and subsistence without conflicts.’ This post received over one hundred ‘likes’, more than any post ever published by Aty Guasu. The latter published a response, denouncing Famasul for its concealment of genocide and racism and rephrasing the text to say, ‘Where there is justice, there is demarcation of indigenous lands . . . there is no genocide’. This response, written in a critical and alarming tone, had zero ‘likes’. In this case, the logic of popularity clearly suppresses dissent in favour of likable and enjoyable content.

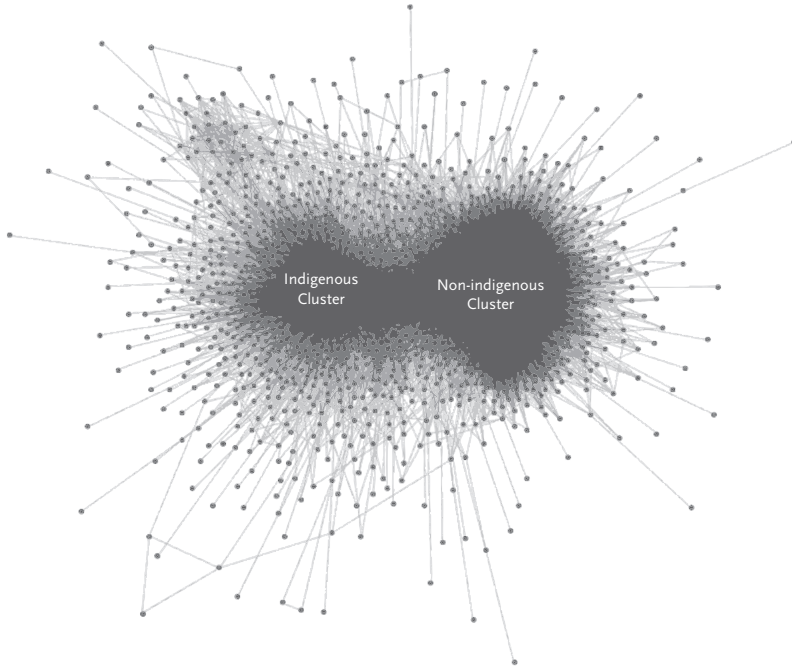
3. *Social ties and segregation.* Although political outreach and coalition-building often require communication between people who do not know each other, some algorithms limit communication between users with weak or absent ties. We found that non-indigenous people rarely see posts by indigenous activists on Facebook. This phenomenon is partly related

to the logic of social ties. In Mato Grosso do Sul, most non-indigenous people have few ties to indigenous people—still less to indigenous activists. Figure 1 shows the friendship network of 2,048 Facebook users connected to an activist group—Coletivo Terra Vermelha (Red Earth Collective)—that provides support to indigenous movements. While all these users are ‘friends’ of the Coletivo, they are not necessarily friends among themselves. Rather, they are clearly separated into distinct clusters based on indigeneity. In the left cluster most users are self-identified Guarani, Kaiowá and Terena people, living on indigenous lands and reservations or in cities like Dourados, Campo Grande and Rio de Janeiro. In the right cluster, most are non-indigenous people living in Brazilian cities, especially Campo Grande. Since Facebook uses the logic of social ties to filter posts, there is little circulation of posts between these two clusters. The composition of clusters is self-selected because individuals choose their ‘friends’, but the lack of communication between clusters is the product of Facebook’s algorithm. They become segregated from each other, reducing exposure to divergent perspectives and limiting opportunities for political coalition-building.

4. *Similarity and the suppression of divergent perspectives.* Similarity filtering also reduces the exposure of users to political perspectives that differ from their own, since Facebook users tend to ‘like’ posts that reflect their views, and algorithms select items deemed similar to those that users have already liked. Our fieldwork found that while some people in Mato Grosso do Sul saw many news and opinion articles about the indigenous mobilizations on their Facebook feeds, the articles in each particular feed all expressed similar political arguments. Activists who were in favour of land recovery constantly saw articles expressing that viewpoint, while everyone else repeatedly saw posts from media conglomerates and agribusiness organizations. The only people who regularly saw articles that conflicted with their own viewpoint were activists, who received sponsored posts from Famasul: perhaps because sponsored posts on Facebook always have a target audience—in this case, those with an interest in the land disputes—or simply because the activists were more likely to notice a sponsored post by Famasul.

When articles by indigenous activists were shown to those who had only seen pro-agribusiness posts, many expressed a willingness to consider alternative views. The political-communication scholar R. Kelly Garrett has drawn a distinction between desire for opinion reinforcement and

FIGURE 1: Facebook friendship network of Coletivo Terra Vermelha



Source: Facebook API; visualization produced using Wolfram Alpha tool.

aversion to opinion challenge, basing his findings on a US national survey, a behaviour-tracking study, and literature reviews on selective exposure.³⁵ According to Garrett, while most people tend to seek opinion-reinforcing information about politics, they do not systematically avoid opinion-challenging information. He argues that 'many of the individuals who choose an opinion-reinforcing partisan source would have preferred a source representing multiple opinions if one was available'.³⁶ The suppression of divergent viewpoints on social media seems to go beyond simple individual preferences, because algorithms generalize

³⁵ Garrett, 'Politically Motivated Reinforcement Seeking: Reframing the Selective Exposure Debate', *Journal of Communication*, vol. 59, no. 4, 1 Dec 2009; and 'Echo Chambers Online?: Politically Motivated Selective Exposure among Internet News Users', *Journal of Computer-Mediated Communication*, vol. 14, no. 2, 1 Jan 2009.

³⁶ Garrett, 'Politically Motivated Reinforcement Seeking', p. 680.

those preferences to new situations about which individuals have yet to express an opinion. The software simply assumes that users always want to see items similar to those they have already liked; thus the logic of similarity excludes opinion-challenging information by design.

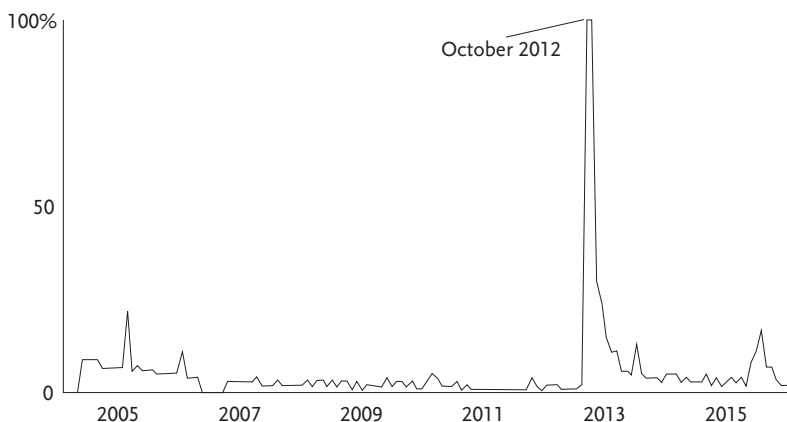
Circumvention and appropriation

Though most platforms do not allow users to manage algorithmic filtering, the sovereignty of software design is far from absolute, and there are opportunities for contestation. Users can transgress imposed limitations by making the technology work in unintended ways, thereby becoming more than mere ‘users’. In Mato Grosso do Sul, activists do so in many ways. According to an indigenous activist and filmmaker who makes videos to document violence and resistance, and who requested anonymity: ‘Indians are gradually dominating the technology of non-Indians.’ Guarani and Kaiowá activists not only use but also appropriate new media creatively, strengthening their voices.

In October 2012, Aty Guasu experienced an unusual moment of visibility on social media. When a federal court ruled to evict a Guarani and Kaiowá community from their land, Aty Guasu published an open letter denouncing the judiciary for its violence.³⁷ Over the following weeks, the Guarani and Kaiowá struggle had a surge of visibility on Facebook. How did this happen, given that, as we have seen, standard posts by Aty Guasu do not appear in many feeds? Rather than circulating by means of conventional posts, the letter and other information about the land disputes spread as sympathetic users and activists deployed a subversive tactic: they changed their last names in their Facebook profiles to ‘Guarani-Kaiowá’. Last names are not subject to algorithmic filtering but are presented automatically, appearing in many places—at the top of every profile and next to every post, comment, private message and notification. Thus users would see these names independently of their feeds. When they did, they became intrigued, because Brazilians broadly recognize Guarani as indigenous but do not normally see Guarani-Kaiowá as a last name. These users contacted the person with this last name and had an exchange about the land disputes, or searched the web for ‘Guarani-Kaiowá’ and thereby learned about the situation. In one week,

³⁷ The letter, titled ‘Carta da comunidade Guarani-Kaiowá de Pyelito Kue/Mbarakay-Iguatemi-MS para o Governo e Justiça do Brasil’, is widely available online.

FIGURE 2: Google search interest for ‘guarani’ and ‘kaiowá’



Source: Google Trends. ‘Interest’ is measured relative to the peak number of searches in Oct 2012.

the volume of Google searches containing the keywords ‘guarani’ and ‘kaiowá’ increased dramatically (Figure 2).

As visibility peaked, the Guarani and Kaiowá struggle gathered much more public support. Activists organized street mobilizations in over fifty cities in Brazil and elsewhere, including New York, Hamburg and Lisbon. Some demonstrators carried banners saying, ‘We are all Guarani-Kaiowá’. The *New York Times* and BBC reported the phenomenon. The discussion also ignited racist reactions in the mainstream press. One columnist wrote in *Folha de S. Paulo*—Brazil’s most widely circulated newspaper—that ‘any defence of a Neolithic mode of life on Facebook is a certificate of mental indigence’.³⁸

This was, however, the only time that the Guarani and Kaiowá struggle had such visibility online, and its circumvention of Facebook’s algorithmic filtering was ephemeral. Not long after the tactic became prominent, Facebook banned the use of ‘Guarani-Kaiowá’ as a last name—even for its Guarani and Kaiowá users—in order to enforce its ‘real names policy’. Activism plummeted. Members of Coletivo Terra Vermelha reported

³⁸ Luiz Felipe Pondé, ‘Guarani Kaiowá de boutique’, *Folha de São Paulo*, 19 November 2012.

that their meetings quickly dropped from around eighty attendees in November 2012 to only three. During the surge of visibility, Guarani and Kaiowá activists lost control of the agenda because sensationalist headlines flooded the discussion. Many articles claimed that the open letter was declaring a ‘collective suicide’, and Aty Guasu had to deny this misinterpretation repeatedly. Many other communities involved with land-recovery struggles, such as Terena, *quilombolas* and *ribeirinhos*, have never experienced a comparable moment of visibility. Despite the occasional circumvention of software systems, most activists remain unheard as long as their voices are algorithmically filtered out by default.³⁹

Politics of audibility

Conceptually, one might consider algorithmic filtering a kind of ‘implicit censorship’ in Judith Butler’s sense.⁴⁰ It is certainly not explicit censorship as conventionally understood in terms of exceptional, state-directed acts of restriction against persons and the content of their speech. As we have shown, such filtering is not only inconspicuous—even invisible—to users, it is also continuous, automatic and generalized. These features constitute a conceptual shift in the regulation of information. In contrast to internet filtering, which is readily recognizable as a form of explicit censorship, algorithmic filtering operates on platforms even in the absence of state regulation because it is part of the software architecture itself.⁴¹ Every feed must establish its selection and order of items by

³⁹ Creative tactics can also serve to weaken opponents’ voices. Christopher Peterson has documented cases in which users have tactically manipulated filtering algorithms to suppress political speech. Such tactics include using multiple accounts to vote posts down and flagging posts as ‘hate speech’ and ‘spam’. See his 2013 master’s thesis, ‘User-Generated Censorship: Manipulating the Maps of Social Media’, Massachusetts Institute of Technology.

⁴⁰ As Butler notes (in *Excitable Speech: A Politics of the Performative*, New York 1997, p. 128), ‘in the conventional view, censorship appears to follow the utterance of offensive speech: speech has already become offensive, and then some recourse to a regulatory agency is made’. However, she inverts this temporal relation to argue that censorship also produces speech. She does so to consider that, as a form of power, censorship operates both explicitly and implicitly. Implicit operations of censorship ‘rule out in unspoken ways what will remain unspeakable’ and may be more ‘efficacious than explicit forms enforcing a limit on speakability’ (p. 130).

⁴¹ When an intermediary explicitly blocks a website or a keyword, it may reveal what it does not want people to see. The censor may thereby unwittingly introduce the censored website or keyword into public discourse. In China, for example, activists publish and circulate lists of websites and keywords blocked by the state.

some criterion, and these favour some voices over others. In this sense, algorithmic filtering is inevitable, and the idea of a free flow of information a fantasy. Thus, the continuous operations of algorithmic filtering to organize feed visibility require a new understanding of information control as rule rather than exception.

It is important to make an additional distinction: while the concept of censorship focuses on speakability, algorithmic filtering does not promote or suppress the speech act itself but rather its audibility, its capacity to reach an audience. On most if not all major social media platforms, getting an account is essentially unrestricted to those with an internet connection, and anyone with an account can speak. The problem is that algorithmic filtering imposes restrictions on what one can hear, on what audience one can reach, and on how speech circulates. In the current regime of online information control, these restrictions are overdetermined by the political economy of advertising. If the crucial issue is not whether one is free to speak but how speech circulates, the contestation of the current regime of information control requires a politics of audibility more than of free speech. The latter is necessary but insufficient to confront the automated and implicit regulation of what voices get heard. The politics of audibility must reveal and question the norms that structure algorithmic filters. As filtering is inevitable, it is necessary to move beyond impossible demands for neutrality and impartiality, and instead to call for strategies that subvert or repurpose existing filters, and for alternative algorithmic logics inspired by different political norms.

Since it is a safe bet that commercial platforms will not voluntarily change their filtering systems in ways that conflict with profits, we propose three plausible strategies to realize alternative norms. The first is the legal regulation of existing platforms that would require them to make public their filtering algorithms and prevent secret feed manipulation. This publicity would not change the promotion and suppression of information on these sites, but at least users would be informed about the filters. The second is the subversion of existing platforms. As we saw in the Guarani and Kaiowá case, this entails the hacking of specific platforms through the circumvention and appropriation of their filtering technologies. There are many imaginative, platform-specific hacks other than those we have documented here, such as using browser extensions that rearrange and reconfigure feeds. But such tactics are mostly 'hit-and-run', and only temporarily effective. The third strategy we propose calls for

the creation of new platforms organized according to more democratic criteria. These would be transparent about, and accountable for, their filtering algorithms. They would publish and explain the source code of their filters, ideally as free and open-source software. These filters could even be auditable by allowing users to use cryptographic methods to verify that their feeds have not been manipulated.⁴²

A new generation of social media platforms could be developed by non-profit organizations or independent collectives which would have less vested interest in proprietary code. More democratically organized platforms would let users have autonomy over the processes that filter the information they see. It is perfectly feasible to design new kinds of user interfaces that enable greater control of filtering logics. Platforms could update their code based on user deliberation, and even let users configure feeds with their own algorithms, thereby encouraging participatory experimentation with new ways of discovering and circulating information. Users could have the ability to access diverse information if they wished, which is currently not the case on any major platform.⁴³ If new communication technologies hold promise for improving the scale and quality of democratic deliberation, both a reconceptualization and a reconfiguration of algorithmic filtering are necessary. Alternatives are needed to the current foundation of social media platforms in advertising for profit. We do not wish to prescribe a new norm here, as the possibilities are multiple and nuanced, but we support those that enable more direct democratic engagement, according to principles of transparency, accountability, diversity, participation and autonomy, all of which challenge the current political economy of information control.

⁴² For an analysis of proposed systems that combine cryptography and public policy to increase accountability in algorithmic processes, see Joshua Kroll's 2015 doctoral dissertation 'Accountable Algorithms', Princeton University.

⁴³ For a discussion of technical strategies to promote exposure to greater diversity, see R. Kelly Garrett and Paul Resnick, 'Resisting Political Fragmentation on the Internet', *Daedalus*, vol. 140, no. 4, 29 September 2011. While translating principles into technical designs is not a straightforward process, the appropriation of existing filtering techniques may provide key ingredients for new kinds of filtering. For example, similarity filtering generates maps of opinion that could be used differently, navigated in a way that includes divergent viewpoints.