

# Marc Argemí

## El sentido del rumor

Cuando las redes sociales ganan a las encuestas



Marc Argemí

# El sentido del rumor

Cuando las redes sociales  
ganan a las encuestas

*ediciones península*

© Marc Argemí Ballbè, 2017

Queda rigurosamente prohibida sin autorización por escrito del editor cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra, que será sometida a las sanciones establecidas por la ley. Pueden dirigirse a Cedro (Centro Español de Derechos Reprográficos, [www.cedro.org](http://www.cedro.org)) si necesitan fotocopiar o escanear algún fragmento de esta obra ([www.conlicencia.com](http://www.conlicencia.com); 91 702 19 70 / 93 272 04 47).  
Todos los derechos reservados.

Primera edición: mayo de 2017

© de esta edición: Grup Editorial 62, S.L.U., 2017  
Ediciones Península,  
Av. Diagonal 662-664  
08034 Barcelona  
[edicionespeninsula@planeta.es](mailto:edicionespeninsula@planeta.es)  
[www.edicionespeninsula.com](http://www.edicionespeninsula.com)

DAVID PABLO - fotocomposición  
ROMANYÀ-VALLS - impresión  
DEPÓSITO LEGAL: B-6.964-2017  
ISBN: 978-84-9942-605-1

# ÍNDICE

Introducción	11
PRIMERA PARTE: DE CÓMO LAS REDES FACILITAN Y CONDICIONAN EL DEBATE POLÍTICO, Y DE CÓMO INTENTAN RESITUARSE LOS MEDIOS	15
1. Votar en tiempos de redes	17
2. Datos que piensan	75
3. Medios, verdades y medias verdades	107
SEGUNDA PARTE: UN EXCURSO (APARENTE) PARA EXAMINAR EL PROBLEMA DESDE OTRO PUNTO DE VISTA	133
4. Cuestión de punto de vista	135
5. El rumor y su contexto	159
TERCERA PARTE: DE CÓMO ADQUIRIR EL SENTIDO DEL RUMOR	189
6. El sentido del rumor	191
Epílogo	221
Agradecimientos	223
Bibliografía	227

## VOTAR EN TIEMPOS DE REDES

Los tuits no mienten. Daniel White, periodista de la revista *Time*, dejó constancia de ello en un artículo después de comprobar que Twitter había predicho los resultados del *caucus* de Iowa mejor que cualquier otra red social. Ocurrió a principios de febrero de 2016, y las reuniones de los votantes de este pequeño Estado —de noventa y nueve condados y casi tres millones de habitantes— acababan de marcar el inicio de la batalla de las primarias entre republicanos y demócratas. Un año más.

Desde que en 1972 se convirtiera en la primera cita electoral para las primarias de los dos grandes partidos, Iowa ha establecido el punto de partida para un relato que acaba con la elección del nuevo presidente de Estados Unidos, en noviembre. Para los medios, los políticos y cualquier persona interesada en la política, los *caucus* de Iowa, estas asambleas políticas, son apasionantes.

Contaba White en su artículo que Twitter había sido un reflejo relativamente fiel de la realidad. En el bando republicano, la conversación digital durante la semana previa mostraba una tendencia al alza de Ted Cruz y Marco Rubio, al tiempo que Donald Trump —el favorito— iba de caída. Al fi-

nal, Cruz obtuvo una cuota del 31 % de los mensajes de Twitter; Trump se quedó en la segunda posición con el 30 %. En cambio, la conversación que tenía lugar en Facebook proporcionaba un retrato bien distinto: Trump acaparaba la mitad de la atención, mientras que Cruz se llevaba un modesto 23 %. El tercero en liza era Rand Paul, quien, con el 11 %, no apareció ni en Twitter ni en la votación real. El resultado, también en la vida real, es bien sabido: ganó Cruz con el 27,6 % de los votos, seguido de Trump, con el 24,3 %, y de Marco Rubio, con el 23,1 %.

Aparentemente, el pronóstico en las filas demócratas era menos complejo a pesar de lo enrevesado de su sistema de votaciones, porque se trataba de una elección a dos: Bernie Sanders o Hillary Clinton. En Facebook, Sanders arrasó con el 73 % de la conversación. Las distancias entre uno y otra en Twitter, en cambio, eran más ajustadas. Como el resultado final del *caucus*: 49,8 % para Hillary y 49,6 % para Sanders.

El sistema de votaciones en Iowa tiene poco de convencional, pero suele marcar la suerte de la mayoría de los aspirantes a la candidatura oficial de uno de los grandes partidos de Estados Unidos. Por cada voto obtenido, se invierten grandes sumas de dinero. Los medios de comunicación han convertido en tradición proporcionarle una cobertura incomprensible para quienquiera que busque proporcionalidad en relación con el número de votantes.

Además, no se trata de unos comicios en sentido estricto. En Iowa, por cada uno de los 1.681 distritos electorales, los ciudadanos registrados en los partidos se reúnen en bibliotecas, escuelas públicas o iglesias. Uno puede registrarse incluso el mismo día de la votación. La única condición es ser mayor de dieciocho años en noviembre de ese mismo año, cuando se celebran las elecciones presidenciales. Son lo más parecido a una reunión de la comunidad de vecinos, donde acude el administrador y todos pueden comentar con él los

problemas de convivencia y las reformas necesarias. Solo que aquí quien comparece es el aspirante a candidato oficial del partido, y el objeto de discusión es la política. Quien pretenda hacer pronósticos no cuenta con un censo cerrado con antelación.

Para complicarlo todavía más, cada partido utiliza un sistema de voto distinto. Los demócratas no cuentan los votos en una urna: optan por repartirse el espacio físico de la reunión y se agrupan según su candidato preferido: los del candidato A a la derecha, y los del candidato B a la izquierda. Los indecisos, en cambio, se quedan donde están, y crean el grupo de los *uncommitted*, «no comprometidos». A este posicionamiento sigue un debate para ganarse nuevos apoyos. Una vez transcurrido un tiempo prudencial, se cuenta a las personas de los grupos y gana el más numeroso. Si ningún grupo suma una proporción significativa de los asistentes, los partidarios de los candidatos con menos apoyos tienen la opción de pasar a engrosar las filas de uno de los principales. Pero aun así les queda la posibilidad de votar como *uncommitted*, alternativa que, de hecho, ganó en dos ocasiones: 1972 y 1976. Por su parte, los republicanos sí cuentan votos, y sus cifras nos permiten hacernos una idea de las dimensiones del acontecimiento. Ted Cruz, el candidato más votado entre los republicanos en 2016, obtuvo tan solo 51.666 votos: el 1,65 % de los ciudadanos de Iowa, el 0,02 % de los estadounidenses.

¿Cómo es posible que una plataforma como Twitter, en la que cualquiera puede opinar tantas veces como desee, donde abundan usuarios falsos y que en campaña electoral se convierte en objeto de una avalancha de mensajes procedentes de los equipos de comunicación de cada candidato, pronostique con acierto los resultados de algo tan complejo como los *caucus* de Iowa? La pregunta no alude a si es posible o no, porque ya se ha demostrado que sí; la pregunta se refiere a de qué modo puede inferirse esta información del galimatías en

el que tantas veces se convierte la actividad en el universo digital.

En pocas palabras, quizá sea porque Twitter funciona de la misma forma que un *caucus* celebrado en las aulas de un instituto de secundaria, pero sin votación final. El reto consiste en leer las señales que surgen de la conversación y saber interpretarlas correctamente.

Muchos nos hemos propuesto responder a este reto. Como otras compañías surgidas durante los últimos años, Sibillare había ido especializándose en el análisis de conversaciones digitales para mejorar la comunicación de empresas e instituciones e identificar y comprender tendencias sociales. Nuestra experiencia, pues, se enmarca en una carrera con muchos y muy buenos exponentes en distintos países. Lo que sigue es un intento de compartir parte de la experiencia.

Para ser sinceros, responder a la cuestión planteada —la de aprovechar esa misteriosa capacidad de la conversación de Twitter para reflejar o incluso predecir tendencias políticas— no estaba dentro de lo que considerábamos factible cuando iniciamos el análisis político de Twitter, casi un año antes de que David White publicara su artículo en *Time*. En unos meses de trabajo, no obstante, logramos extraer la siguiente conclusión: la conversación de Twitter sí refleja tendencias e incluso es válida para realizar estimaciones de voto.

La respuesta metodológica y técnica, sin embargo, no llegó de una sola vez. En mayo de 2015, con las elecciones municipales de Barcelona, aprendimos a reconocer tendencias. En diciembre del mismo año, gracias a un salto tecnológico cualitativo de la mano de Websays, nos atrevimos a dar por buenas estimaciones de voto en las elecciones generales del día 20, experimento que revalidamos con la repetición de las elecciones en junio de 2016. Entre una y otra convocatoria, los amigos de Websays, junto con un consorcio europeo, lograron predecir con un margen de error de apenas milésimas



el resultado del referéndum sobre la permanencia del Reino Unido en la Unión Europea, un hito histórico. Una respuesta, pues, gestada a golpe de elecciones. Y muy probablemente incompleta todavía, a la espera de testar y ajustar el sistema en nuevos comicios. Aún hay mucho que aprender de este inmenso conjunto de opiniones vertidas en internet por miles de personas a todas horas.

El primer paso lo dimos con ocasión de las elecciones municipales convocadas en España para el 24 de mayo de 2015. Por aquel entonces, llevábamos apenas medio año desarrollando la programación que posibilitaría el análisis de datos masivos.

El padre de la criatura, informáticamente hablando, claro, fue Carlos Bort. La plataforma que ideó Bort era lo más parecido a una máquina de palomitas virtuales. Se trataba de un programa informático en el que se cargaba una inmensidad de datos indigeribles —como granos de maíz— y al cabo de minutos generaba un bol de palomitas, o sea, archivos pdf y excel con la información exacta que se buscaba, esto es las respuestas a las preguntas que habíamos formulado.

Si bien el símil de las palomitas puede parecer frívolo, en realidad recoge la percepción compartida entonces por el equipo: el análisis de datos solo funcionaría cuando diera como resultado algo comestible y, de ser posible, sabroso. ¿De qué sirve efectuar grandes análisis, presentados con bonitas diapositivas de Power Point, si no se traducen en algo concreto que suponga una ventaja competitiva o un paso cualitativo hacia la comprensión de un problema?

Nuestra meta desde ese momento fue obtener resultados relevantes. No solo queríamos saber qué nos dicen los datos, sino qué nos sugieren y qué nos recomiendan. Carlos entrevistaba los datos con las preguntas que formulábamos a partir de necesidades concretas.

Para tener la programación a punto si queríamos producir

de forma escalable, estimamos un plazo idóneo de dos o tres años. En lugar de esperar todo ese tiempo, empezamos a producir análisis para clientes al quinto mes. Con sinceridad, esta decisión no vino motivada por ninguna actitud visionaria, sino por algo mucho más prosaico: nuestro presupuesto no daba para más. No participábamos en ningún programa de empresas tecnológicas con inversores que inyectaran líquido.

En definitiva, se trataba de hacer de la necesidad virtud. La opción de probar con fuego real la plataforma que estaba armando Carlos resultó especialmente acertada para comprender un principio básico en el negocio del análisis de datos: el valor añadido que aporta la tecnología es fundamental, pero nunca suficiente. Lo verdaderamente imprescindible es la capacidad de identificar el problema que los datos pueden ayudar a resolver, y la habilidad de formular las preguntas correctas a esos datos.

Una dificultad añadida era la naturaleza misma de los datos. Tratábamos, principalmente, con tuits y posts de páginas de internet. Recogíamos lo que miles de personas dejaban escrito sobre cualquier tema. Lo que corría por las redes. Rumores. El sentir de la gente, por decirlo de alguna manera. Un aluvión de material de calidad muy diversa, donde la sesuda opinión de un intelectual renombrado se mezclaba con la última ocurrencia de un quinceañero con vocación de humorista. Pocas veces podía preverse el volumen de la conversación, porque su elasticidad era altísima: una pequeña anécdota, la intervención de un político en un programa radiofónico o un evento climático imprevisto podían provocar aumentos exponenciales en el número de mensajes. Además, no siempre sabíamos desde dónde escribían sus autores estos, con lo que si se quería un análisis en una zona geográfica determinada, aquello pintaba difícil.

La mayoría de los datos contenían opiniones. El punto de partida de nuestro análisis, pues, debía ser el mismo que el de las

empresas de sondeos y estudios de mercado, que buscan averiguar lo que determinado grupo de gente piensa sobre una idea, una opción política o un producto.

En el ámbito político, el precedente eran los sondeos de intención de voto, de larga tradición en España. Este tipo de estudio se basa en la formulación de preguntas directas a una muestra representativa de la sociedad. El método de recopilación de las respuestas puede variar, pero la clave para su validez está en el diseño de la muestra y en la interpretación de los datos, que permiten obtener conclusiones globales.

En otras palabras, el valor añadido de un sondeo es que encuentra a las mil personas —por poner una cifra— cuyas respuestas sirven para explicar lo que opina el conjunto de la sociedad. Se trata de una información muy valiosa para los partidos que concurren a los comicios, y para los periódicos, que basan extensos reportajes en los resultados durante la campaña electoral.

Decidimos equiparar nuestros análisis de redes junto a los sondeos y, al mismo tiempo, diferenciarnos de ellos. Buscábamos lo mismo que esas encuestas, pero emprendiendo un camino distinto. Si bien ambos recogíamos la opinión política de la gente, nosotros no entrevistábamos a nadie: nos limitábamos a observar. Un sondeo sale a la calle y pregunta a los viandantes si tienen un momento para responder un cuestionario, o bien llama a un teléfono fijo con la esperanza de que alguien descuelgue y acepte atenderle durante unos minutos. Nosotros, en lugar de salir a la calle, entrábamos en Twitter, la cafetería de la plaza mayor, en términos digitales, y poníamos en marcha una grabadora para captar todo comentario relacionado con las elecciones. Cada método tiene sus ventajas y sus inconvenientes. Con la consulta directa, la información que se recoge está estrechamente relacionada con las preguntas cuya respuesta se pretende definir. Va directa al grano. Por el contrario, el cuestionario mismo del sondeo supone un ses-

go, en el sentido de que puede dejar fuera aquellas preocupaciones u opiniones que quizá existen pero nadie ha contemplado. Y al revés: las preguntas pueden provocar preocupaciones que no existían en la mente del encuestado hasta el momento en el que se plantean y llevarle a ofrecer opiniones o respuestas a algo frente a lo que tal vez se mostrara indiferente apenas unos segundos antes.

La observación de redes sociales mediante escucha, en cambio, puede captar mejor la espontaneidad y, de algún modo, recoger opiniones que son tabú o que no entran en el cuestionario. Por contra, el que grita más obtiene más atención, y teniendo presente que las opiniones son siempre las que se comparten públicamente en Twitter, nada impide que se produzcan cierta autocensura por parte de los individuos que no comparten la opinión mayoritaria de un grupo determinado y cierta reiteración por parte de los que dominan. Sobre esta cuestión, la generación de conversaciones homogéneas, abundaremos más adelante. Por no hablar de los profesionales de los partidos, que redifunden mensajes propagandísticos a un ritmo que, de tener que decirlos de palabra, los dejaría afónicos al quinto día de campaña. En resumen: en la pregunta directa, el encuestado representa una muestra; en la escucha sin intervención del analista, cada «encuestado» es una muestra.

Una ventaja clara del análisis de redes es que el registro puede ser constante: cabe obtener información diaria sin necesidad de salir a preguntar. La descripción de las tendencias puede ser más detallada, y el lapso entre la recogida de los datos y su análisis, acortarse a apenas unas horas.

En nuestra comparación, no obstante, surgieron dos objeciones de inmediato. La primera cuestionaba la validez de la muestra. ¿En qué medida resulta representativa una muestra de personas y de mensajes que se ha gestado de manera independiente, que es completamente abierta? Los sondeos se centran en la selección de las personas que van a responder

para garantizar la misma proporción de cada segmento (gente joven, jubilados, hombres y mujeres, etcétera) que en el conjunto de la sociedad. En nuestro caso, formaría parte de la muestra cualquier usuario de redes sociales que mencionara a alguno de los candidatos, con el agravante de que uno puede intervenir en distintos sitios sin que sepamos si el tal de Twitter es el mismo que uno de idéntico nombre en otra red social. Se trata, pues, de una muestra abierta. Participa quien quiera, y nada impide que una sola persona participe tantas veces como desee en cualquier conversación, e incluso que cree usuarios falsos con los que simular el apoyo de un grupo a un candidato, cuando en realidad no hay nadie más que él. Esta es, quizá, la mayor dificultad a la que nos enfrentábamos.

La segunda objeción que surgió giraba en torno al sentido de los mensajes: ¿cómo sabemos si lo que comentan en relación con un candidato es positivo o negativo? De entrada, un análisis de las palabras que utilizan los usuarios para expresarse o del tipo de enlaces que comparten nos daría pistas del sentido favorable o desfavorable de sus mensajes.

Visualizar las redes de relaciones de cada usuario sería asimismo otro modo de arrojar luz al asunto: dime con quién hablas y te diré cómo ha ido la conversación.

Hasta que despejáramos estas dos incógnitas, no ofreceríamos un sustituto de los sondeos de opinión en términos exactos. Parecía claro que el análisis de conversaciones en redes sociales tiene en común con las encuestas la capacidad de recoger opiniones políticas; ahora bien, en el estadio de experimentación en el que nos encontrábamos en mayo de 2015, el salto interpretativo que suponía este análisis era una herejía metodológica.

En cambio, no había ningún inconveniente en dar a conocer lo que analizábamos, si éramos capaces de dotar nuestro análisis de una forma completamente diferenciada de los sondeos. Así que decidimos ofrecer artículos de prensa sobre cómo

se desarrollaba la conversación en las redes sociales durante la campaña. En lugar de seudosondeos, brindamos periodismo de datos.

Para las elecciones de mayo, presentamos unas ideas muy concretas a varios medios. La propuesta consistía en realizar dos reportajes que explicaran, mediante gráficos y texto, la evolución de la conversación en Twitter sobre los candidatos a una alcaldía o un Gobierno autonómico. La base para los artículos serían todos los tuits en los que se mencionara a alguno de los candidatos, con independencia del lugar desde donde se emitieran, del tono —positivo o negativo— y de la influencia del usuario que participara en la conversación.

En las Islas Baleares, el periódico *Última Hora* nos pidió los datos y gráficos en exclusiva para la campaña de las elecciones autonómicas. En Barcelona, el diario digital *El Singular* publicó un análisis sobre la batalla por la alcaldía de Barcelona.

## TRIAS CONTRA COLAU

El proyecto más complejo fue el segundo: el que analizó la conversación en torno a los alcaldables para la ciudad de Barcelona.

Lo interesante de la campaña por la alcaldía de Barcelona es que los sondeos pronosticaban un ajustado enfrentamiento a dos entre Xavier Trias y Ada Colau. Trias, de CiU, centroderecha, se presentaba a la reelección. Colau no se había presentado hasta entonces: era una candidata surgida de los movimientos de defensa de los que no podían pagar la hipoteca tras el estallido de la crisis, agrupados en las Plataformas de Afectados por las Hipotecas y conocidos popularmente como «los antidesahucio». Algunos veían aquí un duelo entre la vieja política y la nueva, entre los partidos tradicionales y los movimientos sociales surgidos durante los años de la crisis.

Más allá de los programas y propuestas que defendían, por supuesto distintos, algo en lo que se diferenciaban unos y otros era en el uso que hacían de las redes sociales. La nueva política llevaba tiempo utilizando Twitter como herramienta de movilización y protesta, y solo recientemente habían tenido acceso a los canales habituales de comunicación de masas. Su activismo se veía animado y coordinado a través de los canales que ponía a su disposición la tecnología, ya fuera mediante redes sociales o mediante mensajería instantánea.

La política convencional en sentido comunicativo —encarnada por Xavier Trias— también estaba presente en las redes, pero de un modo más profesionalizado, con una base menor de personas movilizadas políticamente en este nuevo entorno. Contrariamente, contaba con un mayor acceso a los medios de comunicación convencionales.

Este fenómeno se dio, en mayor o menor medida, en otras contiendas locales durante aquellos comicios. Aparecieron en escena actores nuevos, por partida doble: nuevos por tratarse de caras desconocidas, lideradas desde instancias al margen de las estructuras tradicionales de partido, y nuevos porque eran «nativos digitales», en términos de organización.

Por consiguiente, si se observaba la conversación sobre aquellas elecciones en las redes sociales, el sesgo previsible era una sobredimensión de los partidarios de Ada Colau.

Y así fue. La aspirante alternativa a la alcaldía era la más mencionada. Su nombre era el centro de un gran debate, con usuarios muy activos. Estaba dominando la conversación digital de modo apabullante: durante la primera semana, 19.921 usuarios mencionaron a Colau 61.980 veces. Algunos lo hicieron en una sola ocasión y otros, incontables veces. Aquello suponía el 78 % de los usuarios y el 54 % de los mensajes que mencionaban a cualquiera de los candidatos. Muy por detrás se situaba Xavier Trias, el candidato que se presentaba a la reelección, con 6.282 usuarios y 20.301 tuits.

En tercera posición se hallaba el candidato de ERC, Alfred Bosch, con 2.916 usuarios y 10.368 tuits. Le seguían Jaume Collboni, del PSC (2.806 usuarios, 8.942 tuits); Carina Mejías, de Ciudadanos (2.178 usuarios, 5.478 tuits), y Alberto Fernández Díaz, del PP (1.554 usuarios, 7.006 tuits). A gran distancia quedó María José Lecha, la candidata de la CUP, a la que mencionaron 924 usuarios en 2.110 tuits.

Se daba la circunstancia de que los candidatos compartían una parte importante de la audiencia: a lo largo de la semana, la mayoría de los usuarios habían aludido a más de uno de los líderes políticos, ya fuera en el mismo tuit o en mensajes distintos.

Nos pareció interesante obtener información precisa acerca de las audiencias compartidas, es decir, quién compartía audiencia con quién y cuántos usuarios hablaban de uno de los cabezas de lista de modo exclusivo.

La presencia de Ada Colau estaba sobredimensionada en relación con voto real. Era algo que intuíamos sin lugar a dudas, pero que Twitter no corroboraba, al menos no en el nivel de análisis que habíamos realizado. Sabíamos que los movimientos sociales de los que procedía estaban utilizando Twitter para apoyar las campañas de sus candidatos tanto en Barcelona como en otras ciudades. Cabía barajar la hipótesis de que la estaban mencionando muchas personas en otros puntos de España para darle ánimos. Si éramos capaces de aislar esta parte de la conversación, quizá se salvaría la distancia entre ella y su perseguidor inmediato, Xavier Trias.

A pesar de esto, Twitter sí ofrecía un orden de notoriedad. Ada Colau era la más popular, Trias quedaba segundo, Alfred Bosch tercero, Jaume Collboni cuarto, Carina Mejías quinta, Alberto Fernández Díaz sexto y María José Lecha séptima.

Publicamos estas primeras cifras el día 15 de mayo, en un reportaje para un cliente —un medio de comunicación— que



titulamos «Ada Colau arrasa en Twitter en la primera semana de campaña».<sup>1</sup>

La semana siguiente ahondamos en los detalles. Para conocer mejor ese presunto sesgo que se generaba en la muestra por apoyos externos a Colau, consultamos los campos «ciudad» y «país» de todos los usuarios que participaban en la conversación. Se trata de una parte de la biografía que tiene toda cuenta de Twitter, pero que no es obligatorio rellenar correctamente ni exige ningún documento ni prueba que documente la veracidad de lo que se indica. Es, pues, en términos generales, una información cierta, pero mucho menos fiable que las coordenadas de los tuits cuyos usuarios han permitido que sean geolocalizados, algo que ocurre muy raramente.

Al mismo tiempo, identificamos a los usuarios más activos y sus hábitos de conversación. Aquellos, como es lógico, no eran propiedad de nadie, al menos en apariencia. Dentro de cada grupo se mezclaban tipos de personas y de organizaciones muy distintas.

Por otro lado, para afinar el tono de la conversación analizamos las palabras más repetidas en torno a cada candidato y los enlaces a noticias o vídeos que más se compartían.

Este segundo análisis se efectuaba no solo con los datos de la segunda semana, sino con la conversación de la semana precedente también. En total se recogieron 199.283 tuits, emitidos por 38.504 usuarios desde el inicio de la campaña electoral hasta el día 21 por la mañana, esto es, dos días antes de la jornada de reflexión. Con las cifras totales, Ada Colau seguía en cabeza, pero con una proporción menor de usuarios: había pasado del 78 % de la audiencia activa total al 72 %; y con respecto a los tuits, se hallaba presente en el 48,9 % de la conver-

1. [http://www.mon.cat/cat/notices/2015/05/la\\_batalla\\_per\\_l\\_alcaldia\\_de\\_barcelona\\_es\\_trasllada\\_a\\_twitter\\_108947.php](http://www.mon.cat/cat/notices/2015/05/la_batalla_per_l_alcaldia_de_barcelona_es_trasllada_a_twitter_108947.php)

sación, cinco puntos por debajo de la semana anterior. Además, entre los contenidos más compartidos había una noticia sobre los pagos que la plataforma de Colau había recibido del propio ayuntamiento por valor de 350.000 euros. Era la más popular, pero también despertaba controversia. Por otro lado, entre los conceptos más repetidos en los tuits se contaban los nombres de Pablo Iglesias y de Manuela Carmena, aliados políticos, y el hashtag #PrimaveraDemocratica, usado por diversas candidaturas que, en distintos puntos de España, sostenían la misma propuesta de alternativa política.

En lo referente al perfil de la muestra, detectamos que Ada Colau era la candidata cuya audiencia menos se presentaba como ubicada en Barcelona: solo el 13 %. La segunda localidad más citada era Madrid (6,2 %), seguida de Valencia (1,2 %) y Sevilla (0,9 %). Como ya he comentado, en Twitter no es obligatorio ni la geolocalización ni explicitar auténticamente el lugar de donde uno es natural. La única información de la que se dispone es aquella que los usuarios quieren compartir, y en los términos en los que ellos quieren compartirla. La forma de presentar la propia ciudad o país admite, pues, una variedad que daría para otro libro: lugares inventados, localizaciones inverosímiles... Para el caso, baste saber que un tercio de la muestra no proporcionaba dato alguno, ni siquiera falso. De todos los lugares que sí aparecían mencionados, la inmensa mayoría aglutinaban audiencias muy pequeñas.

Sea como fuere, un 13 % para Colau en Barcelona eran pocos, en comparación con el 24,66 % de Trias. Otra cosa, claro está, es que el 13 % de 27.988 usuarios son 3.638, mientras que la proporción de los que se ubican en Barcelona y mencionan a Trias significa, en términos absolutos, 2.385. Se acortaban las distancias, pero parece claro que entre el primero y el segundo era imposible un *sorpasso*.

## LOCALIZACIÓN EXPLÍCITA DE LOS USUARIOS EN TWITTER

Candidato	Audiencia activa	Dicen ser de Barcelona	
<b>Ada Colau</b>	27.988	13%	3.638
<b>Xavier Trias</b>	9.675	24,66%	2.385
<b>Carina Mejías</b>	5.284	18,19%	961
<b>Alfred Bosch</b>	4.716	26,55%	1.252
<b>Alberto Fernández Díaz</b>	4.425	17,90%	792
<b>Jaume Collboni</b>	4.191	26,39%	1.106
<b>María José Lecha</b>	1.855	24,85%	460

Esta ponderación no pareció concluir prácticamente nada, salvo la confirmación de que en la conversación sobre Ada Colau no solo participaban activamente detractores suyos, sino también aliados de fuera de Barcelona. Sin embargo, la procedencia de esas voces foráneas no podía cuantificarse más que en un porcentaje muy pequeño. Además, tampoco queríamos caer en el error de contar a los que no expresan ubicación en algún grupo en particular: ya fuera como usuarios procedentes de Barcelona o como foráneos. Si no lo dicen, no lo dicen. Hacer lo contrario supondría algo así como adaptar al mundo digital el clásico truco de atribuirse a los abstencionistas al que nos tienen acostumbrados algunos políticos. Ejercicios de estadística ficción, para otro momento.

Repetimos el análisis para cada uno de los candidatos, y el titular final fue «Colau domina, Trias aguanta y Mejías busca hacerse un hueco en Twitter».<sup>2</sup> La clasificación por volumen de audiencia quedaba así:

2. [http://www.mon.cat/cat/notices/2015/05/colau\\_domina\\_trias\\_aguanta\\_i\\_mejias\\_busca\\_fer\\_forat\\_109160.php](http://www.mon.cat/cat/notices/2015/05/colau_domina_trias_aguanta_i_mejias_busca_fer_forat_109160.php)

## EL SENTIDO DEL RUMOR

### USUARIOS QUE MENCIONARON A CADA CANDIDATO DURANTE LA CAMPAÑA

	Audiencia total	
<b>Ada Colau</b>	27.988	<b>1</b>
<b>Xavier Trias</b>	9.675	<b>2</b>
<b>Carina Mejías</b>	5.284	<b>3</b>
<b>Alfred Bosch</b>	4.716	<b>4</b>
<b>Alberto Fernández Díaz</b>	4.425	<b>5</b>
<b>Jaume Collboni</b>	4.191	<b>6</b>
<b>María José Lecha</b>	1.855	<b>7</b>

Quedaban por recoger los dos últimos días de campaña, que suelen ser especialmente intensos, pero con estos datos ya podíamos medir la evolución del peso de cada candidato de una semana a la otra.

En la banda baja de crecimiento, se percibía una ralentización en la campaña de Jaume Collboni, y mucho más en la de Ada Colau. Por arriba, se registraban dos crecimientos muy acentuados: la candidata de Ciudadanos, Carina Mejías, y el candidato del PP, Alberto Fernández Díaz. Estas variaciones podían deberse a un crecimiento orgánico o a la existencia de un contenido viral que disparase la audiencia, aunque durante un período de tiempo muy breve. Si nos hallábamos ante el primer caso, cabía interpretar los crecimientos como tendencia al alza de un candidato. Si, por el contrario, teníamos delante un viral circulando a todo gas, no había que tomárselo tan en serio.

En la conversación sobre Carina Mejías no se percibía nada especialmente escandaloso. Circulaban enlaces a dos noticias de medios convencionales y un tuit que había emitido ella misma años atrás. (Bastante gracioso, por cierto.)

Alberto Fernández Díaz, en cambio, era protagonista de un mensaje viral. El usuario @hibai\_ publicó, el 14 de mayo a las 16.04, una fotografía de un panfleto electoral del PP sobre

EVOLUCIÓN DE LA CAMPAÑA ELECTORAL

	15 de mayo		21 de mayo		Variaciones	
	Audiencia	Tuits	Audiencia	Tuits	Crecimiento audiencia	Crecimiento tuits
<b>Ada Colau</b>	19.921	61.980	27.988	97.465	40,49%	57,25%
<b>Xavier Trias</b>	6.282	20.301	9.675	37.597	54,01%	85,20%
<b>Alfred Bosch</b>	2.916	10.368	4.716	18.378	61,73%	77,26%
<b>Jaume Collboni</b>	2.806	8.942	4.191	15.794	49,36%	76,63%
<b>Carina Mejías</b>	2.178	5.478	5.284	15.616	142,61%	185,07%
<b>Alberto Fernández Díaz</b>	1.554	7.006	4.425	15.789	184,75%	125,36%
<b>María José Lecha</b>	924	2.110	1.855	4.829	100,76%	128,86%

inmigración, con el siguiente comentario: «Se puede ser mezquino, repugnante, un racista de mierda y luego ya viene @albertofdezxbc, que es todo lo anterior». En el momento de recoger los datos, este tuit ya aparecía en más de un millar de registros, y en el conjunto de la conversación las palabras «mierda», «racista», «repugnante» y «mezquino» se contaban entre las seis primeras.

Así pues, podíamos confirmar que Carina Mejías era la candidata con tendencia al alza, tanto como para aspirar a la tercera posición. Alfred Bosch y Jaume Collboni estaban separados por muy pocos usuarios, pero la tendencia favorecía al primero.

Con respecto a Colau y a Trias, la segunda semana confirmaba el dominio de la candidata alternativa, aunque nos extrañaba la enorme ventaja que, en cuanto a volumen, sacaba a su contrincante. Es un fenómeno al que hemos vuelto en más ocasiones y para el que aún no tenemos una explicación irrefutable. La hipótesis es que se genera un fenómeno de refuerzo de la parte del debate que se percibe como mayoritario —en este caso, Ada Colau— ya sea por el silencio de los que discrepan, ya sea porque se ha impuesto un encuadramiento que sitúa a una de las partes en un papel preponderante.

Si hubiéramos querido hacer una estimación de voto, algo que entonces se nos antojaba demasiado arriesgado, con la mera observación de la conversación que habíamos recogido (199.283 tuits, 38.504 usuarios) habríamos afirmado:

- Que Ada Colau ganaba las elecciones.
- Que Xavier Trias quedaba segundo.
- Que Carina Mejías sería una sorpresa y podía quedar tercera.

- Que Alfred Bosch y Jaume Collboni quedarían muy cerca uno del otro, pero acabaría imponiéndose el primero, porque su tendencia de crecimiento era mucho mayor. Y que Alberto Fernández Díaz, aunque superara a Jaume Collboni en

usuarios de Twitter, no obtendría un resultado superior, porque el motivo del crecimiento era un fenómeno viral que, en cualquier caso, no podía atribuirse a posibles votantes suyos.

ESTIMACIÓN PROVISIONAL DE RESULTADOS

	Posición por volumen de usuarios en Twitter en fecha 21/5	Sentido de la tendencia
<b>Ada Colau (BCNC)</b>	1	Se confirma su dominio
<b>Xavier Trias (CiU)</b>	2	Segunda posición asegurada
<b>Carina Mejías (C's)</b>	3	Tercera posición confirmada
<b>Alfred Bosch (ERC)</b>	4	Muy ajustado con PSC
<b>Alberto Fernández Díaz (PP)</b>	5	Seguramente menor, por sesgo detectado
<b>Jaume Collboni (PSC)</b>	6	Probable adelanto por sesgo en PP
<b>María José Lecha (CUP)</b>	7	Demasiado lejos del resto. Quedará última

De ser cierta nuestra identificación de tendencias, los resultados nos darían una caída del candidato del PP a la sexta posición, en beneficio del candidato del PSC. De ninguna manera podíamos afirmar que el apoyo popular sería proporcionalmente similar al número de usuarios que hablaban de cada candidato: demasiado ruido entre los que criticaban y los que participaban en la conversación pero no podrían participar en la votación porque no eran de Barcelona. Por otro lado, nuestros datos solo alcanzaban hasta el día 21 —cuando debíamos entregar el análisis al medio de comunicación que lo publicaba—, de modo que no reflejaban la recta final de la campaña, que fue muy intensa.

Seguimos muy atentamente las encuestas que fueron apareciendo, pendientes de si esta simple observación tenía sentido o no. Los principales sondeos publicados ofrecían el orden de

COMPARATIVA DE ENCUESTAS Y ESTIMACIÓN VÍA TWITTER

	CIS abril	La Vanguardia 3/5	El Periódico 11/5	ABC 12/5	El Mundo 16/5	El País 16/5	Twitter (hasta 21/5)	Resultado electoral	Aciertos
<b>Ada Colau</b>	1	2	2	2	2	1	1	1	3 de 7
<b>Xavier Trias</b>	2	1	1	1	1	2	2	2	3 de 7
<b>Carina Mejías</b>	3	3	5	3	3	3	3	3	6 de 7
<b>Alfred Bosch</b>	5	5	6	4	6	4	4	4	3 de 7
<b>Alberto Fernández Díaz</b>	6	6	4	6	5	7	6	6	4 de 7
<b>Jaume Collboni</b>	4	4	3	5	4	6	5	5	2 de 7
<b>María José Lecha</b>	7	7	7	-	7	5	7	7	5 de 7
<b>aciertos</b>	5 de 7	3 de 7	1 de 7	4 de 7	2 de 7	4 de 7	7 de 7		



candidatos por apoyo popular que puede verse en la tabla anterior. En la penúltima columna figuran las tendencias reconocidas en Twitter, y en la última, los resultados finalmente registrados en las urnas.

Sorprendente, ¿no? La tendencia que identificábamos en las redes parecía guardar relación con la que se registraba en las preferencias de los votantes de la ciudad. Las encuestas y los sondeos, contra lo que lleva a pensar el sentir práctico de lectores y muchos periodistas y opinadores, no nos anticipan lo que ocurrirá en las urnas. Ni siquiera nos cuentan lo que está ocurriendo, si no más bien lo que ocurría en el momento en que se formularon las preguntas a los entrevistados. Sí pueden, por el contrario, indicar una tendencia. Dados los aciertos y los errores en las estimaciones, la tendencia que seguía la conversación en Twitter respondía mejor al resultado final que las encuestas según el método tradicional.

Los resultados, en número de votos, fueron los siguientes:

RESULTADOS ELECTORALES EN BARCELONA (MUNICIPALES)

	votos
<b>Ada Colau</b>	176.337
<b>Xavier Trias</b>	158.928
<b>Carina Mejías</b>	77.279
<b>Alfred Bosch</b>	76.988
<b>Jaume Collboni</b>	67.380
<b>Alberto Fernández Díaz</b>	60.887
<b>María José Lecha</b>	51.889

Twitter no mentía en cuanto al orden de preferencias de los votantes, como tampoco mentiría en el *caucus* de Iowa, diez meses más tarde.

La distancia entre Ada Colau y Xavier Trias fue mucho menor en las urnas que en las redes. En el caso de la última clasi-

ficada, María José Lecha, el volumen de conversación en redes quedó a una distancia mucho mayor del candidato inmediatamente superior de lo que, en número de votos, resultó al final. Las redes sobredimensionaban por un lado, en el liderazgo de la conversación, y castigaban por el otro, en el último de la fila. Al menos en este caso.

Si se analizaban las diferencias entre los candidatos del centro de la tabla (excluyendo a Alberto Fernández Díaz, por el sesgo ya mencionado), Twitter indicaba la proximidad entre tres opciones: Carina Mejías, Alfred Bosch y Jaume Collboni. Por ese orden, el mismo que en las urnas.

Nuestra escucha de redes y análisis se amplió unos días más tarde, cuando tomamos el conjunto de datos de toda la campaña. En los días que faltaban para el segundo artículo periodístico, la tendencia al alza de Alfred Bosch continuó hasta llegar a un empate técnico con Carina Mejías: esta obtenía 6.090 usuarios, y Bosch la superaba con 6.097. Estarían muy pegaditos, como ocurrió en las urnas. Collboni, por su parte, se rezagó ante el empuje de los dos anteriores, y si el día 21 quedaba separado del candidato de ERC por 525 usuarios, en los últimos tres días la diferencia casi se duplicaba, hasta alcanzar 1.017. Alberto Fernández Díaz, que le sacaba 234 usuarios de ventaja a Jaume Collboni, cerraba la campaña con solo 89 de ventaja: se confirmaba que el del PSC adelantaría al del PP.

En la parte alta, constatamos el efecto de sobredimensión para la opción que obtiene el liderazgo. La victoria de Ada Colau iba viéndose paulatinamente posible entre sus partidarios y la opinión pública, y la movilización fue aumentando de manera exponencial. Este fenómeno se ha producido en otras contiendas, pero no me atrevería a señalarlo como patrón seguro: quizá sea matizable según el punto desde el cual parte cada opción, es decir, si se presenta desde la oposición o desde el Gobierno saliente.