News and Political Information Consumption in Mexico: Mapping the 2018 Mexican Presidential Election on Twitter and Facebook

COMPROP DATA MEMO 2018.2 – ONLINE SUPPLEMENT/ JUNE 29, 2018
comprop@oii.ox.ac.uk

Monika Glowacki
Oxford University
@MoniGlow

Vidya Narayanan
Oxford University
@vidunarayanan

Sam Maynard
Oxford University
@slaynard

Gustavo Hirsch
Oxford University
@gustavoh1985

Bence Kollanyi
Oxford University
@bencekollanyi

Lisa-Maria Neudert
Oxford University
@lmneudert

Phil Howard
Oxford University
@pnhoward

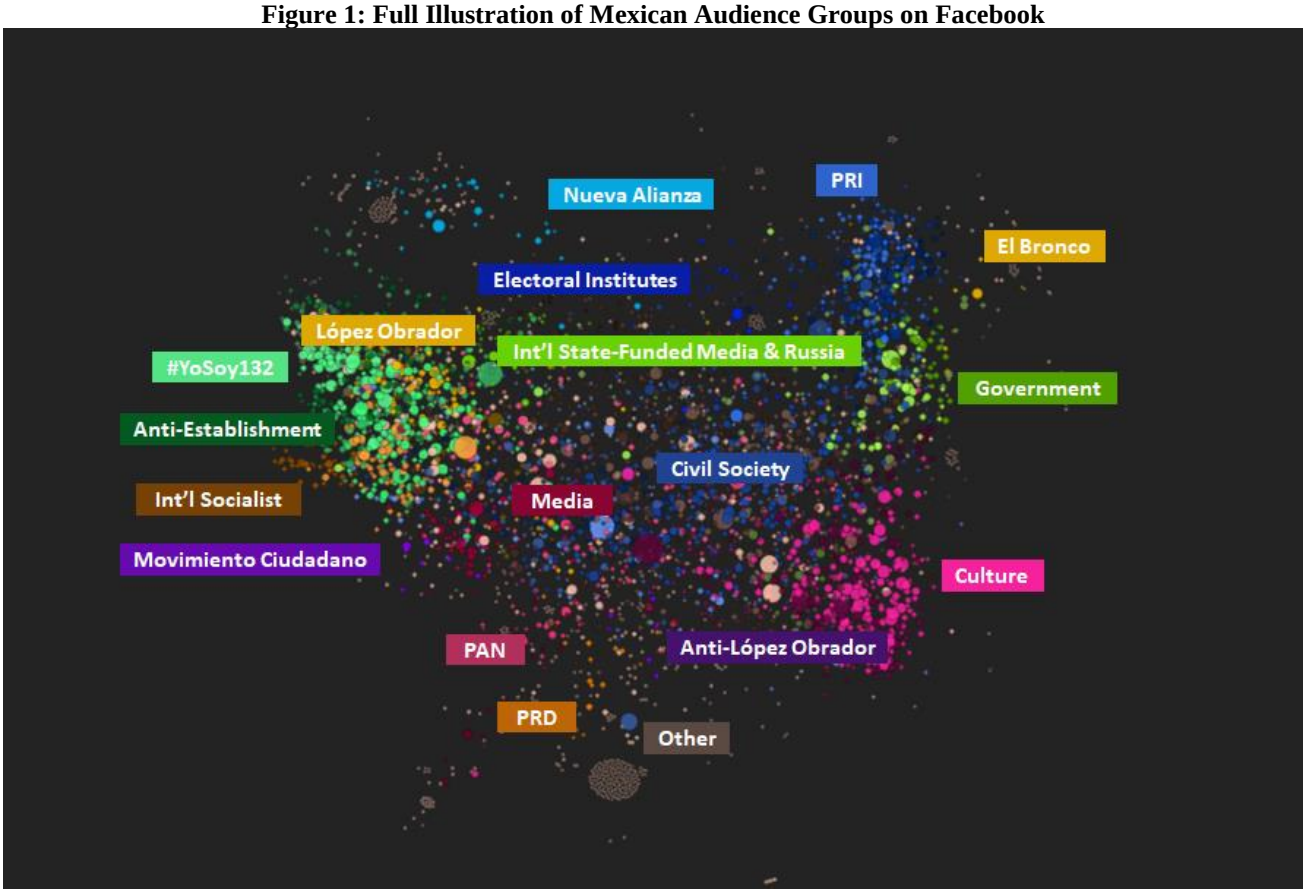
Thomas Lederer
Graphika

Vlad Barash
Graphika
@vlad43210

1. Mexican Audience Groups on Facebook

Each node in this network represents a public page on Facebook. Each node belongs to both a broad *group* and a smaller *cluster* within that group. The nodes are placed within the map using a Fruchterman-Reingold visualization algorithm. This works to place nodes into the map according to two principles: first, a centrifugal force acts upon each node to push it to the edge of the canvas; second, a cohesive force acts upon every connected pair of nodes to push them closer together. A cluster is a collection of nodes with a shared pattern of interest, while a group is a collection of clusters that are politically, culturally, geographically or socially similar. The size of the node is proportional to the number of other map nodes that *like* it on Facebook. The color of the node is based on its parent cluster. *Figure 1* is the full color visualization of Mexican audience groups found on Facebook. *Table 1* gives a full list of Facebook groups and associated clusters considered in this study.

Figure 1: Full Illustration of Mexican Audience Groups on Facebook



Source: Authors' calculations from data sampled 07/03/18—05/06/2018. Note: Segment labels are determined by patterns of likes and groups are determined by shared segment characteristics.

Table 1: List of Groups and Clusters for the Facebook Visualization	
Groups	*Clusters*
#YoSoy132	#YoSoy132 and Other Movements
	#YoSoy132
	#YoSoy132 and López Obrador
López Obrador	López Obrador, MORENA Politicians and Political Orgs, Anti-PRI and -Peña Nieto
	López Obrador, MORENA Party and Politicians
Anti-López Obrador	Anti-López Obrador
Anti-Establishment	Spain Anti-Establishment, Social Movements
	LatAm Anonymous
	North America Occupy Movement
	Int'l Antifa, Anarchist
	Int'l Occupy Movement
	Anti-Establishment, Revolutionary
Civil Society	Int'l Animal Rights
	Entrepreneurship, Environmental Sustainability
	Human Rights
	Civil Society for Media
	Civil Society for Culture
	Int'l Development Agencies, US Gov't
	Int'l Civil Society for Journalism and Culture
	Civil Society
	Spain Civil Society
Culture	Music
	Tourism, Travel
	Film
	Art, Culture
	Art, Culture, Literature
	UNAM
El Bronco	El Bronco
Electoral Institutes	Electoral Institutes
Government	Nat'l Gov't
	State Gov't México
	Local Gov't CDMX
	Local Gov't Guadalajara
	Local Gov't Nuevo León
International Socialist	Venezuela Chavista
	Pro-Cuba
International State-Funded Media and Russia	Int'l Russia RT, Iran HispanTV
	Russia Media and Gov't
Media	Traditional and New Media
	Misc Media
	Alternative Media
Movimiento Ciudadano	Movimiento Ciudadano Party
Nueva Alianza	Nueva Alianza Party and Politicians
Other	Misc Parties
	Misc
	Int'l Colombian Media, Paranormal Encounters
	LatAm Media and Culture
	Mexico Cycling, Urban Mobility
	US Media, Latinos, Environmental Justice
	Peru Media, Lima Universities
	Spain El Pais
	LatAm Fashion
	Int'l Misc
	Int'l Entertainment, Soccer
	Arabic Misc
	Int'l Architecture, Design
	Int'l Rock Music
	Int'l Science, Spirituality
PAN	PAN Party, Politicians and Political Orgs
	PAN Party and Politicians
PRD	PRD Party and Politicians
PRI	Meade, PRI Party and Politicians
	PRI Party and Politicians
	Meade, PRI Party, Politicians and Political Orgs
	Meade, PRI Party and Gov't

2. Clustering Algorithm for Determining Groups and Clusters

To generate clusters and groups for the map it is necessary to employ a clustering algorithm. This involves first building a graph between nodes in the map and the rest of the social medium in question. This graph provides a structural similarity metric between nodes in the map. This is then used in combination with a hierarchical agglomerative clustering algorithm to segment a map into distinct communities. This is a *bottom-up* approach whereby each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy. Facebook networks are clustered based on page likes.

3. Heterophily Index

For every pairing of groups within a network map, a heterophily value can be calculated. This is a measure of the level of connection between the groups. To determine the heterophily score, a ratio is calculated of the actual ties between two groups compared to the expected ties between the groups if all the pages in the Facebook map were evenly distributed. The *natural log* of these ratios is then taken, along with a *zero correction* to create a balanced index and ensure that all values are displayed in a positive form.

$$Ratio\ of\ Ratios_T = \frac{\frac{Connections_{\text{pairing}}}{\sum_{\text{all pairings}} Connections}}{\frac{Connections_{\text{pairing}}}{\sum_{\text{all pairings}} Connections}}$$

Expression A: Ratio of Two Ratios

This heterophily index is therefore created through a ratio of two ratios. The ratio of these two ratios reveals whether two nodes have about the proportion of links they should have given its size. This is displayed in *Expression A*, where a pairing of groups is calculated as having a measure of connections in balance with its share of all the connections.

Half the distribution of possible values from this ratio of ratios ranges from 0 to 1 (a disproportionately small share of connections in a group given its size) and the other half ranges from 1 to +infinity (a disproportionately large share of connections in a group given its size). However, by taking the natural log of the ratio of ratios the index will become more balanced: from -infinity to 0 becomes less than proportionate share, and from 0 to +infinity becomes more than proportionate share. For example, take a three-group network A, B and C. If nodes in group A have a total of ten connections, and there are ten nodes in each group, then the expected connections between A and B will be 3.33. If, in reality, the nodes in group A actually have all ten connections to nodes in group B then this connection is stronger than expected. The heterophily score for groups A and B = 10 / 3.33 = 3.0. The natural log of this is then taken along with a zero correction across the range of heterophily values.

A greater heterophily index indicates a denser pattern of connections between the two groups. It is important to note however that these scores indicate only first order connections, not second or third order connections.

Table 2 shows the heterophily score between each pairing of Mexican audience groups in our Facebook map.

Table 2: Heterophily Index for Mexican Audience Groups on Facebook

Group	#YoSoy132	López Obrador	Anti- López Obrador	Anti-Establishment	Civil Society	Culture	El Bronco	Electoral Institutes	Government	Int'l Socialist	Int'l SF Media & Russia	Media	Movimiento Ciudadano	Nueva Alianza	Other	PAN	PRD	PRI
#YoSoy132	15.4	2.5	0.0	2.2	0.4	0.2	0.1	0.1	0.2	1.0	0.0	1.5	0.1	0.0	0.4	0.1	0.1	0.0
López Obrador	0.0	11.7	3.6	2.3	0.4	0.2	0.1	0.2	0.2	0.8	0.0	1.5	1.8	0.1	0.7	0.3	1.0	0.0
Anti- López Obrador	0.0	0.0	60	0.2	0.4	0.0	0.0	0.8	0.0	0.0	0.0	0.1	0.0	2.1	0.7	1.1	0.0	1.1
Anti-Establishment	0.0	0.0	0.0	5.5	0.6	0.2	0.1	0.0	0.1	2.4	3.5	1.2	0.3	0.0	0.6	0.1	0.0	0.0
Civil Society	0.0	0.0	0.0	0.0	2.1	1.0	0.4	0.8	0.9	0.5	0.3	1.0	0.4	0.2	0.9	0.7	0.8	0.5
Culture	0.0	0.0	0.0	0.0	0.0	5.8	0.2	0.4	1.4	0.2	0.0	0.8	0.2	0.2	1.0	0.5	0.2	0.3
El Bronco	0.0	0.0	0.0	0.0	0.0	0.0	63	0.0	5.7	0.0	0.0	0.5	7.4	0.0	1.0	0.4	0.0	0.5
Electoral Institutes	0.0	0.0	0.0	0.0	0.0	0.0	0.0	78	1.7	0.1	0.0	0.6	1.5	2.7	0.9	1.0	1.3	0.8
Government	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.1	0.1	0.0	0.5	0.5	0.3	0.9	1.0	0.4	3.8
Int'l Socialist	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10.7	1.8	1.0	0.2	0.0	0.6	0.1	0.0	0.0
Int'l SF Media & Russia	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	28.5	0.2	0.0	0.0	0.8	0.0	0.0	0.0
Media	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.7	1.3	0.0	0.8	1.4	0.4	0.3
Movimiento Ciudadano	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	138.7	0.0	0.9	0.2	0.7	0.4
Nueva Alianza	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	50.5	0.9	0.2	0.0	0.4
Other	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.0	0.6	0.6
PAN	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	24.5	0.3	0.6
PRD	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	78.0	0.2
PRI	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.1

Source: Authors' calculations from data sampled 07/03/18—05/06/2018. Note: Int'l is abbreviated from International. SF is abbreviated from State-Funded.

4. List of Hashtags Tracked

Table 3: Hashtags Tracked for Twitter Data Collection on Mexican Election

Candidate	*Candidate Handle*	*Candidate Hashtags*	*Party*	*Party Handle*
Ricardo Anaya	@RicardoAnayaC	#MéxicanosAlFrente, #DeFrenteConAnaya, #MiCausaConAnayaEs, #PorMéxicoAlFrente, #AnayaPresidente, #MujeresConAnaya, #DeFrenteAlFuturo, #RicardoAnaya, #Anaya, #MxEnPaz, #AnayaGana, #AnayaMiente, #MentirasDeAnaya, #DiNoaRicardoAnaya	PAN (Partido Acción Nacional)	@AccionNacional
Andrés Manuel López Obrador	@lopezobrador_	#YoConAMLO, #JuntosHaremosHistoria, #AMLO, #MORENAva, #AMLOpresidente2018, #TodosConAMLO, #AMLOpresidente, #LópezObrador, #AMLOve, #AMLOvers, #AMLOporLaPaz, #AMLOnoPagaBots, #VotoMasivoMORENA, #AMLO2018, #MéxicoConAMLO, #AMLOContesta, #AsíNOAndrés, #NiUnVotoaAMLO, #NiUnVotoaMORENA, #amloNO, #DiNoaLópezObrador, #AMLOprotegeDelincuentes	MORENA (Movimiento Regeneración Nacional)	@PartidoMorenaMx
José Antonio Meade	@JoseAMeadeK	#YoConMeade, #MeadePresidente, #GanóMeade, #AvanzarContigo, #MeadeEsElMejor, #Meade, #MujeresAliancistasConMeade, #DeCiudadanoACiudadanas, #YoPorMeade, #PepeMeade, #JoseAMeadeK, #Meade2018, #VotaMeade, #MeadeMiente, #DisculpateYRenuncia, #NiUnVotoAlPRI, #MeadeMentiroso	PRI (Partido Revolucionario Institucional)	@PRI_Nacional
Jaime Rodríguez Calderón	@JaimeRdzNL	#MéxicoBronco, #PropuestasBroncas, #PláticaBronca, #ElBronco, #Bronco, #JaimeRdz, #JaimeRodriguezCalderon, #ElBroncoMiente, #NiUnVotoAlBronco	Independent	—
General Election	—	#EleccionesMéxico, #EleccionesMéxico2018, #VotoMX, #VotoPorMéxico	—	—

5. Size, Coverage and Consistency of Mexican Audience Groups on Facebook

We also computed the size, and coverage and consistency scores for each group based on the number of links they shared to content from Major News Brands, New Media and Start-ups, as well as Junk News and Information and Russia, all categories outlined in our grounded typology in the data memo. The coverage of a group refers to the percentage of news domains, within one of the categories stipulated above, that a group posts links to. The consistency of a group refers to the percentage of the total number of links to news sources, within one of the categories listed above, that is shared by the group. A high coverage score indicates the group shares a wide range of news sources within one of the categories above, while a high consistency score indicates the group plays a key role in spreading them.

Table 4: Size, Coverage and Consistency of Mexican Audience Groups Sharing Types of News on Facebook

Group	Pages N	Pages %	Coverage for Major News Brands	Consistency for Major News Brands	Coverage for New Media & Start-ups	Consistency for New Media & Start-ups	Coverage for Junk News & Information & Russia	Consistency for Junk News & Information & Russia
#YoSoy132	135	3	0.3	0.0	0.4	0.1	0.5	0.1
López Obrador	104	2	0.2	0.0	0.3	0.0	0.3	0.0
Anti-López Obrador	9	0.2	0.1	0.0	0.1	0.0	0.2	0.0
Anti-Establishment	330	7	0.9	0.2	0.9	0.3	1.0	0.4
Civil Society	724	15	0.8	0.2	0.8	0.2	0.6	0.1
Culture	449	9	0.4	0.1	0.4	0.0	0.0	0.0
El Bronco	18	0.4	0.1	0.0	0.0	0.0	0.0	0.0
Electoral Institutes	25	0.5	0.0	0.0	0.0	0.0	0.0	0.0
Government	162	3	0.2	0.0	0.1	0.0	0.0	0.0
Int'l Socialist	124	3	0.4	0.0	0.5	0.1	0.6	0.1
Int'l SF Media & Russia	69	1	0.1	0.0	0.1	0.0	0.3	0.0
Media	210	4	0.6	0.1	0.5	0.1	0.6	0.1
Movimiento Ciudadano	15	0.3	0.1	0.0	0.1	0.0	0.0	0.0
Nueva Alianza	30	0.6	0.1	0.0	0.0	0.0	0.0	0.0
Other	2067	42	1.0	0.3	0.9	0.2	0.9	0.2
PAN	76	2	0.1	0.0	0.1	0.0	0.1	0.0
PRD	20	0.4	0.0	0.0	0.0	0.0	0.0	0.0
PRI	311	6	0.2	0.0	0.2	0.0	0.1	0.0
Total	4878	100	—	1.0	—	1.0	—	1.0

Source: Authors' calculations from data sampled 07/03/18—05/06/2018.
Note: Int'l is abbreviated from International. SF is abbreviated from State-Funded.
Percentages have been rounded to the nearest whole number unless they were below one percent.

6. Limitations

Regarding our Twitter methods and analysis, it is possible that we missed relevant tweets that did not use our selected hashtags and captured irrelevant tweets that did. We ran a two-day test data collection to minimize this risk. Additionally, detecting automation on Twitter is a challenging task and it is difficult to determine whether high frequency tweeting activity was positive amplification of or an attack on a candidate's campaign.

For both Twitter and Facebook, we acknowledge that coding the base URL of a source has limitations, as the story might meet our criteria for junk content while the institution behind it does not. However, our approach is logistically more feasible and gives a better sense of the likelihood of landing on junk content when clicking through the site. Our typology and coding of sources is not intended to be a comprehensive list of all types of information providers in the Mexican media environment.