WORLD WIDE WEB
FOUNDATION

THE INVISIBLE CURATION OF CONTENT:

Facebook's News Feed and our
Information Diets

April 2018

CONTENTS

WORLD WIDE WEB FOUNDATION

FOREWORD

In March 2017, our founder and inventor of the World Wide Web, Sir Tim Berners-Lee, expressed concerns regarding the loss of control over personal data, and the role that algorithms and artificial intelligence are increasingly playing in society. Over the last year, the Web Foundation has embarked on research to provide frameworks and examples to think about how the use of personal data, algorithms and AI are creating risks and opportunities for low- and middle-income countries.[1]

One of the concerns that has emerged from this research is the degree to which opaque algorithms are managing information that is critical to our societies, shaping public debate and defining access to public services.

Given the Web Foundation's mission to bridge the digital divide, we are worried about the imbalance in power dynamics between those who build algorithms, and those who are impacted by them. We see this study as an effort to put the spotlight on algorithms affecting our societies and people's experiences online.

"Algorithms have been delegated a critical curatorial function that defines online users' information diets and what can best be called the online 'public square."

This report focuses on one of the algorithms that online users often interact with: how social media platforms curate and deliver information. To better understand this process, we ran a controlled experiment through six identical Facebook profiles that followed the same news sources on Facebook. The findings demonstrate how algorithms have been delegated a critical curatorial function that defines online users' information diets and what can best be called the online 'public square'. While the focus is on Facebook, there are equal concerns that other social media platforms, including YouTube and Twitter, are shaping and framing the information diet of users — without their knowledge or control — due to a reliance on similar algorithms.

We must work to reverse this trend and to ensure people who use these platforms are put back in control and that their online rights are protected and respected. We envision this report as a first step toward these ends, allowing the general public to better understand how the platform algorithms work and the mechanics behind the process. In doing so, we hope this document helps to trigger a broader conversation regarding the role of online users, governments and platforms in defining the values that inform how algorithms are designed, leveraged, and made accountable to the people and communities they affect.

---

1   WF published three white papers on these topics. For more information, see: https://webfoundation.org/research/white-paper-series-opportunities-and-risks-in-emerging-technologies/.

EXECUTIVE SUMMARY

Facebook — the world's largest online social media platform with 2 billion active monthly users — has developed a broad range of technologies which inherently shape its relationships with users.

This includes defining the information Facebook users will be exposed to (or not) via its platform. This study builds on the findings of past research efforts to show how each Facebook user has a bespoke 'information diet.'[2] The experiments included in this report have been done to inspire a conversation about what this information diet means, and how it is structured by Facebook. Despite recent changes announced by Facebook,[3] its News Feed feature still only allows limited customisation features that users can control.[4]

The findings from the experiments included in this report show substantial variations in the information users receive via the platform — even when users have the exact same profile characteristics. The research shows:

- **Large gaps between what stories were published and seen as posts:** Our profiles were shown an average of one out of every six posts from across all the pages that they were similarly following.[5] The broadness of this gap between what is seen and hidden serves as a proxy for the potential level of curation taking place on the platform.[6]

---

2   Kincaid, Jason. 2010. "EdgeRank: The Secret Sauce That Makes Facebook's News Feed Tick." http://techcrunch.com/2010/04/22/facebook-edgerank/; Taylor, Dan. 2011. "Everything You Need to Know about Facebook's EdgeRank." http://thenextweb.com/socialmedia/2011/05/09/everything-you-need-to-know-about-facebooks-edgerank/; Napoli, P. (2014). Automated Media: An Institutional Theory Perspective on Algorithmic Media Production and Consumption. *Communication Theory, 24*(3), 340-360.; Devito, M. (2017). From Editors to Algorithms: A values-based approach to understanding story selection in the Facebook news feed. *Digital Journalism, 5*(6), 753-773.; Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin.

3   Facebook is rolling out a more transparent advertising system, which has already begun as a pilot in Canada. See: https://www.propublica.org/article/facebook-experiment-ad-transparency-toronto-canada.

4   The option to opt out of the curated News Feed ("Top Stories") gets reset every time the user exits the application. The option to prioritise certain Pages basically allows users to place the content from a specific person or Page "first" on the News Feed. It is unclear what happens when several Pages or profiles are marked to appear first on the News Feed, or how the News Feed is working as a default, or for posts published by actors not marked as a priority.

5   This figure is equivalent to 18% of all the posts published during the period but varies between Pages. The average share of posts seen pre-polarisation was 14%. See https://public.tableau.com/profile/web.foundation#!/vizhome/Theinvisiblecurationofcontent-Full-English/SeenNotseenBars.

6   This percentage is sensitive to the relationship between the type and number of Pages followed (how much content is published-supply) on the one hand, and the amount of time spent scrolling down Facebook's timeline on the other (demand). Facebook's own estimates of how users experience this gap are similar. https://www.facebook.com/business/news/Organic-Reach-on-Facebook Our research provides a window for users to see how this takes place, and examples of what types of posts might be seen and excluded.

- **Entire lack of exposure to certain news stories:** The decision to rely on algorithms to curate content means critical social and political posts that fall outside of these metrics may not be shown to some users. For example, during the period observed none of the stories about femicide published by the pages our profiles were following surfaced on their feeds.

- **Different levels of exposure to different articles:** Even users with the same profile and who follow the same news sites may not be exposed to the same set of stories. It is as if two people buy the same newspaper but find that different sets of stories have been cut out. The algorithm places each user into a separate and individualised version of what should be an open public square for information.

These findings are based on tests conducted in Argentina between October and December 2017 using a series of Facebook profiles and automated tools to control and monitor their online behaviour.[7] The study presents a methodological approach that could help expand research on the dynamics of algorithms used by social media platforms. Academic studies and news articles that look at Facebook's role in news distribution have tended to focus on the European Union and US contexts. The country-level results for Argentina have the potential to be extrapolated to Facebook users in other countries given the standardised manner in which the platform delivers information across users and countries (see "methodology").

Recommendations

These findings suggest that a degree of curation takes place on the platform that appears disconnected from users' own profiles and interests. The Web Foundation believes that users should have greater control over their News Feed[8] and suggests these measures to help Facebook address the gaps found by the study:

- **Explore the use of real-time transparency dashboards** for users to compare what posts were seen versus what was published over a given period of time, in order to increase public understanding of how the algorithm works.

- **Make chronological ordering the default mode** for content delivery via Facebook.[9]

- **Allow users an array of options to manage their own News Feed settings** and play a more active role in constructing their own information diet.[10]

- **Be responsive to Facebook users** by working with them to actively shape platform decisions and feedback on their user experience as a result of the algorithms being used.

- **Facilitate independent research** on the externalities and systemic effects of the algorithmic curation of content.

The current model of top-down control over 'information diets' is in need of revision. Facebook is already looking into restructuring its News Feed, which could provide an important opening for such changes as it attempts to offer users the "most personalised experience."[11] However, the announcement has not offered details around how users can exert greater control over their News Feed content, or opt out from the changes.

We hope that this report encourages Facebook to redefine its role and to empower users of its platform to determine their own information diets.

---

7 Two experiments were run - the first with nine user profiles and the other with six. The second test adjusted for challenges that were found in the first experiment in order to minimise the variables being assessed.

8 Global surveys show that almost 1 in 2 people do not feel comfortable providing information to online companies in return for personalised services and products. Also nearly 60% of people globally mind companies using information provided to them automatically. See: https://www.ipsosglobaltrends.com/wp-content/uploads/2017/04/Slide02-5.jpg and https://www.ipsosglobaltrends.com/wp-content/uploads/2017/04/Slide03-5.jpg.

9 Current settings allow users to opt-out from algorithmic curation. Yet, in the context of informal testing, we observed that re-launching the application reset this preference to the default of algorithmic curation. See https://www.facebook.com/help/community/question/?id=10152476808951132.

10 See for example gobo.social.

11 "We don't think of it in the terms of 'neutral,'" Stretch continued, pointing out that Facebook tries to give users a personalized feed of content. "But we do think of ourselves as (...) open to all ideas without regard to viewpoint or ideology." Transcript of Facebook VP Stretch in Senate hearing on Russian 2016 elections interferences https://www.recode.net/2017/10/31/16579072/live-update-facebook-google-twitter-testimony-russia-congress-hearing.

01

FACEBOOK'S 'INFORMATION DIET': THE ARGENTINA EXPERIMENT

With over two billion monthly active users around the world, Facebook is indisputably the leading social platform globally.

This footprint means that Facebook has a trove of users' personal data, which it uses to refine user profiles and customise content delivery through proprietary algorithms. These algorithms are used to determine what information is delivered and when — from advertising the latest online sale to alerting us to friends' posts about their dogs.[12]

Facebook provides users a curated experience, sifting through the mountains of information circulating within its platform. This curation is designed to increase user engagement.[13]

When it comes to news and information delivery, such curation has profound impacts. It is like buying a newspaper or magazine on the street and having the vendor cut out the articles that he or she decided would not interest you. These personalised information diets, that users are unable to control, risk exacerbating divisions between and among communities that increasingly lack a shared body of news and therefore a shared understanding of reality.

Online users' reading experience is radically changing as a result of algorithms personalising what information is delivered. As users, we are only aware of the content the algorithm produces for our News Feeds: the articles left inside a clipped out magazine. Yet users are still largely unaware of what is not appearing in their News Feed and why.[14] While Facebook enables users to opt out of the News Feed, or prioritise content of certain Pages, these options seem limited and not user friendly as they fail to provide enough information for users to make informed choices about their information diets.[15]

12 Facebook estimates that "Of the 1,500+ stories a person might see whenever they log onto Facebook, News Feed displays approximately 300. To choose which stories to show, News Feed ranks each possible story (from more to less important) by looking at thousands of factors relative to each person." See: https://www.facebook.com/business/news/Organic-Reach-on-Facebook.

13 For related research on algorithms, see: http://www.pewinternet.org/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/.

14 Powers, E. (2017). My News Feed is Filtered?: Awareness of news personalization among college students. Digital Journalism, 5(10), 1315-1335.

15 The option to opt out of the curated News Feed ("Top Stories") gets reset every time the user exits the application. The option to prioritise certain Pages basically allows users to place the content from a specific person or Page "first" on the News Feed. There is no indication regarding how the algorithm will act when several Pages or profiles are marked to appear first on the News Feed, or how the News Feed is working as a default, or for posts published by actors not marked as a priority.

It is estimated that in the United States alone, 45% of people look to Facebook as their main online news source. In Argentina, the country of focus for this research, two-thirds of the population turn to Facebook as their social platform for news. Decisions about the information users receive can have profound effects on our politics, societies, economies and cultures. Recent research shows Facebook's News Feed might impact voter turnout, and even have a small — but statistically significant — impact on users' moods. The effects of Facebook are increasingly moving beyond the social, impacting power, politics and the well-being of citizens.

This report hopes to provide researchers with an experimental method to gain further insights into how the News Feed algorithm operates. It also aims to give users a window into how their information diets are being determined. A set of interactive dashboards have been prepared with the report to explore the findings, including to see which posts were hidden and shown to our profiles.

ARGENTINA: CONTEXT

- **Market:** Connectivity rates in the country are nearly 70%, even in rural areas. The previous and current government have committed to bridge the remaining digital divide — including between women and men. Internet costs are relatively affordable — Argentina meets the Alliance for Affordable Internet's "1 for 2" affordability target, with 1 GB of mobile data available for 1.58 percent of the average monthly income.

- **Institutional protection of information flows:** Freedom of expression is average. Argentina ranks 50th out of 180 countries on the World Press Freedom Index. However, internet freedom is seen as mixed and there are increasing concerns regarding media consolidation as traditional media, telecom, and cable providers attempt to merge.[16]

- **Growth of social media:** Social media platforms are quickly taking on a mediating role between users and web content, including online news sites: Facebook is the most visited website in Argentina.[17] Facebook has achieved 75% penetration among Argentine internet users, of which 61% access the platform at least daily.

- **Role of social media in activism:** Social networks, mostly Facebook and WhatsApp, have been key platforms for public debate and for coordinating collective action, from the platforms to the streets. This has been the case for feminist protests that led to #NiUnaMenos ("Not One Less") — a rallying cry for effective policies against femicides and violence against women — as well as public demonstrations of thousands to condemn the disappearance of a young activist.

- **Role of social media in electoral politics:** The political party currently in power has relied heavily on social media, particularly Facebook, to target voters.[18] Facebook has highlighted how the current President, Mauricio Macri, reached more than 92% of Facebook users in Argentina during the last electoral race.

---

16 https://www.wsj.com/articles/telecom-argentina-cablevision-announce-merger-1498924578 , https://www.publicknowledge.org/news-blog/blogs/argentinas-dangerous-path-toward-media-and-communications-dominance ; http://www.ambito.com/907340-preocupante-aprobo-el-enacom-fusion-telecom-cablevision.
17 Slide 35 https://www.slideshare.net/wearesocialsg/digital-in-2017-south-america See also: https://www.digitalhouse.com/noticias/las-redes-sociales-mas-elegidas-los-argentinos-marketing-digital/.
18 https://www.facebook.com/business/success/mauricio-macri#u_0_0; See also http://www.lapoliticaonline.com/nota/99814/.

02

METHODOLOGY: WHAT WE DID

The project was designed to provide Facebook users and researchers a window into the process of algorithmic curation carried out by social media platforms during critical moments when public opinion is being shaped.

The project aims to showcase a method and set of tools that can help dig deeper into these questions, and to highlight a number of key aspects of our online experience in which Facebook could improve user control and choice.
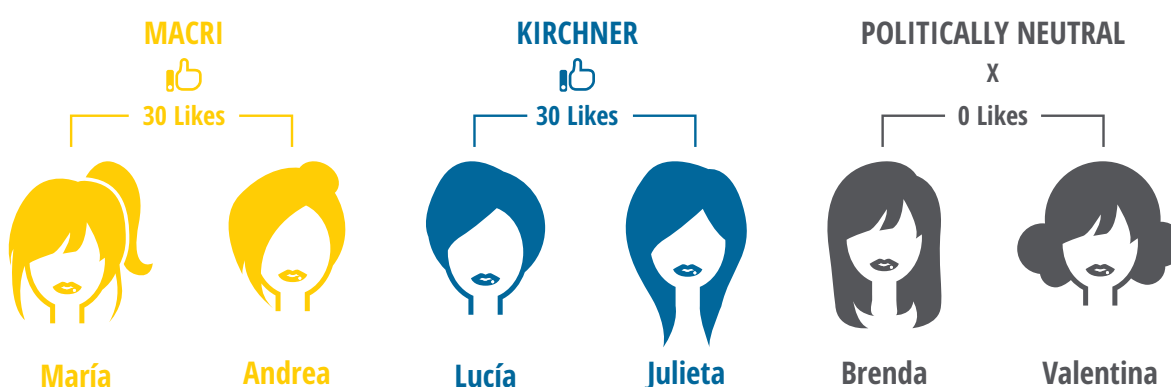
Beginning in October 2017, some initial experiments were run to test our tools and methods.[19] We then ran the next iteration in early December, setting up six Facebook profiles in Buenos Aires on virtual machines. All six profiles were created on the same day in December. To reduce the scope of variables potentially influencing the results, all six profiles were set up as 30-year-old women: Lucía, Julieta, Brenda, Valentina, María and Andrea.[20]

---

19 Ahead of Argentina's legislative elections on October 22, 2017, nine Facebook profiles were set up specifically for this research project to follow 40 online news and social media sources. The tool was used to collect data for over two weeks (some days before the election day and some days after the election). To read all the technical details, please refer to the methodology document, which provides a detailed account of the steps and iterations of the experiment.

20 All six hypothetical profiles on Facebook were created as new profiles on the same day. No personal data other than name (female) and age (30) were included. Sim cards were purchased in Buenos Aires to create the profiles using unique phone numbers. Any other personal data required to be provided to Facebook was to be completed in such a way that all profiles provided the same data.

These six profiles were set to follow the same 20 Facebook Pages – run by online news sources – and two political pages (one belonging to Mauricio Macri, the current president, and one set up for Cristina Fernandez de Kirchner, a former president and a key figure of the opposition). On December 15 — one week into the experiment — we polarised our profiles. Of the six profiles, two "liked" 30 posts published by Macri's Facebook Page, two "liked" 30 posts published by Kirchner's Facebook Page, and two remained politically neutral. The purpose of the variation was to determine whether this slight polarisation would lead to changes in the news delivered to a Facebook user.

MACRI		KIRCHNER		POLITICALLY NEUTRAL	
👍		👍		X	
30 Likes		30 Likes		0 Likes	
María	Andrea	Lucía	Julieta	Brenda	Valentina

The data collection spanned 11 days in two phases: (1) from December 8-15, during which time data was collected before the users were polarised, and (2) from December 16-19, during which time data was collected after the user profiles had been polarised.[21] The experiment looked at the information that these six users were exposed to through their News Feed. This period of time was of particular interest given the meetings of the World Trade Organization in Buenos Aires, and protests around proposed reforms of Argentina's public pension system.

The black-box nature of the News Feed algorithm creates a broad set of challenges that discourage key research from being carried out on its nature and implications.[22] In this context we chose to employ an auditing approach, common to these scenarios.[23] The data was collected using Facebook Tracking Exposed, an open-source, web-based extension installed via Google Chrome.[24] The tool collects the full HTML component of every public post appearing on the browser. We configured the six users to access Facebook daily and at the same time. The access was scheduled with a UserScript,[25] which automated the six users for their daily visits

---

21 The profiles were set up on Sunday, December 5, 2017. The experiment ran December 8-15 2017.

22 Devito, M. (2017). From Editors to Algorithms: A values-based approach to understanding story selection in the Facebook news feed. Digital Journalism, 5(6), 753-773.

23 See for example, Chen, L., Mislove, A., & Wilson, C. (2015). Peeking Beneath the Hood of Uber. Proceedings of the 2015 Internet Measurement Conference, 495-508.; Bertrand, M., & Mullainathan, S. (2004); and, Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. The American Economic Review, 94(4), 991-1013.

24 https://facebook.tracking.exposed/.

25 UserScript technology https://en.wikipedia.org/wiki/Userscript; Our Auto-scroller script https://facebook.tracking.exposed/autoscroll.user.js.

to Facebook that entailed scrolling through their News Feed. This scrolling exposed each user to an average of 45 impressions (posts) per connection.[26] During the first week of data collection, each user connected and performed three sets of scrolls per day. After the users were polarised, the automation was set so that each user performed seven scrolls per day.[27] The scrolls were set to occur during the day to increase external validity, and coordinated across the profiles to increase internal validity.[28] The times were set to reflect the activity of real users and were as follows:

First week: 08:01, 16:01, 23:55

Final days: 8:01, 10:01, 12:01, 14:01, 16:01, 18:01, 20:01

In order to document the complete information sphere from which Facebook was defining the posts it would show on each user's News Feed, we collected the 11,603 posts published by the 22 Facebook Pages under observation through Facebook's public API.[29] A set of dummy variables was set up for keywords that were relevant during this time for the country's news cycle.[30]

---

26 The variation is explained by the type of post: Pictures and videos often take more space than text updates. We call them impressions, instead of posts, because a user can be exposed to the same post more than once.

27 Due to factors outside of our control, the data collection period was cut short. To increase the number of data points for the post polarization phase we increased the number of daily scrolls, thus having a stronger base of data to analyze.

28 Due to technical reasons, there were some seconds or minutes of difference between the scrolls. However this difference should not have significantly impacted the profiles' exposure to posts given the frequency of new posts by the pages that were being followed.

29 All the collected data by the profiles can be found here: https://github.com/tracking-exposed/experiments-data. Following the data collection, by relying on the concatenation of Page name, and date and time of publication we joined the dataset collected using Facebook Tracking Exposed and the dataset retrieved through Facebook's API. Based on this column we could observe which of the published posts were shown to which profiles, and which posts were hidden.

30 The keywords were defined by the researchers, relying on an automated script that counted repetition of certain keywords across the underlying news articles tracked. The raw keywords can be seen by exploring the data available at the Facebook Tracking Exposed GitHub T https://github.com/tracking-exposed/experiments-data and the report's Annex: https://docs. google.com/spreadsheets/d/1pJLixlZtI6vz3faVxUV3wF8ugZpKQzvZyXUiWe60u-w/edit#gid=1967898836. The raw dataset can be sorted for: location data, every time each user refreshed the News Feed, the date and time each time a user refreshed the News Feed, the time of the impression of every post displayed by News Feed to a user, the HTML code of every public post in the users' News Feed, and metadata related to the posts.

Fig. 1 — Information sphere: All 11,603 posts published by the
22 Facebook Pages[31] (December 8-19, 2017)
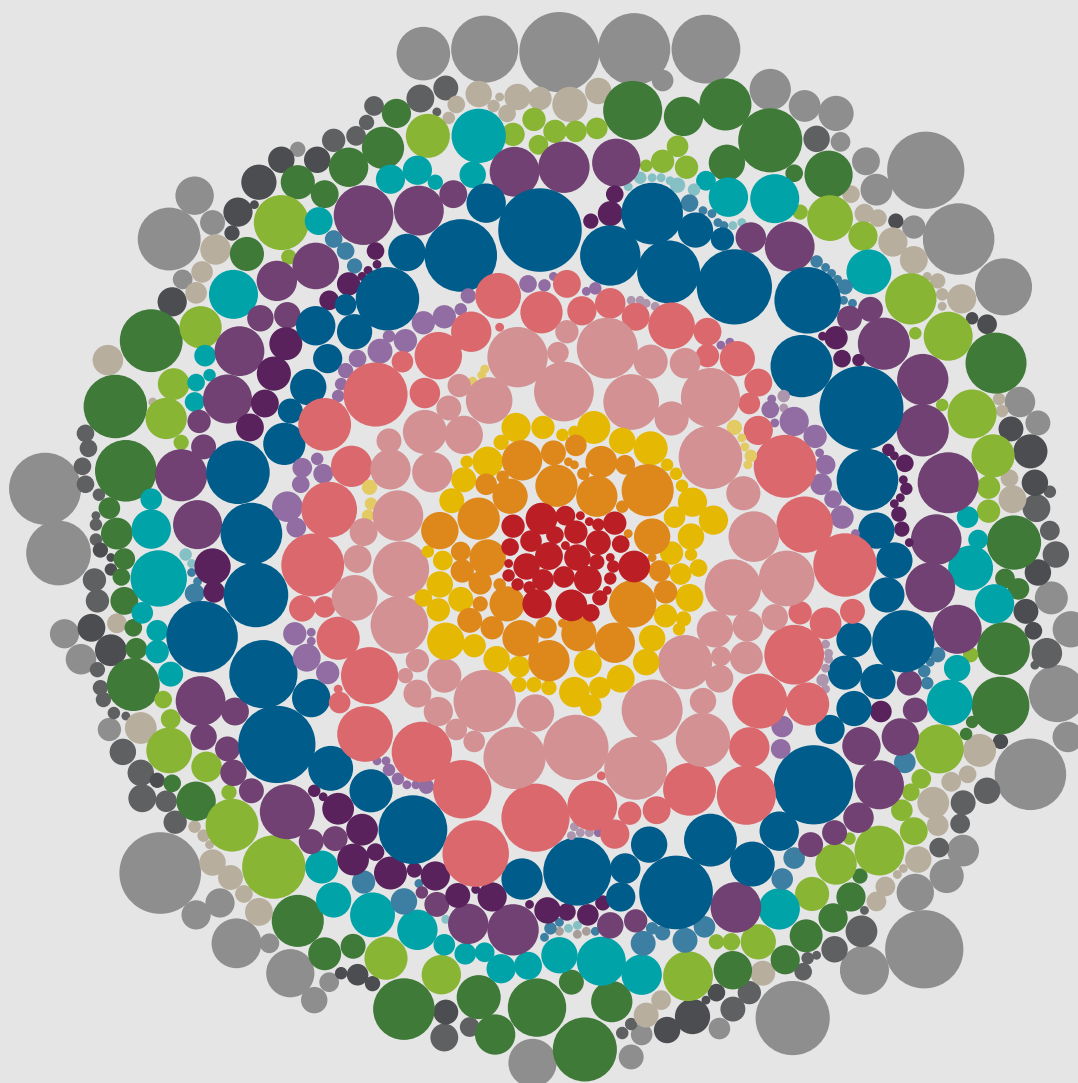
Page Name

🟥	Agencia Paco Urondo	🟦	La Política Online
🟧	AmericaTV	🟦	Mauricio Macri (President)
🟨	C5N	⬜	Noticiero Trece
🟫	Clarín	🟦	Pagina 12
🟨	Cristina CFK (former-president)	🟩	Perfil
🟥	Cronica	🟩	Radio Mitre
🟪	Diario Popular	⬜	Radio10
🟪	Garganta Poderosa	⬛	Telenueve Noticiero
🟦	Infobae	⬛	Tiempo Argentino
🟪	INFOnews	⬜	TN Todo Noticias
🟪	La Nacion		

Fig. 1 —

03

ANALYSIS: WHAT WE FOUND

A. Facebook continues to heavily curate what we see

Perhaps the most interesting aspect revealed was the differing scope of information that users with the same profile receive — either by design or defect. The result is that even users with the same profile who are following the same sites have different information diets.

It is no secret that Facebook plays a filtering function based on user profiles to sort through the immense quantity of information that is posted to the social media platform. The findings of our research suggest that this curation extends to users with seemingly identical profiles. As Figure 2 shows, of the 11,603 posts published during the 11 days observed, Andrea only saw 1192 posts, and our six profiles combined were only exposed to 2,071 — just 18% of all the posts published across the 22 Facebook Pages. This is the equivalent of being exposed to one out of every six posts from across all the sites being followed. Andrea also saw very different posts when compared to Lucía. Despite following the same pages and going online at the same time, roughly half of Andrea's information diet was excluded from Lucía's timeline, and vice-versa.[32] It is important to understand the decisions behind which of the 11,603 posts are made visible to users. The broadness of this gap between what is seen and hidden represents the breadth of the curation potentially taking place via the platform that is outside the control of users.
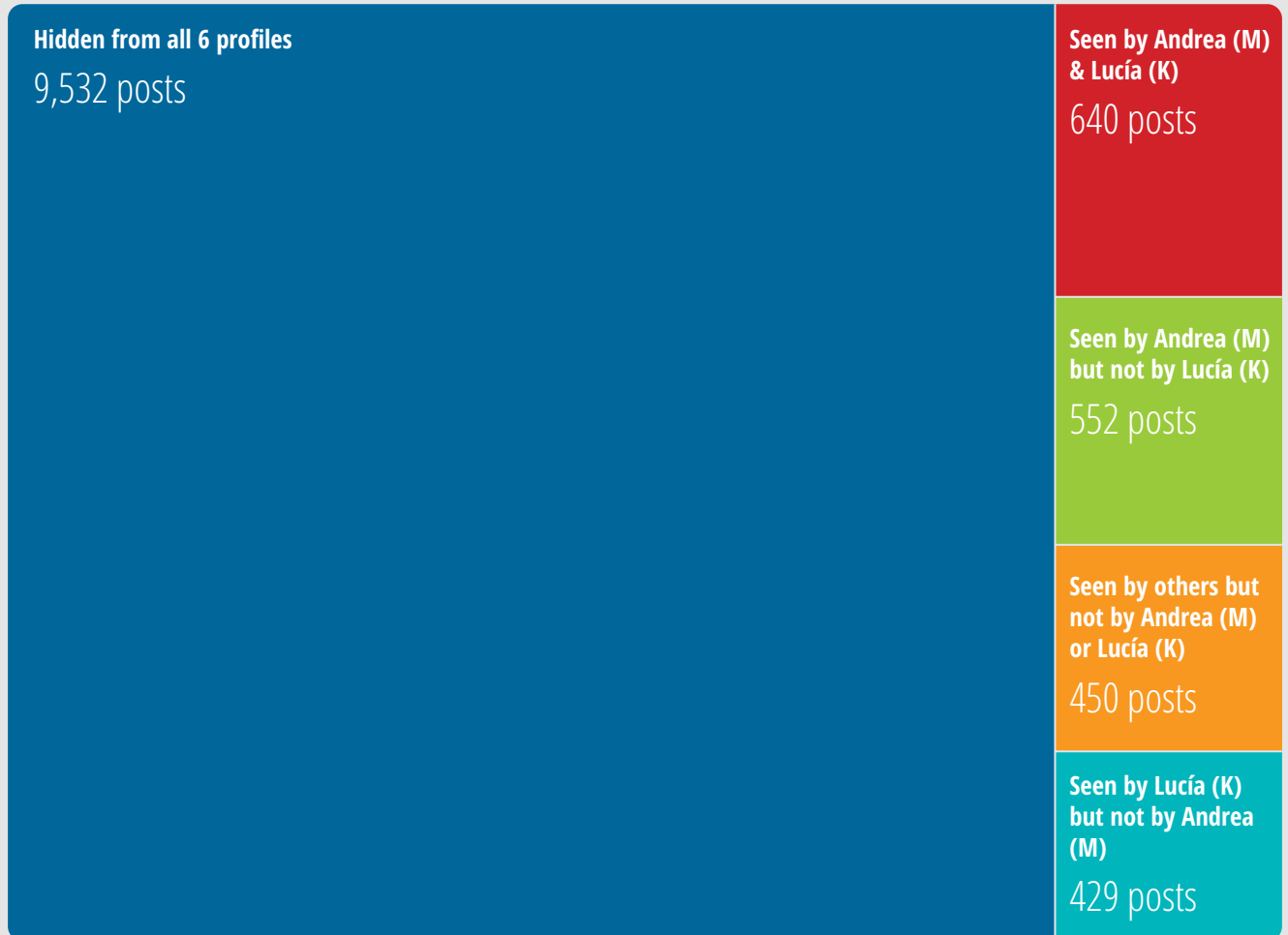
Moreover, it is important to understand whether Facebook users are aware that such curatorial decisions are being made. A recent study based on a survey of 147 students in the US suggests that most users are unaware their News Feed is filtered.[33]

---

32 This figure is based on time bins 1-17, before the profiles were polarised. See the interactive graph here https://public.tableau.com/profile/web.foundation#!/vizhome/Theinvisiblecurationofcontent-Full-English/PieChartProfiles.

33 "Only 24% were aware that Facebook prioritizes certain posts and hides others from users' feeds. (...) 37% believed every post is included in the newsfeed." See Powers, E. (2017). My News Feed is Filtered?: Awareness of news personalization among college students. *Digital Journalism, 5*(10), 1315-1335.

Fig. 2 — Treemap showing what our profiles saw compared to what they did not see.

(Size of each box is proportional to the number of posts).

Hidden from all 6 profiles

9,532 posts

Seen by Andrea (M) & Lucía (K)

640 posts

Seen by Andrea (M) but not by Lucía (K)

552 posts

Seen by others but not by Andrea (M) or Lucía (K)

450 posts

Seen by Lucía (K) but not by Andrea (M)

429 posts

The public square as seen by Facebook

The public square as seen by each of our users

Whereas digital platforms often claim they are allowing the public square to go to scale, what we are observing is that, through the use of algorithms, they are placing each user into separate versions of the public square, which are micromanaged by their algorithms. These filtering decisions put Facebook (and other platforms) in a place of control over the public discourse, which brings with it key corporate responsibilities.[34]
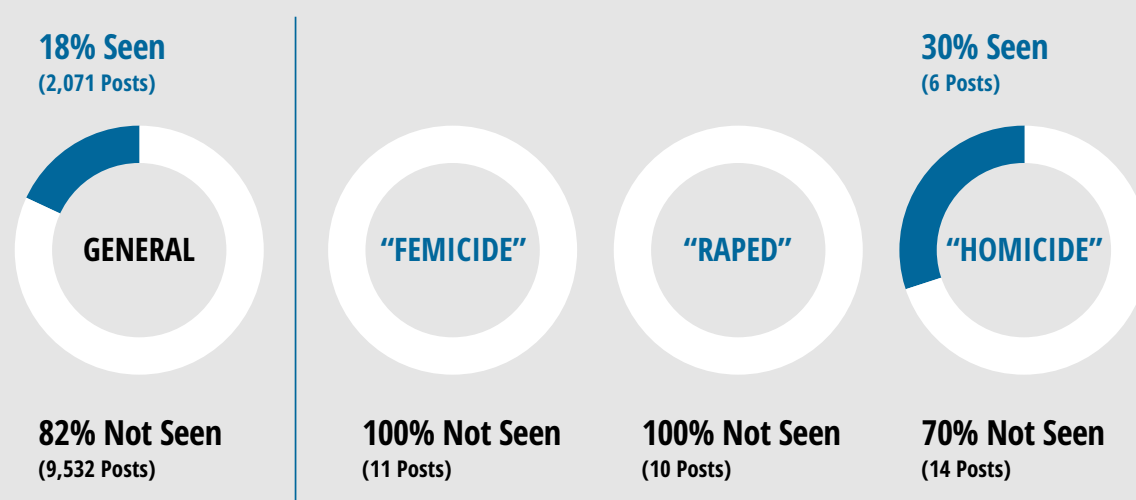
---

34 Koene A. et al. (2015) Ethics of Personalized Information Filtering. In: Tiropanis T., Vakali A., Sartori L., Burnap P. (eds) Internet Science. INSCI 2015. Lecture Notes in Computer Science, vol 9089. Springer, Cham; Taddeo, M., & Floridi, L. (2016). The Debate on the Moral Responsibilities of Online Service Providers. Science and Engineering Ethics, 22(6), 1575-1603.; Koene, A., Vallejos, E. P., Webb, H., Patel, M., Ceppi, S., Jirotka, M., & McAuley, D. (2017). Editorial responsibilities arising from personalization algorithms.

B. Algorithmic curation can hide critical social
and political news from some users.

There are some posts that might be relevant for society at large, or for an online user, that might not surface on Facebook, given the variables it has chosen to prioritise in the design of its algorithm.[35] When a vast proportion of the population relies on the platform to access news, this can quickly turn into a serious problem. Given these risks, Facebook must design tools that allow for full transparency regarding what information is seen and not seen by users.

Fig. 3 — Percentage of posts containing certain
keywords our profiles were exposed to[36]

18% Seen
(2,071 Posts)

30% Seen
(6 Posts)

GENERAL   "FEMICIDE"   "RAPED"   "HOMICIDE"

82% Not Seen
(9,532 Posts)

100% Not Seen
(11 Posts)

100% Not Seen
(10 Posts)

70% Not Seen
(14 Posts)

Three out of every five Argentines use Facebook to get their online news. If Facebook was the only source of information relied on by our six Argentine users they likely would not have been made aware, for example, of stories of femicide and rape happening in the country. Both these issues have taken on heightened importance in the public and political discourse, leading to social protests (on social media and in the streets) against femicide and rape.[37]

---

35 To explore the differences in the post topics appearing - or not appearing - on the profiles' feeds, see: https://public.tableau.com/profile/web.foundation#!/vizhome/LaManoInvisible-TableroAvanzado/DashReactionsShares.
36 Explore further at https://public.tableau.com/profile/web.foundation#!/vizhome/Theinvisiblecurationofcontent-Full-English/PickabooDash . Version in Spanish available at https://public.tableau.com/profile/web.foundation#!/vizhome/LaManoInvisible-TableroAvanzado/DashReactionsShares.
37 As expected, the algorithm is particularly sensitive to the number of reactions and shares, which is in itself an editorial decision. Yet, as the graph shows this is not the only factor, and in the specific case under scrutiny some of the stories on homicide performed worse than omitted stories on femicide and rape that did not surface. See: https://public.tableau.com/profile/web.foundation#!/vizhome/Theinvisiblecurationofcontent-Full-English/DashboardSharesReacts.
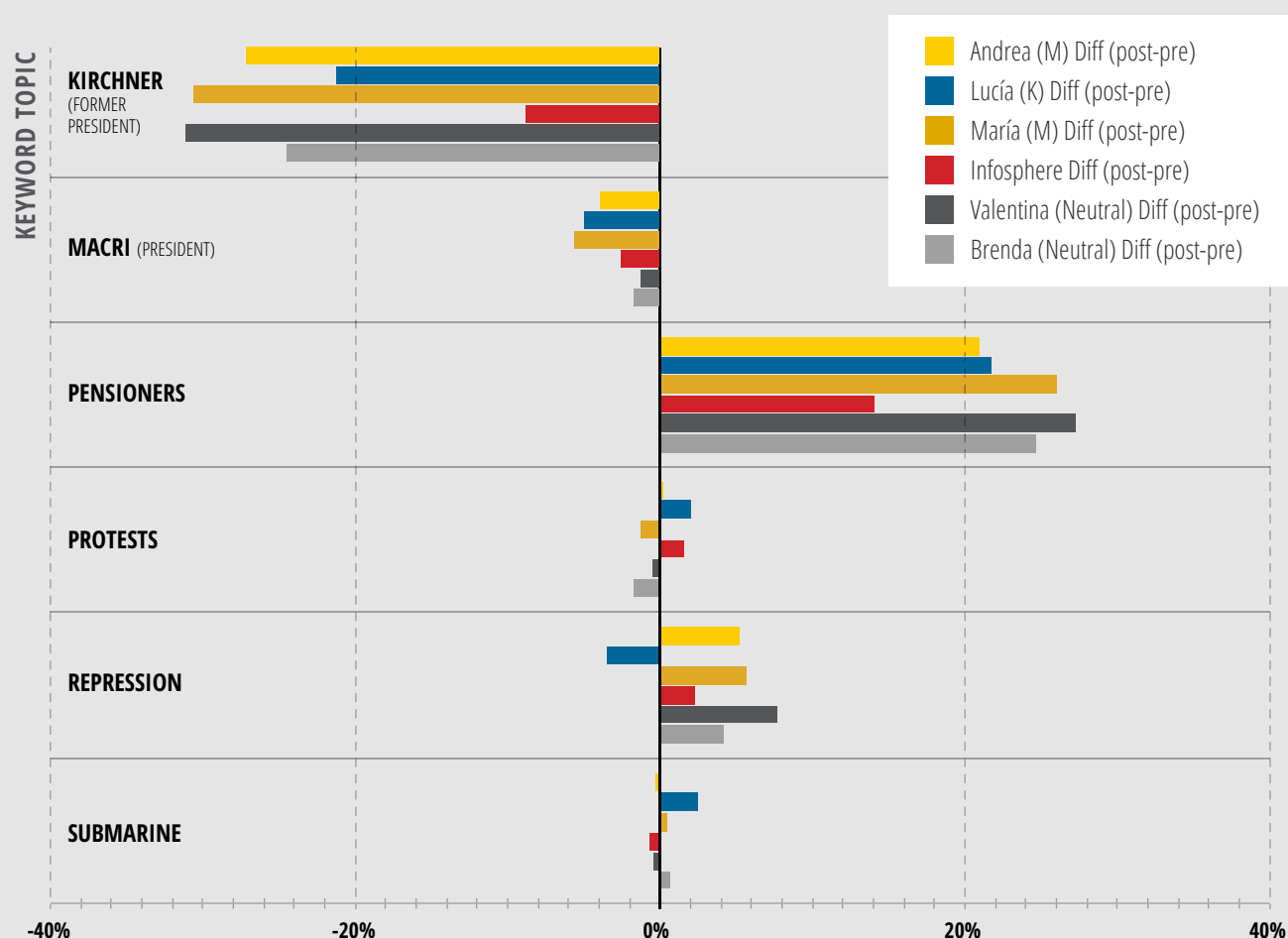
C. Facebook's curation means that users with the same profile following the same news pages are not being exposed to the same stories published on the platform

Facebook's footprint in Argentina and globally as the main news source for many citizens means its algorithms — and changes to them — affect information flow, public discourse and society as a whole.

The experiments found that different profiles were shown different stories for a specific group of topics: protests, repression, and the disappearance of a military submarine.[38] For example Lucía (K), who liked posts from a key figure of the opposition, saw more posts about protests and the missing submarine than others in the study. Lucía (K) was also exposed to less posts about the repression that took place in the context of such protests, which would seem counterintuitive to the hypothesis that stories are selected based on political profiling. As a result, while this study is able to demonstrate that the delivery of stories differs across similar profiles, it is unable to confirm whether Facebook's algorithm reacts to small signs of political interest (see Figure 4). Further research should be done to test this hypothesis.[39]

Fig. 4 — Percentage points change in posts containing selected keywords, before and after polarisation[40]

Legend:
- Andrea (M) Diff (post-pre)
- Lucía (K) Diff (post-pre)
- María (M) Diff (post-pre)
- Infosphere Diff (post-pre)
- Valentina (Neutral) Diff (post-pre)
- Brenda (Neutral) Diff (post-pre)

KEYWORD TOPIC:
KIRCHNER (FORMER PRESIDENT)
MACRI (PRESIDENT)
PENSIONERS
PROTESTS
REPRESSION
SUBMARINE

X-axis: -40%, -20%, 0%, 20%, 40%

38 The ARA San Juan, which became a national tragedy and led to public criticism of the current Macri government for its handling of the matter. See: https://www.theguardian.com/world/2017/nov/30/argentina-calls-off-missing-submarine-rescue-effort.

39 For more information on these results see Annex II. Also: Interactive graph showing variation in these percentages over time https://public.tableau.com/profile/web.foundation#!/vizhome/Theinvisiblecurationofcontent-Full-English/Selectedkywdsdash and distribution https://public.tableau.com/profile/web.foundation#!/vizhome/Distributions-Invisiblecurationofcontent-English/AdvancedDlst.

40 To reduce noise this graph relies on a smaller subset of posts, during which data collection was most stable. Pre-polarisation: Dec 6 at 9pm to Dec 12, 5pm. These are represented as Time Bins 1-17 in the dataset and interactive graphs. Post-Polarisation: , which took place between Dec 15 at 5pm and Dec 18 at 7pm. These are represented as Time Bins 27-47 in the dataset and interactive graphs. Note: Julieta's profile was excluded from this analysis since her profile crashed on Dec 17, and failed to collect data at 3pm, as was expected (Time Bin 38).

04

CONCLUSIONS AND RECOMMENDATIONS

This study documents the changing dynamic between an online user, the news sources she follows, and the news received via social media platforms as a result of:

- **Curation of news content:** This implies that information diets — whether on Facebook or other online platforms — are defined by algorithms that take control over how users are informed of, and view, the public sphere, often making seemingly relevant news invisible to users.

- **The opaque nature of algorithms:** Facebook appears to define users' information diets based on criteria that are not visible to users. Furthermore, the algorithm is complex or obscure enough that even seemingly identical users following the same information sources are getting different posts on their timelines.

What these developments mean is a very different information diet for each person.

Following Facebook news pages — as our six users did — was the equivalent in the past of being subscribed to various newspapers and having them delivered every day. Before social media, it would have meant that these six users would have been exposed to exactly the same news — like neighbours getting the same papers delivered to them. The only 'curation' would be based on what any given reader decided was of interest and what she had time for. She determined her information diet.

Our research finds that Facebook's algorithms appear to decide which newspapers you get, when you get them and what you'll find inside. The same algorithms also affect what your 'neighbours' get — meaning that no two people will find the same papers or the same articles for them waiting outside their doorsteps.

In the case of Argentina, a country with intense social media engagement, this creates a scenario where Facebook users' understanding of the world is potentially curated by algorithms.

These findings suggest that:

- Algorithmic curation is persistent and appears to rely on seemingly arbitrary factors that users do not control or inform.

- Users need to better understand how technology works and impacts their information diets and perception of reality. This can be done by explaining the phenomenon with examples that show the dynamics behind the algorithmic curation of content.

- Platforms, especially those playing a prominent role in defining a country's news diet, need to disclose their curation practices, be open to reviewing those mechanisms with users, and inform users about how their information diets are shaped.

- There is a need to better understand whether laws and regulations are adequate to face the challenges posed by algorithms that personalise information delivery.

Some suggestions for Facebook to begin addressing these challenges could include to:

- **Explore the use of real-time transparency dashboards** for users to compare what posts were seen versus what was published over a given period of time, in order to increase public understanding of how the algorithm works.

  - This could include filters to navigate the information by source and keyword.

- **Make chronological ordering the default mode** for content delivery via Facebook.[41]

- **Allow users an array of options to manage their own News Feed settings** and play a more active role in constructing their own information diet.[42]

- **Be responsive to Facebook users by working with them** to actively shape platform decisions and feedback on their user experience as a result of the algorithms being used.

- **Facilitate independent research** on the externalities and systemic effects from the algorithmic curation of content.

These suggestions are a way for Facebook to achieve its mission to "give people the power to build community and bring the world closer together." They are critical steps towards putting the control of users' information diets back in the hands of the community and the users themselves.

---

41 Current settings allow users to opt-out from algorithmic curation. Yet, in the context of informal testing, we observed that re-launching the application reset this preference to the default of algorithmic curation. See https://www.facebook.com/help/community/question/?id=10152476808951132.

42 See for example gobo.social.

05

REFERENCES

Annex I

List of Facebook Pages followed

TYPE	NAME	FB HYPERLINK	FOLLOWERS (2 DEC 2017)
Blog	Agencia Paco Urondo	https://www.facebook.com/agenciapacourondo/	120,636
TV	America 2	https://www.facebook.com/AmericaTV//	1,435,601
TV	C5N	https://www.facebook.com/C5N.Noticias	4,082,608
TV	Canal 13	https://www.facebook.com/NoticieroTrece/?ref=br_rs	93,675
TV	Canal 7	https://www.facebook.com/tvpublicanoticias/	144,586
TV	Canal 9	https://www.facebook.com/telenueveC9/?ref=br_rs	81,314
Newspaper	Clarin	https://www.facebook.com/clarincom/	5,684,851
Politician	Cristina Fernandez de Kirchner	https://www.facebook.com/CFKArgentina/	2,358,232
Newspaper	Cronica	https://www.facebook.com/cronicadiario/	1,435,565
Newspaper	Diario Popular	https://www.facebook.com/populardiario/	112,030
Blog	Garganta Poderosa	https://www.facebook.com/La-Garganta-Poderosa-213440425391495/?fref=ts	505,251
Newspaper	Info News	https://www.facebook.com/INFOnewscom	103,130
Newspaper	Infobae	https://www.facebook.com/Infobae/	2,312,842
Newspaper	La Nacion	https://www.facebook.com/lanacion	3,342,628
Blog	La Politica Online	https://www.facebook.com/pages/lapoliticaonline/304176325478	132,228
Politician	Mauricio Macri	https://www.facebook.com/mauriciomacri/	4,488,499
Radio	Mitre	https://www.facebook.com/radiomitre/	1,458,622
Newspaper	Página12	https://www.facebook.com/Pagina12ok/	391,303
Newspaper	Perfil	https://www.facebook.com/perfilcom	644,724
Radio	Radio 10	https://www.facebook.com/radio10/	398,785
Newspaper	Tiempo Argentino	https://www.facebook.com/DiarioTiempoArgentino/	331,903
TV	TN	https://www.facebook.com/todonoticias/?ref=br_rs	6,825,397
TOTAL			**36,484,410**

Total Posts By Site (during the period of observation):

Page Name

Site	Posts
Infobae	1,856
TN Todo Noticias	1,457
Clarín	1,245
Cronica	1,232
Radio Mitre	1,087
La Nacion	1,060
Perfil	735
Pagina 12	605
AmericaTV	420
C5N	323
Radio10	308
Tiempo Argentino	268
INFOnews	262
Telenueve Noticiero	207
Diario Popular	189
Agencia Paco Urondo	180
La Política Online	111
Mauricio Macri (President)	19
Cristina CFK (former-president)	18
Garganta Poderosa	18
Noticiero Trece	3

Keywords selected for dummy variables (used to test during polarization)

VARIABLE NAME	BOOLEAN	EXTRA (WITH MANUAL CONTROL OF RESULTS)
CFK mention	CFK|Cristina Fernandez|Cristina Kirchner	Cristina
Macri mention	Mauricio Macri|Macri	el presidente|Mauricio
Movilización (y vars)	Protesta|Movilizació|marcha|cacerolazo	None
Jubilados	previsional|jubila	None
Cacerolazo	Cacerola	None
Represión	Represio|reprimi|balas de goma|gas lacrimógeno	None
Submarino	ARA San Juan|Submarino	None
Inflación	inflacio	None

(Annex II)

Technical documentation

http://webfoundation.org/docs/2018/04/Annex-FB-Tests-Final.xlsx

Collaborative research

Share your results back and let's build a collective effort to understand better how Facebook works. If you are using the data set, we encourage you to share the data set by using the hashtag #LaManoInvisible #InvisibleHand or simply share your research sending us an email to contact@webfoundation.org.

Interactive graphs available at: https://public.tableau.com/profile/web.foundation#!/vizhome/Theinvisiblecurationofcontent-Full-English/PickabooDash (English)

https://public.tableau.com/profile/web.foundation#!/vizhome/LaManoInvisible-TableroAvanzado/DashReactionsShares (Spanish)

WORLD WIDE WEB
FOUNDATION