

Twitter, una sonda permanente de opinión y una vía para canalizar acciones ciudadanas  
*Documento para su presentación en el IV Congreso Internacional en Gobierno,  
Administración y Políticas Públicas GIGAPP-IUIOG. Instituto Nacional de  
Administración Pública (Madrid, España) 23 y 24 de septiembre 2013.*

Congosto Martínez, M. Luz. Grupo de investigación WebTlab, Departamento de  
Telemática Universidad Carlos III  
Email: [mcongost@inv.it.uc3m.es](mailto:mcongost@inv.it.uc3m.es)

**Resumen/abstract:** Twitter proporciona un caudal público de información, recogiendo opiniones totalmente espontáneas que reflejan el pulso de la sociedad. El volumen de datos supera con creces al de cualquier sondeo y el muestreo es continuo. Sin embargo, la información es desestructurada, dispersa y difícil de segmentar. El objetivo de esta investigación es validar un método que permita extraer la información más relevante dentro de los múltiples debates sociales que se están desarrollando en Twitter para comprender cómo se está usando esta red social. Esta relevancia se ha medido tanto por la difusión de contenidos como por la persistencia de publicación de los usuarios. Este método es un primer paso hacia una metodología más sofisticada que permita complementar a las actuales encuestas que realiza el CIS para el barómetro social

**Palabras clave:** Twitter, análisis de redes sociales, metodología, relevancia, demografías

**Nota biográfica:** Licenciada en Informática por la UPM de Madrid y Máster en Telemática por la Universidad Carlos III. Realiza el doctorado en la Universidad Carlos III en el Departamento de Telemática sobre “Caracterización de usuarios y propagación de mensajes”. Ha sido profesora asociada y trabajado en grandes proyectos de I+D de las telecomunicaciones (TESYS, MORE, INFOVÍA, EOC..) durante más de 25 años.

[Comentarios]

## Introducción

La generalización del uso de las redes sociales y los smartphones por amplios segmentos de la sociedad ha contribuido que en pocos años los ciudadanos hayan transformado la manera de consumir información, de difundirla y de interactuar con otros. Las personas están más horas conectadas a la Red, la información fluye en tiempo real y la capacidad de reacción ante los eventos es cada vez más rápida.

La participación masiva de personas en la Red es consecuencia del nacimiento de la Web 2.0 en 2004. Hasta esa fecha estaba ocupada por Webs profesionales, el entorno académico y Webs personales de apasionados de la tecnología. La aparición de blogs rompió las barreras tecnológicas que impedían que personas sin bagaje tecnológico pudieran crear su propia Web y publicar sus contenidos, siendo la unidad de publicación el “post”, del tamaño de una página. Sin embargo, el mundo de los blogs era aún minoritario porque requería de una dedicación importante de tiempo y un esfuerzo en la creación de contenidos propios. El nacimiento de las redes sociales fue parejo por esos tiempos. Twitter (2006), Facebook (2005), LinkedIn (2003) pero la participación seguía siendo minoritaria hasta el 2009, cuando comenzó el despegue de su utilización. Actualmente las redes sociales han adquirido tal popularidad que han roto la tradicional brecha de acceso a la Red por género y acortado la de edad según muestra un estudio de realizado por Pew Internet (2012).<sup>1</sup>

Las redes sociales aportan tres características nuevas a los medios tradicionales de publicación (artículos/ post):

- Fragmentación del mensaje (mensajes cortos y espaciados en el tiempo)
- Dedicación intermitente (ocupación de espacios de tiempo libre)
- Propagación de los mensajes (difusión de mensajes de otros).

Las redes sociales son una fuente de datos valiosísima para las Ciencias Sociales y al alcance de los investigadores. Disponen de APIs (Interfaz de Programación de Aplicaciones) con la que se pueden recabar los datos. Se requieren unos conocimientos de programación medios para obtenerlos. También existen muchos servicios de Internet que ofrecen información ya elaborada.

---

<sup>1</sup> [http://www.pewinternet.org/~media/Files/Reports/2013/PIP\\_SocialMediaUsers.pdf](http://www.pewinternet.org/~media/Files/Reports/2013/PIP_SocialMediaUsers.pdf)

Actualmente este torrente de información no se está aprovechando en todo su potencial porque existen las siguientes barreras:

- Cantidades masivas de información
- Mezcla de relevancia con irrelevancia
- Mensajes desestructurados
- Fuentes no contrastadas

De todas las redes sociales merece un especial interés para su estudio Twitter por su naturaleza pública, lo que permite que cualquier mensaje emitido por un usuario en esta red pueda ser potencialmente leído por los demás. Los mensajes de Twitter, llamados *tweets*, se caracterizan por su brevedad siendo su tamaño máximo de 140 caracteres, lo que obliga a la concisión en la escritura. Esta característica ha favorecido su utilización para la difusión de noticias, opiniones, quejas, pensamientos, y un cauce para la ironía. Por otra parte, la posibilidad de etiquetar los mensajes con *hashtags* (palabra precedida por el símbolo #) le permite agrupar mensajes dentro de un contexto.

Twitter proporciona un caudal público de información, recogiendo opiniones totalmente espontáneas que reflejan el pulso de la sociedad. El volumen de datos supera con creces al de cualquier sondeo y el muestreo es continuo. Sin embargo, la información es desestructurada, dispersa y difícil de segmentar. El objetivo de esta investigación es validar un método que permita extraer la información más relevante dentro de los múltiples debates sociales que se están desarrollando en Twitter para comprender mejor cómo se está usando esta red social. Esta relevancia se ha medido tanto por la difusión de contenidos como por la persistencia de publicación de los usuarios.

El método se ha aplicado a debates de larga duración como los Recortes, el Rescate o el caso Bankia y otros más recientes como el llamado caso Bárcenas o la ley de Transparencia. Se han obtenido interesantes conclusiones de los datos analizados y pensamos que puede aplicarse para medir el pulso social de la sociedad española.

### **Marco teórico**

De la inmensa información que circula por Internet una pequeña parte, aunque no por ello menos importante, es de naturaleza social o política. Poder conocer el pulso social por las aportaciones de los ciudadanos a la Red es un desafío y una oportunidad para los investigadores en Ciencias Sociales y Comunicación.

El uso de Internet relacionado con la política ha sido ampliamente estudiado aunque generalmente se ha enfocado hacia el análisis de las campañas electorales por ser periodos especialmente intensos de actividad en todos los medios de publicación de la Red. Los foros, los blogs y las redes sociales se convierten en cauces de debates políticos durante esas épocas.

A pesar de una abundante oferta de información en la Red, en España el porcentaje de penetración de Internet como fuente de información política en periodo electoral es tan sólo de un 20%, muy inferior a Estados Unidos y Europa. Los más propensos a informarse por este medio son los que tienen mayores recursos (hombres con mayor nivel de educación, habilidad en el uso de Internet y frecuente lectura de la prensa) o los de mayor motivación política (interés por un partido o por asuntos políticos) (Anduiza et al., 2012). Por otro lado, el uso de Internet está más relacionado con un mayor interés político y un incremento de la eficacia política interna (la creencia de que se puede participar eficazmente en la política), aunque no ocurre lo mismo con la eficacia de la política externa (la percepción que tiene el ciudadano de que el gobierno responde a sus necesidades) (Colombo et al., 2012). La actividad frecuente en Internet se da entre ciudadanos más críticos y en los que la probabilidad de participar en movilizaciones sociales no institucionales es mayor (Cantijoch, 2009).

No todos los usuarios participan de la misma manera, ni tienen las mismas motivaciones, ni el grado de compromiso ni la vinculación con los partidos políticos. Desde una participación débil como retransmisores de mensajes hasta un compromiso social o político, existen otros grados de participación intermedios en los que predomina alguna faceta como el humor, la controversia o la opinión (Sampedro et al., 2013).

Pero cualquier tipo de participación es susceptible de adquirir relevancia, incluso el llamado *clicktivism* (participación de tan poco esfuerzo como realizar un clic: firmar una petición, pulsar un “me gusta” o “hacer RT”) en su vertiente colectiva adquiere una preeminencia como emisor cuando concentra un gran número de clics hacia un único receptor (Peña-López, 2013-b).

Centrando el estudio en Twitter, también se ha encontrado que una parte importante de las publicaciones sobre su utilización en temas políticos se enfoca al análisis de campañas electorales, en este caso para predicción de resultados. Muchos artículos surgieron como reacción al análisis de las elecciones alemanas de Tumasjan (Tumasjan et al., 2010) que afirmaba que era posible predecir los resultados de una forma tan sencilla como contar el número de menciones de los partidos en Twitter. Sus afirmaciones fueron rebatidas por su compatriota Jungherr que replicó el experimento incluyendo otros partidos políticos (Jungherr. et al., 2011). Otros autores también intentaron aplicar su método sin éxito en otros países como USA 2010 (Metaxas et al., 2011) (Gayo-Avello, 2011), Holanda (Tjong, 2012), Irlanda (Bermingham et al., 2011) y se mostraron muy escépticos con este método con tan poca base científica.

En España se han estudiado las elecciones al parlamento celebradas en noviembre de 2011 (20N) desde la participación de los ciudadanos en Twitter. Dos investigaciones independientes y con distintos enfoques llegan a las mismas conclusiones. La primera de ellas (Barberá et al., 2012) descubrió que las demografías no son uniformes, el perfil más participativo fue hombre, residente en zona urbana y con unas preferencias ideológicas intensas. También se observó que ciertos partidos estaban sobre-representados y que la estrategia del Partido Popular fue más jerarquizada usando lemas de campaña (*#hashtag*) frente a la del PSOE que fue más horizontal e interactiva. El segundo estudio (Congosto et al., 2012) aplicando el método de Tumasjan obtuvo unos resultados bastante ajustados aunque un análisis más profundo descubrió que los resultados diferían según se constasen menciones o aparición de *hashtags* de campaña y por tanto los resultados eran sensibles al tipo de estrategia de comunicación del partido. Al igual que en el primer estudio se encontraron partidos sobre-representados y un análisis de la polaridad de los usuarios detectó un “voto oculto” deducido de la diferencia entre la orientación política encontrada y los resultados. En el caso de los votantes de izquierda se refuerza esta afirmación al encontrarse una descompensación entre el alto porcentaje de noticias de prensa de medios de izquierda frente a la polaridad de izquierda detectada. Asimismo, otro análisis de las elecciones andaluzas del 2012 (Detell et al., 2013) consiguió una predicción en los dos partidos mayoritarios más ajustada que la obtenida en las encuestas electorales. Sin embargo, el resultado para

los partidos más pequeños fue menos preciso debido nuevamente a la sobre-representación.

En periodos no electorales también está latente esa participación en Twitter en temas políticos y sociales. Las movilizaciones, los sucesos que han causado una alarma social por la situación económica o por los escándalos de corrupción, asimismo han intensificado la actividad de los ciudadanos en la Red. Estos eventos están menos estudiados por su diversificación y porque que entrañan una serie de dificultades como una extensa duración, una cantidad ingente de datos y una falta de metodologías para abordar este problema. Por otro lado, la forma de medir cómo la ciudadanía está participando en las movilizaciones es insuficiente, tanto por el tipo de preguntas realizadas como por la limitación las opciones de respuestas (Fernández-Prados, 2012).

El uso de *hashtags* ha sido una de las piezas clave para hacer de Twitter una herramienta de movilización social. La primera vez que se utilizó uno de ellos fue en diciembre de 2009, para organizar una protesta social contra la Ley de Economía Sostenible, conocida popularmente como la Ley Sinde. La causa de la polémica fue una disposición final relativa a la regulación de Webs y la protección de la propiedad intelectual. Se utilizó el *hashtag* #manifiesto que en tres meses aglutinó a 37.131 *tweets* publicados por 6.449 usuarios, según datos del blog Barriblog.com<sup>2</sup>, una cantidad muy significativa teniendo en cuenta el desarrollo de Twitter en esas fechas. En el año 2010 el uso de Twitter aumentó espectacularmente, con cien millones de usuarios nuevos<sup>3</sup>, crecimiento que se ha mantenido hasta la fecha con más de quinientos millones de usuarios activos. Durante el año 2011, el uso de Twitter se globalizó como herramienta de movilización en la Primavera Árabe, el 15M en España y Occupy Wall Street en U.S.A. Actualmente, Twitter forma parte de la logística de cualquier convocatoria y por tanto es una fuente imprescindible para conocer las inquietudes sociales.

El caso del 15M, objeto de muchas publicaciones, marcó un antes y un después en las movilizaciones a través de las redes sociales. Por primera vez en España una convocatoria desde la Red consiguió movilizar a un público no asiduo a

---

<sup>2</sup><http://www.barriblog.com/index.php/2010/04/05/analizando-el-tag-manifiesto/>

<sup>3</sup><https://blog.twitter.com/2010/meaningful-growth>

manifestaciones compuesto en su mayoría por jóvenes con estudios, muchos de ellos mujeres y desempleados. Una comparación del 15M con otras manifestaciones que tuvieron lugar durante el 2011 muestra claramente la diferencia en tres aspectos: las organizaciones convocantes, el proceso de movilización y las características de los participantes (Anduiza, E 2012).

La complejidad que entraña extraer conclusiones de la movilización social en Twitter requiere de un análisis multidisciplinar. En el caso del 15M el grupo @Dataanalysis15m lo abordó combinando la experiencia interna de los participantes en el movimiento con el análisis de redes y datos. La investigación se desarrolló en cinco ejes: el histórico-experiencial, el filosófico-sociológico, el analógico, el de la matemática de los sistemas complejos y el empírico (Toret et al, 2013)

La reacción en Twitter a las convocatorias es muy dinámica, cada una de ellas es capaz de movilizar a un público nuevo o crear nuevas plataformas ciudadanas. El descontento existe y es posible canalizarlo por distintas llamadas. En el transcurso de poco más de un año, las demografías detectadas en el 25S (convocatoria de “Rodea el Congreso”) se transformaron respecto a las del 15M: aumentó el número de usuarios de zonas no urbanas, disminuyó el número de profesionales a la vez que se incrementó el de estudiantes, se redujo la diferencia entre hombres y mujeres y creció la participación de los partidos políticos y sindicatos (Peña-López et al., 2013-a). Este dinamismo hace que nunca se tenga una foto fija de lo que está sucediendo, ya que los usuarios de Twitter van y vienen. Por ese motivo es imprescindible conocer la evolución para tener una referencia.

### **Metodología**

La metodología que se propone está basada en un muestreo continuo en Twitter de distintos eventos sociales de los que se extraen pautas de comportamiento y demografías de usuarios que pueden facilitar el análisis de nuevos sucesos y servir de referencia.

Es una metodología fundamentalmente cuantitativa aunque en algunos casos ha sido necesario un tratamiento manual. El proceso que se ha seguido consta de tres partes:

1. Captura de los *tweets* de distintos eventos
2. Análisis de los picos de actividad de cada evento
  - a. Obtención de los mensajes más propagados
  - b. Tipificación de los usuarios
  - c. Clasificación de los *tweets* más difundidos
3. Selección de los usuarios persistentes
  - a. Demografías: Género, localidad y profesión
4. Uso del grupo de usuarios persistentes para analizar otros *datasets*

### ***Captura de los tweets de distintos eventos***

Se han obtenido los datos mediante el Streaming API de Twitter que proporciona los *tweets* con una completa información de contexto que incluye el perfil del autor en el instante de la emisión. Para cada *tweet* se han seleccionado los siguientes datos:

- **Date-time:** fecha emisión
- **Text:** texto del mensaje
- **id\_user:** identidad del autor
- **screen\_name:** apodo del autor
- **name:** nombre del autor (dato no obligatorio)
- **followers\_count:** número de seguidores al emitir el *tweet*
- **friends\_count,** número de followers al enviar el *tweet*
- **statuses\_count:** número de *tweets* publicados hasta el momento
- **location:** localización declarada del autor (dato no obligatorio)
- **description:** bio del autor (dato no obligatorio)

Por este procedimiento se han obtenido los siguientes datasets

**Tabla 1. Dataset utilizados**

<b>Data set</b>	<b>Fecha recogida</b>	<b>Tamaño de la muestra</b>	<b>Palabras clave</b>
-----------------	-----------------------	-----------------------------	-----------------------

<b>Data set</b>	<b>Fecha recogida</b>	<b>Tamaño de la muestra</b>	<b>Palabras clave</b>
Recortes	17-02-2012 al 9-08-2013	2.757.060 <i>tweets</i> 591.941 usuarios	#19F, #contrareforma, #reformalaboral, #19Ftomalacalle, #razones19F, #huelgageneral, #29F, #11M, #29M, #29NoaLaHuelgaGeneral, #noalahuelgageneral, #pideunpiquete, #piqueteempresarial, #huelgadespierta, recortes, reforma laboral, retallades,#15S,#14N
Bankia	7-05-2012 Hasta 14-8-2013	2.617.257 <i>tweets</i> 538.301 usuarios	#bankia, Rodrigo Rato y Caja Madrid
Rescate	23-07-2012 Hasta 14-8-2013	1.713.586 <i>tweets</i> 651.823 usuarios	prima de riesgo, rescate
Bárcenas	18-01-2013 Hasta 814-08-2013	2.643.483 <i>tweets</i> 526.317 usuarios	#lospapelesdebárcenas, #barcenas, Bárcenas, Barcenas, corrupción del pp, Gürtel, Gurtel.
Ley de transparencia	30-05-2013 Hasta 14-08-2013	75.354 <i>tweets</i> 37.953 usuarios	Ley de transparencia, #LeyDeTransparencia, #LaLeyDeTransparencia, #transparencia.

### ***Análisis de los picos de actividad de cada evento***

Se han agrupado los *tweets* por días, calculando el número de mensajes y de usuarios únicos de cada uno de ellos, seleccionando las fechas de mayor actividad para su posterior análisis.

Analizar el contenido del gran volumen de *tweets* de cada una de estas reacciones sólo sería abordable por procedimientos de clasificación automáticos, sin embargo la tecnología para poder realizarlo de una manera precisa no está aún disponible, por este motivo solo se ha examinado una muestra de forma manual. No se ha tomado una muestra aleatoria sino los cuarenta mensajes más difundidos de cada uno de los picos de reacción. Los mensajes en Twitter se difunden mediante el mecanismo de retransmisión de *tweets* (RT). En estudios de campañas electorales en Twitter se ha detectado la polaridad política de los ciudadanos analizando el uso del RT debido a que se utiliza

generalmente cuando se está de acuerdo con el contenido del mensaje (Conover, 2011). Por este motivo los mensajes más propagados desvelan un sentir mayoritario y proporcionan una visión de la relevancia en el conjunto de todos los mensajes emitidos. Se ha definido una catalogación propia estableciendo los siguientes tipos de tweets:

- Foto-vídeo, cuando contiene una imagen multimedia o un enlace a un contenido multimedia
- Noticia, si incluye una noticia o un enlace a una noticia de un medio
- Opinión, en el caso de que manifieste una opinión o un enlace a un artículo de opinión
- Ironía, si esconden dobles significados o la intención que va más allá del significado más simple o evidente
- Movilización, cuando incitan a alguna acción determinada
- Otros, para los que no pertenecen a ninguno de los grupos anteriores

Para la clasificación de usuarios se ha utilizado una tipificación sencilla pero muy clarificadora (Osteso et al., 2013) de tres tipos de usuarios:

- Personales: creados por personas físicas que corresponden a una identidad real
- Institucionales: asociados a cualquier tipo de institución (Partidos políticos, sindicatos, medios, Administración Pública, plataformas ciudadanas y asociaciones de todo tipo)
- Personajes: de identidad virtual al amparo del anonimato. Dentro de este grupo se incluyen las parodias o *fakes*

### ***Selección de los usuarios persistentes***

Al seleccionar a los usuarios que participan en distintas líneas de discusión se obtiene una muestra de los usuarios que siguen de cerca lo que ocurre en su entorno y por tanto más propicios a difundir o comentar sucesos de impacto social o a participar en movilizaciones. A este grupo de usuarios se les ha denominado persistente.

El grupo persistente es minoritario debido a que la participación de los usuarios en Twitter sigue una ley de potencias donde la mayoría de los *tweets* publicados provienen una minoría (Cheng et al., 2009). Lo mismo sucede con la perseverancia, solo una minoría participa de una forma continuada, como se ha podido constatar en la participación de las convocatorias del 15M y del 25S (Peña-López et al., 2013-a).

Para el grupo persistente se han obtenido, en la medida de lo posible, datos demográficos como género, profesión y localidad. Debido a la escasa información de perfil en Twitter, los atributos para caracterizar a los usuarios han tenido que deducirse indirectamente (Mislove et al., 2011), estando en algunos casos incompletos. Cada usuario se ha caracterizado con los siguientes atributos mediante algoritmos propios desarrollados en la herramienta *t\_hoarder*<sup>4</sup>.

- **Género:** deducido de dato “name” [no obligatorio]. Recurso utilizado: INE, tabla de nombres Hombres/mujeres por edad media<sup>5</sup>.
- **Localización:** deducido del dato “location” [no obligatorio], estructurándolo en autonomía, provincia y localidad. Recurso utilizado: INE, tabla de municipios y provincias<sup>6</sup>
- **Profesión:** deducido del dato “description” [no obligatorio]. Utilizando la clasificación de sectores económicos de Wikipedia<sup>7</sup>.

#### *Uso del grupo persistente para analizar otros datasets*

Uno de los problemas para generar la red de seguidores de un conjunto de usuarios Twitter es el *ratelimit*<sup>8</sup>, un límite temporal de accesos a la API. A partir de cierta cantidad de usuarios el tiempo para obtener la red se alarga de tal manera que es imposible de conseguir. Muchos de los usuarios de un *dataset* son usuarios pocos activos (Ley de potencias) y algunos de ellos falsos positivos. Cualquier método que descartase ambos tipos de usuarios facilitaría el proceso de obtención de la red al disminuir su tamaño, además de obtener un conjunto de usuarios relevantes.

Se ha aplicado el grupo persistente para segregar de un *dataset* los usuarios más persistentes de forma que sea posible obtener su red y las comunidades que forman.

## **Resultados**

### *I- Análisis de los picos de actividad de cada evento*

---

<sup>4</sup><http://t-hoarder.com/>

<sup>5</sup>[http://www.ine.es/daco/daco42/nombyapel/nombres\\_por\\_edad\\_media.xls](http://www.ine.es/daco/daco42/nombyapel/nombres_por_edad_media.xls)

<sup>6</sup><http://www.ine.es/jaxi/menu.do?type=pcaxis&path=/t20/e245/codmun&file=inebase>

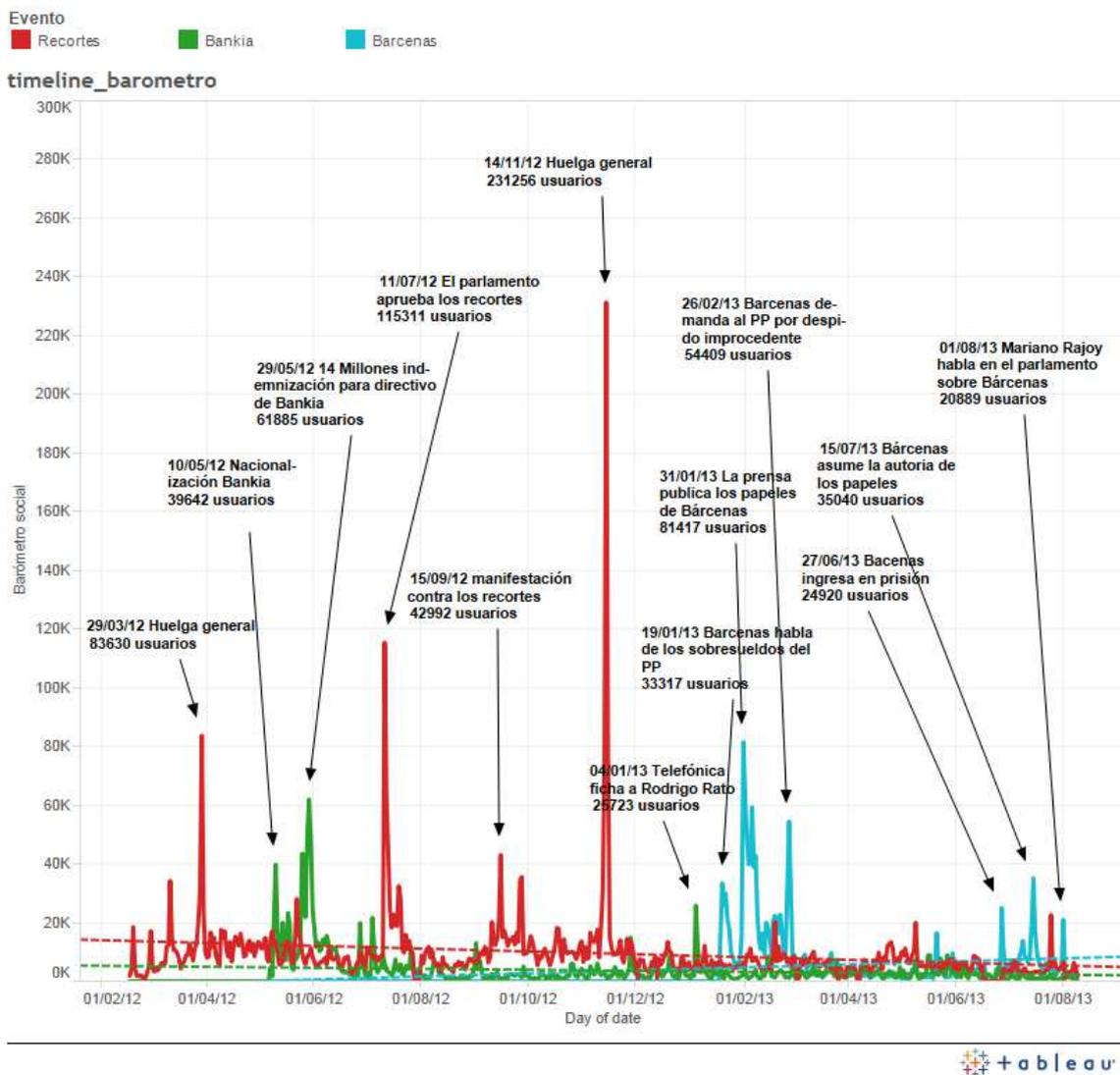
<sup>7</sup>[http://es.wikipedia.org/wiki/Sector\\_econ%C3%B3mico](http://es.wikipedia.org/wiki/Sector_econ%C3%B3mico)

<sup>8</sup><https://dev.twitter.com/docs/rate-limiting/1.1>

Este análisis se ha realizado sobre tres contextos significativos: “Los recortes” y el “caso Bankia” iniciados en 2012 y el “caso Bárcenas” surgido en el 2013. Para cada uno de ellos se ha calculado el número de usuarios que han participado cada día.

En la imagen 1 se muestra la evolución en el tiempo del número de usuarios que han participado en estos eventos con anotaciones en las fechas de mayor actividad. Se puede observar un volumen de publicación sostenido con una oscilación entre el fin de semana y los días laborables, con unos picos de actividad de duración limitada generalmente a un día, que corresponden a momentos críticos del evento. Con el transcurso del tiempo la participación se atenúa como se puede apreciar por las líneas de tendencia.

**Figura 1 Barómetro social 2012-2013**



Para determinar el nivel de reacción se ha establecido un coeficiente  $k_r$  que permite comparar reacciones en eventos diferentes

$$K_r = \frac{\text{usuarios pico}}{\text{media geométrica}}$$

Se utiliza una media geométrica para que sea menos sensible a los valores extremos. Los valores obtenidos son: 7.229 en los “Recortes”, 2.407 en “Bankia” y 4.330 en Bárcenas. Ordenados los picos por su coeficiente  $k_r$  de mayor a menor se obtiene el siguiente ranking que muestra la tabla 2.

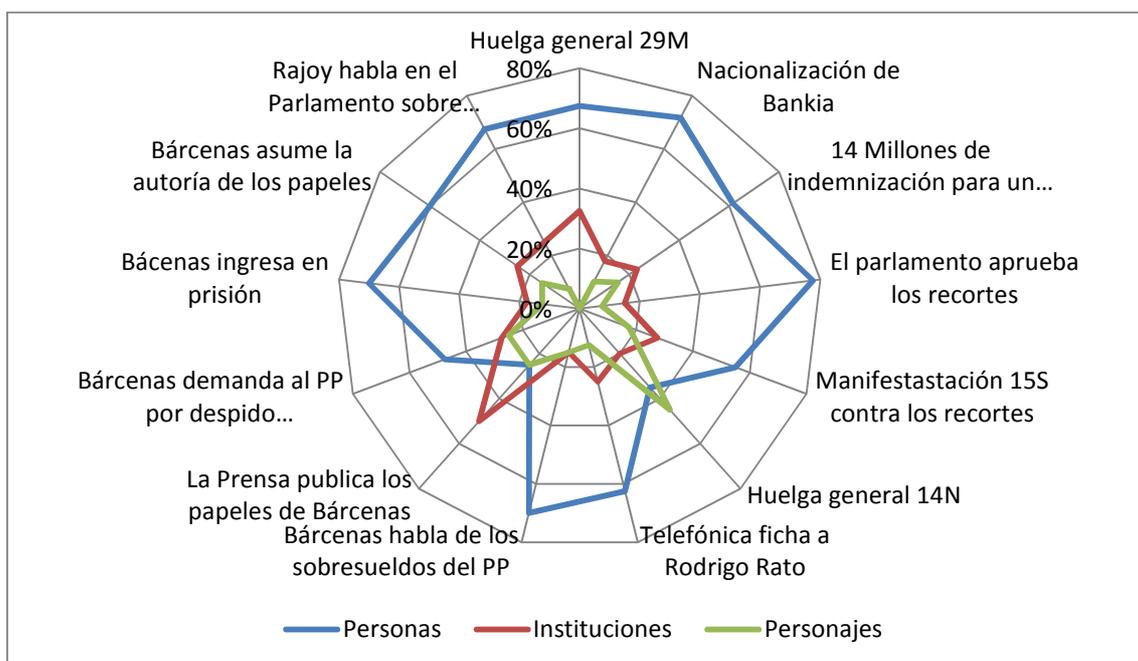
**Tabla 2 Ranking de reacción de usuarios ante sucesos**

<b>Fecha</b>	<b>Pico</b>	<b><math>k_r</math></b>	<b>Usuarios</b>
14/11/2012	Huelga general 14N	31,99	231.256
29/05/2012	14 Millones de indemnización para un directivo de Bankia	25,71	61.885
31/01/2013	La Prensa publica los llamados papeles de Bárcenas	18,80	81.417
10/05/2012	Nacionalización de Bankia	16,47	39.642
11/07/2012	El parlamento aprueba los recortes	15,95	115.311
26/02/2013	Bárcenas demanda al PP por despido improcedente	12,57	54.409
29/03/2012	Huelga general 29M	11,57	83.630
04/01/2013	Telefónica ficha a Rodrigo Rato	10,69	25.723
15/07/2013	Bárcenas asume la autoría de los papeles	8,09	35.040
19/01/2013	Bárcenas habla de los sobresueldos del PP	7,69	33.317
27/06/2013	Bárcenas ingresa en prisión	5,76	24.920
15/09/2012	Manifestación 15S contra los recortes	5,14	42.992
01/08/2013	Rajoy habla en el Parlamento sobre Bárcenas	4,82	20889

Como se aprecia en este ranking, la reacción es mayor ante los escándalos que en los sucesos que afectan al bienestar o a la economía del País. La indemnización de un ejecutivo de Bankia o los llamados papeles de Bárcenas aparecen en el ranking por encima de la nacionalización de Bankia o los recortes aprobados por el parlamento. Estas medidas permiten conocer la intensidad de la participación de los usuarios que son debidas a múltiples estímulos, por lo que es necesario profundizar en el contenido de los *tweets* para determinar las causas.

De acuerdo a la metodología propuesta se han clasificado manualmente los autores de los cuarenta *tweets* más difundidos de cada uno de los picos de reacción en: personas, instituciones y personajes. En la figura 2 se muestra el resultado obtenido, apareciendo los eventos relevantes ordenados en el tiempo conforme a las agujas del reloj con el porcentaje de los distintos tipos de usuarios en cada uno de ellos.

**Figura 2 Tipos de usuarios por pico de reacción**

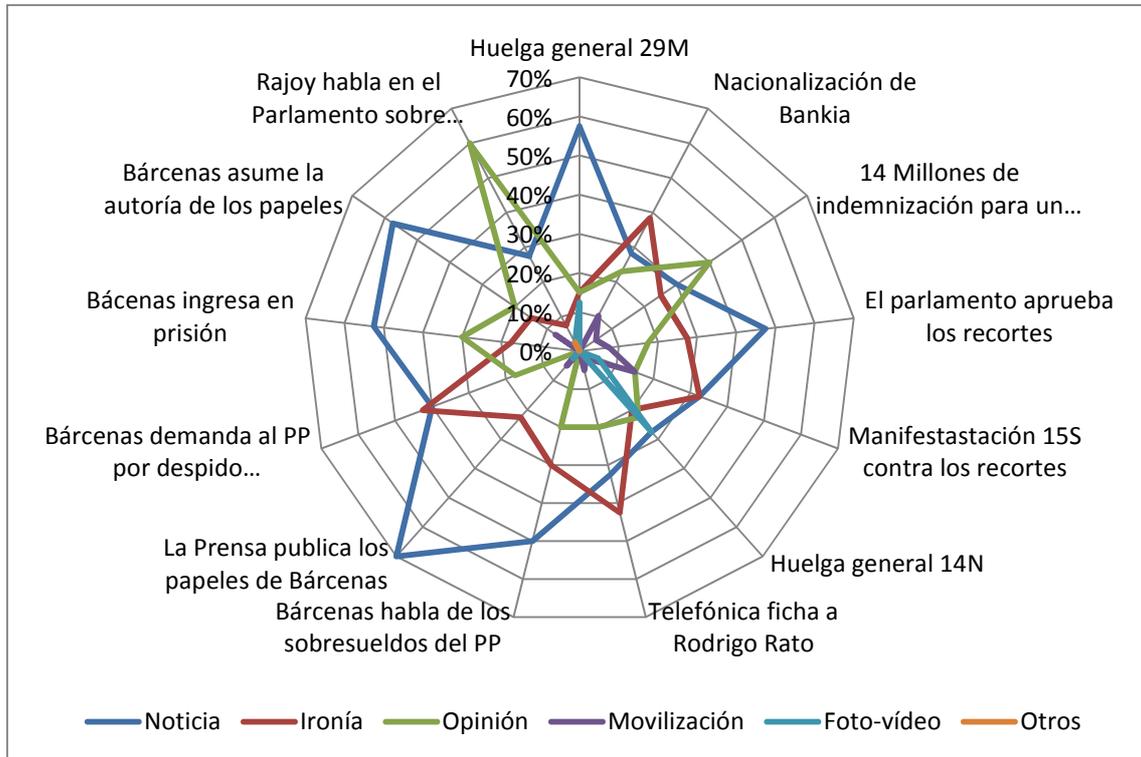


Como se puede apreciar mayoritariamente es el grupo “personas” el que consigue que sus *tweets* sean los más difundidos salvo en dos eventos en los que el patrón cambia dramáticamente.

- “La prensa publica los llamados papeles de Bárceñas”, en la que se observa un fuerte predominio de las instituciones (generalmente medios) en detrimento de las personas reales.
- La huelga del 14N, en la que aparece destacados los personajes frente a las personas reales.

Aplicando nuevamente la metodología propuesta se han clasificado manualmente los cuarenta *tweets* más difundidos en cada uno de los picos de reacción en los siguientes tipos: Foto-vídeo, Noticia, Opinión, Ironía, Movilización y Otros. El resultado queda recogido en la figura 3, apareciendo de igual forma que en la figura anterior, los eventos ordenados en el tiempo conforme al giro el de las agujas del reloj con el porcentaje de los distintos tipos de mensajes en cada uno de ellos.

**Figura 3 Tipos de mensajes por pico de reacción**



Como se puede observar en la imagen, el tipo de mensaje predominante es la noticia. Se da la paradoja de que los medios (instituciones) que generan las noticias no son los más difundidos, como se deduce de la clasificación de usuarios, sino que la propagación se realiza indirectamente por personas reales. Posiblemente es debido a que las noticias fluyen por las redes de usuarios y que la capacidad de propagación de los medios es limitada. Por otro lado, existen costumbres y reglas no escritas en el manejo de Twitter en que los usuarios de cierta relevancia tienden ser emisores de noticias más que difusores de los medios. Por ese motivo, el caso anómalo de “Los llamados papeles de Bárceñas” en que los periódicos fueron los más propagados requeriría un estudio más detallado de su comportamiento que desafortunadamente no se puede abordar en esta investigación.

También se observa en la imagen un aumento de la ironía ante dos sucesos quedan lugar a la misma: “Telefónica contrata a Rodrigo Rato” y “Bárceñas demanda al PP por despido improcedente”. En ambos eventos no solo los “personajes” son los que despliegan su ironía (su recurso de captación de seguidores) sino que son las personas

reales las que dan rienda suelta a su imaginación para asimilar con humor sucesos tan surrealistas.

Cabe destacar que en el evento “Rajoy habla en el parlamento sobre Bárcenas” la ironía desaparece y da lugar a la opinión, no parece que en este evento tenga cabida el humor e incluso los personajes opinan más que ironizan.

Finalmente, la información multimedia suele aparecer generalmente en eventos que corresponden a manifestaciones (14N, 15S y 29M) y los mensajes de movilización son muy escasos.

### ***Selección de los usuarios persistentes***

Este análisis se ha realizado sobre cuatro contextos significativos: “Los recortes”. “el rescate” y el “caso Bankia” iniciados en 2012 y el “caso Bárcenas” surgido en el 2013. Las personas que han publicado *tweets* sobre estos temas muestran una preocupación por la economía, los asuntos sociales y los escándalos políticos.

En la primera aproximación se ha aplicado el coeficiente de similitud de Jaccard a los *datasets* para ver el grado de similitud de audiencias.

$$J(A, B) = \frac{A \cap B}{A \cup B}$$

Como se puede apreciar en la tabla 3 la mayor coincidencia de usuarios se encuentra entre Recortes-Bankia con un coeficiente 26,45 y la menor entre Rescate-Bárcenas con 16,86

**Tabla 3 Coeficiente de similitud de Jaccard entre eventos**

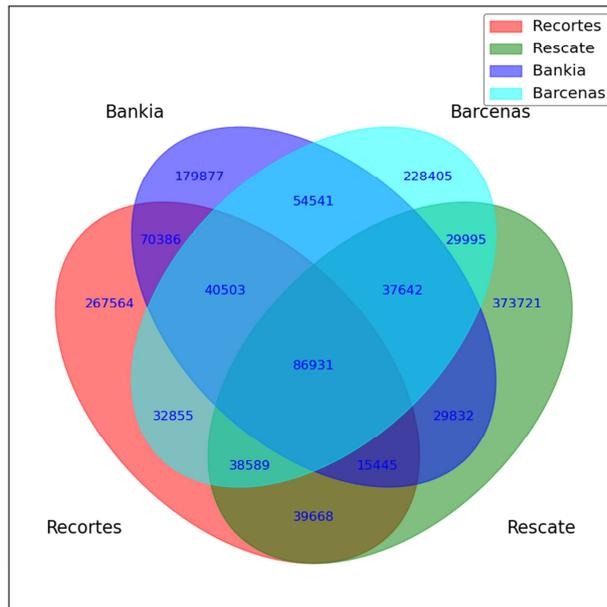
	<b>Recortes</b>	<b>Bankia</b>	<b>Rescate</b>	<b>Bárcenas</b>
<b>Recortes</b>	-	<b>26,45</b>	16,99	18,65
<b>Bankia</b>	<b>26,45</b>	-	19,35	25,99
<b>Rescate</b>	16,99	19,35	-	<b>16,86</b>
<b>Bárcenas</b>	18,65	25,99	<b>16,86</b>	-

La figura 4 muestra el diagrama de Venn con la intersección de usuarios de los cuatro *datasets*, del que se ha obtenido el grupo persistente.

$$\text{Grupo persistente} = \text{Recortes} \cap \text{Bankia} \cap \text{Rescate} \cap \text{Bárcenas}$$

**Figura 4 Diagrama de Venn de los cuatro datasets**

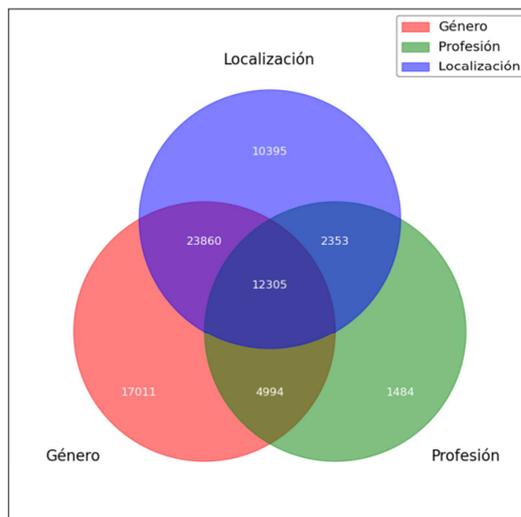
Recortes-Rescate-Bankia-Barcenás: usuarios únicos



El grupo persistente consta de 86.931 usuarios (5,52%), de los cuales al 83.2% se les ha detectado al menos la localización, el género o la profesión. En la figura 5 se muestra el diagrama de Venn de los distintos datos obtenidos. Se ha podido deducir la localización, el género y la profesión de 12.303 usuarios (14,15%).

**Figura 5 Datos obtenidos de grupo persistente**

Recortes-Rescate-Bankia-Barcenás-1111\_profiles: datos



Los datos de localización se recogen en la tabla 4 por total y desglosados por género. Como se puede observar el número de hombres es más del doble que el de mujeres en todos los casos, salvo en los perfiles en los que no se ha detectado localización. El porcentaje de usuarios de Madrid, Andalucía y Cataluña (28,62%) es mayor que el del resto de los usuarios localizados (22,15%), siendo de origen desconocido un 48,23%

**Tabla 4 Localización del Grupo persistente**

Autonomía	% Total	Género		
		% Mujeres	% Hombres	% Desconocido
Desconocido	48,23	10,68	19,60	17,95
Madrid	12,02	2,88	5,86	3,28
Andalucía	8,97	2,05	4,80	2,12
Cataluña	8,63	1,95	4,66	2,02
C. Valenciana	4,53	0,98	2,46	1,09
Galicia	2,80	0,66	1,38	0,76
Castilla y León	2,74	0,58	1,46	0,70
Murcia	1,87	0,39	1,01	0,46
Aragón	1,55	0,32	0,81	0,43
Asturias	1,50	0,34	0,80	0,37
Castilla-La Mancha	1,29	0,29	0,65	0,35
Canarias	1,29	0,30	0,70	0,29
Extremadura	1,28	0,28	0,69	0,31
País Vasco	1,14	0,20	0,64	0,29
Baleares	0,76	0,16	0,40	0,20
Cantabria	0,56	0,12	0,31	0,12
Navarra	0,46	0,11	0,25	0,10
La Rioja	0,27	0,05	0,16	0,06
Ceuta y Melilla	0,12	0,03	0,06	0,03

Los datos de las profesiones se recogen en la tabla 5 por total y desglosados por género. A diferencia de las localizaciones se ha encontrado que los medios de comunicación y la sanidad están balanceados, sin embargo en el subsector de tecnología es casi seis veces superior el número de hombres que el de mujeres. Respecto a los más jóvenes, los estudiantes, se mantienen en la proporción de doble número de hombres que mujeres

por lo que parece que la brecha digital de género se mantiene en las nuevas generaciones.

**Tabla 5 Profesiones del grupo persistente**

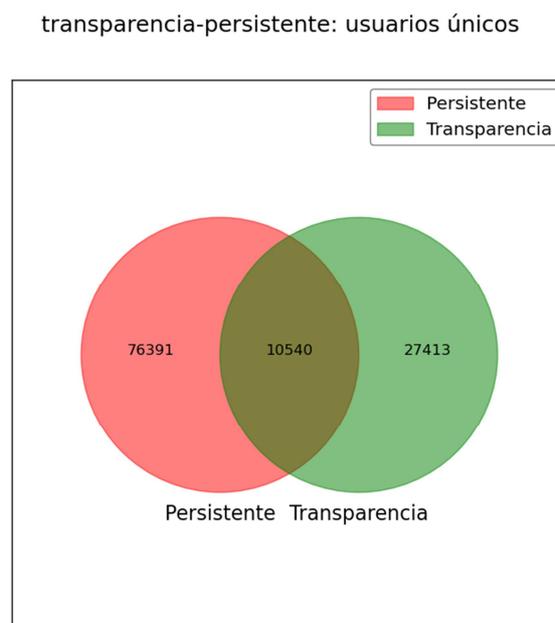
Subsector	% Total	Género		
		% Mujeres	% Hombres	% Desconocido
Desconocido	74,90	16,10	32,43	26,37
Medios comunicación	5,81	2,22	2,95	0,64
Estudiantes	5,00	1,28	2,82	0,90
Educación	2,91	0,62	1,74	0,55
Cultura y deporte	2,71	0,42	1,58	0,71
Servicios a empresas	2,22	0,37	1,44	0,41
Tecnologías	1,77	0,20	1,17	0,41
Admon Pública	1,56	0,38	0,95	0,23
sanidad	1,35	0,50	0,56	0,29
Construcción	0,62	0,09	0,40	0,13
Transporte	0,17	0,02	0,11	0,05
Comercio	0,16	0,03	0,10	0,03
Hostelería	0,14	0,02	0,08	0,04
Industria	0,13	0,01	0,10	0,02
Ejército	0,12	0,02	0,06	0,04
Ganadería	0,11	0,02	0,06	0,03
finanzas	0,07	0,01	0,05	0,02
Seguridad	0,07	0,01	0,03	0,03
Agricultura	0,06	0,01	0,04	0,01
Estética	0,05	0,03	0,01	0,01
Artesanía	0,02	0,01	0,00	0,01
Pesca	0,02	0,00	0,01	0,00
Trabajo social	0,01	0,00	0,01	0,00

*Uso del grupo persistente para analizar otros datasets*

La ley de Transparencia<sup>9</sup> tiene por objetivo restaurar la confianza en las instituciones y mejorar la calidad de nuestra democracia. Sin embargo, la opacidad de su tramitación y la no exclusión inclusión en la ley de algunas entidades ha creado un debate en Internet. Una petición canalizada por change.org de incluir a los partidos políticos en la futura ley alcanzo casi 180.000 firmas<sup>10</sup>

Se ha examinado cuántos de los usuarios del grupo persistente, que ya han mostrado una sensibilidad hacia temas sociales y económicos, han participado en la discusión de la Ley de Transparencia en Twitter. Sorprendentemente tan solo se han encontrado 10.540 usuarios (12.12%) como muestra el diagrama de Venn de la figura 6

**Figura 6 Grupo persistente vs. grupo Transparencia**



Analizando las demografías de conjunto  $Persistente \cap Ley\ transparencia$  se ha encontrado una mayor presencia de personas de Madrid y Andalucía mientras que disminuyó la población catalana como se observa en la tabla 6, lo que demuestra que la

---

<sup>9</sup><http://www.leydetransparencia.gob.es/index.htm>

<sup>10</sup><https://www.change.org/es/peticiones/a-la-vicepresidenta-del-gobierno-soraya-s%C3%A1enz-de-santamar%C3%ADa-que-se-incluya-a-los-partidos-pol%C3%ADticos-en-futura-ley-de-transparencia>

evolución de la Ley despierta diferente interés por zonas geográficas. Se ha marcado en negrita los valores que han aumentado respecto a la referencia del grupo persistente.

**Tabla 6 Localizaciones del grupo persistente  $\cap$  Transparencia**

<b>Autonomías</b>	<b>%Persistente <math>\cap</math> Transparencia</b>	<b>Grupo Persistente</b>
Desconocido	46.47	48.23
<b>Madrid</b>	<b>14.61</b>	12.02
<b>Andalucía</b>	<b>10.40</b>	8.97
Cataluña	7.09	8.63
<b>Comunidad Valenciana</b>	<b>4.72</b>	4.53
Castilla y León	2.57	2.74
Galicia	2.18	2.80
<b>Murcia</b>	<b>2.01</b>	1.87
<b>Canarias</b>	<b>1.55</b>	1.29
Aragón	1.44	1.55
<b>Castilla-La mancha</b>	<b>1.37</b>	1.29
<b>Extremadura</b>	<b>1.30</b>	1.28
Asturias	1.14	1.50
<b>Baleares</b>	<b>0.81</b>	0.76
País vasco	0.79	1.14
Cantabria	0.52	0.56
<b>Navarra</b>	<b>0.50</b>	0.46
<b>La Rioja</b>	<b>0.33</b>	0.27
<b>Ceuta y Melilla</b>	<b>0.20</b>	0.12

Lo mismo ocurre con las profesiones donde se encuentra una mayor presencia de los subsectores Medios de comunicación, Educación, Admon. Pública, Servicios para empresas, Sanidad, Ejército, Finanzas, Seguridad y Agricultura mientras que los Estudiantes pierden participación y el sector de tecnologías disminuye ligeramente.

**Tabla 7 Profesiones grupo Persistente  $\cap$  Transparencia**

<b>Subsector</b>	<b>%Persistente <math>\cap</math> Transparencia</b>	<b>% Grupo Persistente</b>
Desconocido	73.72	74.90
<b>Medios comunicación</b>	<b>6.71</b>	5.81

<b>Subsector</b>	<b>%Persistente <math>\cap</math> Transparencia</b>	<b>% Grupo Persistente</b>
<b>Educación</b>	<b>3.72</b>	2.91
Estudiantes	3.38	5.00
<b>Admon Pública</b>	<b>2.82</b>	1.56
<b>Servicios a empresas</b>	<b>2.61</b>	2.22
Cultura y deporte	2.02	2.71
Tecnologías	1.76	1.77
<b>Sanidad</b>	<b>1.47</b>	1.35
Construcción	0.58	0.62
<b>Ejército</b>	<b>0.25</b>	0.12
Transporte	0.17	0.17
Comercio	0.13	0.16
Industria	0.11	0.13
Ganadería	0.11	0.11
Hostelería	0.11	0.14
<b>Finanzas</b>	<b>0.10</b>	0.07
<b>Seguridad</b>	<b>0.09</b>	0.07
<b>Agricultura</b>	<b>0.08</b>	0.06
Estética	0.04	0.05
Artesanía	0.02	0.02
Pesca	0.01	0.02

Para profundizar en el conocimiento del conjunto de usuarios *Persistentes*  $\cap$  *Ley transparencia* se han estudiado sus vínculos. Las cuentas en Twitter están conectadas por dos tipos de redes. La red estática que está formada por las relaciones de declaradas de seguidores-seguidos y determina los caminos por los fluyen los *tweets*. La red dinámica que emerge de las relaciones de interacción (respuestas, RTs o menciones) que muestra el grado de atención que reciben los mensajes que circulan por la red estática.

- Izquierda + Plataformas: es la comunidad mayor, casi la tercera parte de los usuarios pertenecen a ella (30,4%). Aglutina a los simpatizantes de partidos, sindicatos,

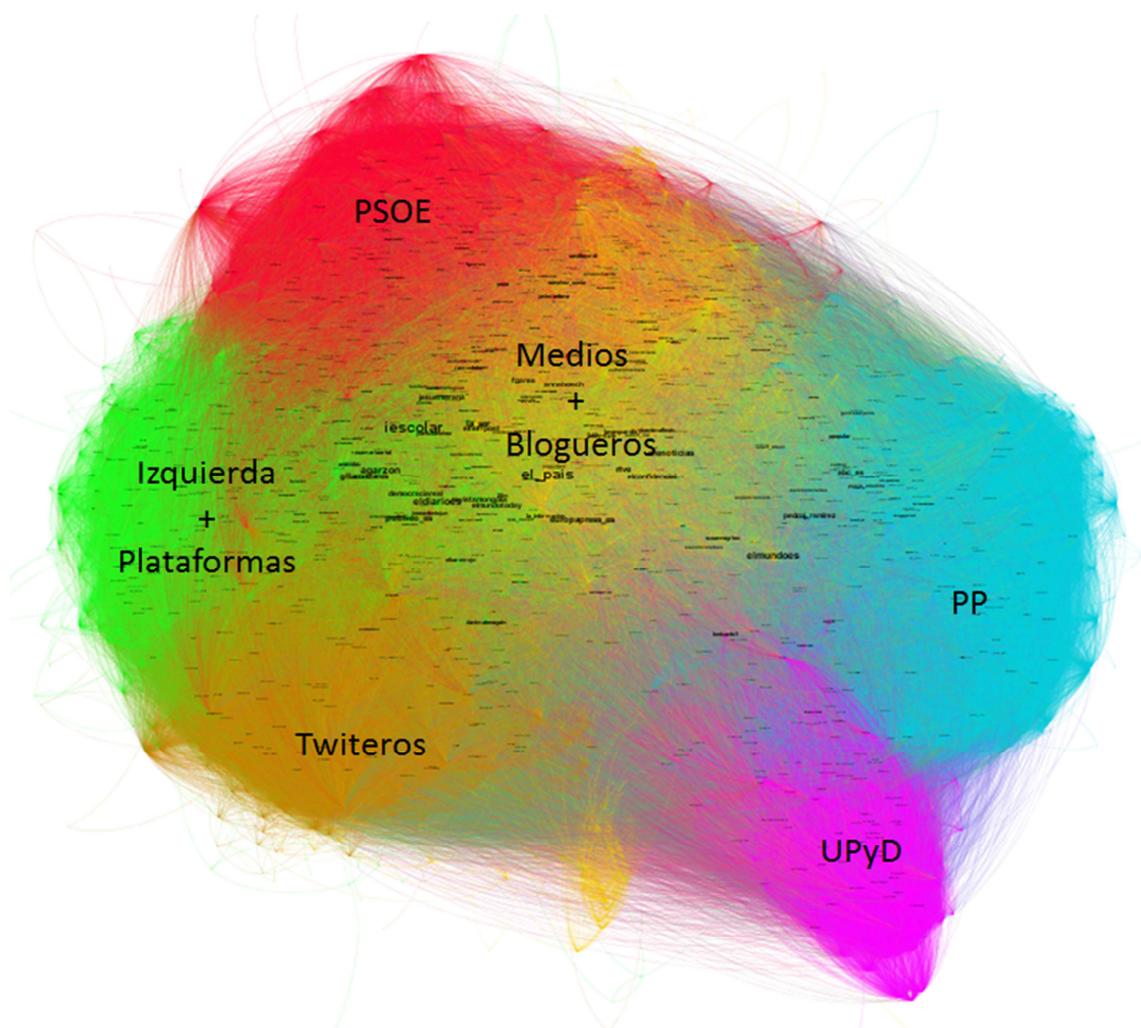
políticos, medios y periodistas de izquierda o de las plataformas ciudadanas y activistas.

- Medios y Blogueros: agrupa a los seguidores de los principales medios de comunicación (País, Mundo, TV, etc.), de periodistas o de blogueros. Este grupo (25,72%) muestra interés por la actualidad social pero no está definido políticamente.
- PP: congrega a los seguidores del Partido Popular, políticos, periodistas y medios de derechas. Estos usuarios (16,9%) manifiestan una polaridad política de derechas.
- PSOE: reúne a los seguidores de políticos del PSOE, UGT y algunos medios y periodistas afines. Muestran una polaridad política afín al PSOE 11,05% de los usuarios.
- Tuiteros: se le ha denominado así porque el factor común que los relaciona es que usan Twitter. En el grupo no se han encontrado ni partidos políticos, ni sindicatos, ni plataformas ciudadanas, ni medios. Si se ha encontrado algún “personaje” del tipo parodia y algún activista. Estos usuarios (10,18%) no muestran polaridad política definida ni siguen a los medios generalistas ni a blogueros destacados.
- UPyD: es el grupo minoritario (5,76%) e incluye a los seguidores de UpyD y sus políticos.

Además de estas comunidades, se han encontrado otras tan pequeñas que se ha descartado del análisis.

La figura 7 muestra el gráfico formado por la red estática de conjunto *CORE*  $\cap$  *Transparencia*. Cada uno de los nodos es un usuario, existiendo una conexión entre ellos cuando uno sigue a otro. Las relaciones son asimétricas debido a que en Twitter un usuario puede seguir a otro sin que exista correspondencia. Las comunidades están delimitadas por color. La cercanía entre los grupos está relacionada con una mayor afinidad. Como se puede apreciar en la imagen muestran afinidad las comunidades PP-UPyD y PSOE- Izquierda + plataformas. También se ha detectado cierta endogamia en los simpatizantes de PSOE, PP y UPyD en las que se encuentra un “un núcleo duro” que solo se relaciona entre ellos frente a otros grupos como Izquierda + Plataformas, Tuiteros y Medios + Blogueros que son más abiertos y mantienen conexiones con otras comunidades.

**Figura 7 Gráfico de la red estática del conjunto *persistente*  $\cap$  *Transparencia***



## Discusión

Se ha aplicado la metodología descrita a un conjunto de debates sociales de larga duración y con una alta participación de usuarios, en algunos casos superiores al medio millón, para obtener información relevante. De los resultados obtenidos se puede concluir:

1. En los momentos de máxima actividad de los eventos sociales el tipo de mensajes que más se difunden son noticias, generalmente publicadas por usuarios identificados como personas reales.
2. Los personajes (*fakes* o parodias) que basan su publicación en la ironía no alcanzan el máximo protagonismo en los momentos de crisis lo cual rebate la fama de frivolidad de los usuarios de Twitter.
3. En cuanto a las instituciones, aunque suelen ser las fuentes de las noticias no se las difunden directamente, son las personas reales las que las incluyen en sus *tweets* y

estos son propagados. Este comportamiento, sin llegar a ser una auto-comunicación de masas (Castells, 2009), si manifiesta una tendencia a la auto-selección de las noticias de los medios desde las masas.

4. Las reacciones de los usuarios son más fuertes ante los escándalos que en los sucesos que afectan al bienestar o a la economía del país porque son noticias que impactan más y que mueven a la participación.
5. La persistencia de los usuarios es baja, en el caso más favorable solo la cuarta parte ha participado en dos debates sociales. Si ampliamos la persistencia de usuarios de dos a cuatro eventos, la participación disminuye hasta el 5,52%. Aun así, este grupo, que hemos denominado persistente, corresponde a la nada desdeñable cantidad de 86.931 personas, cifra muy superior a cualquier sondeo.
6. El perfil de los usuarios del grupo persistente es urbano, casi la mitad de ellos de Madrid, Andalucía y Cataluña. Predominan los periodistas, los estudiantes y profesores. Existe una brecha de género, por profesión el número de hombres duplica al de mujeres incluso entre los estudiantes. Únicamente el sector de la sanidad y el de la comunicación están balanceados.
7. Utilizado el grupo persistente como referencia en el análisis de otros eventos, en este caso la Ley de Transparencia, se ha detectado una variación de la participación tanto de las zonas geográficas como de las profesiones. Esto denota distintas sensibilidades por colectivos profesionales o localidades, lo que podría ayudar a construir un mapa de intereses por estos conceptos.
- Analizando las conexiones del grupo persistente que ha participado en la Ley de Transparencia, se ha encontrado una estructura muy definida de comunidades agrupadas en su mayoría por ideología, casi dos tercios (64,56%), por lo que el debate está muy politizado.

La metodología propuesta permite tener una visión general de lo que está ocurriendo en los debates sociales, propone un coeficiente para comparar reacciones, clasificar la información más relevante, tener una referencia para poder comparar, conocer las características el grupo de usuarios más sensibilizados con los temas sociales y servir de referencia para el análisis de otros eventos.

Aunque los datos demográficos del grupo persistente no son tan completos como los de las encuestas son suficientes para determinar la sensibilización ante ciertos temas según

el género, la localización geográfica, y la profesión, como se ha demostrado en el caso de La ley de Transparencia. Este método es un primer paso hacia una metodología más sofisticada que en un futuro permita complementar a las actuales encuestas que realiza el CIS para el barómetro social<sup>11</sup>. A favor de Twitter está el sondeo continuo y la opinión espontánea, cada vez con más información sobre la que marcar referencias, tan solo queda ir avanzando paso a paso en la mejora de las metodologías.

## **Bibliografía**

- Anduiza, E., Gallego, A. y Jorba, L. 2009. "The Political Knowledge Gap in the New Media Environment: Evidence from Spain". Prepared for the seminar Citizen Politics: Are the New Media Reshaping Political Engagement?" Barcelona, 2009. Barcelona: IGOP.
- Anduiza, E., Cristancho, C. y Sabucedo, J.M. 2012. "Mobilization through Online Social Networks: the political protest of the indignados in Spain". *Arbor. Ciencia, Pensamiento y Cultura*, 188(756). Barcelona: UAB.
- Anduiza, E., Cristancho, C. y Cantijoch, M. 2012. "La exposición a información política a través de Internet En *Arbor. Ciencia, Pensamiento y Cultura*", 188 (756), 673-688. Berkeley Edición Digital.  
<http://arbor.revistas.csic.es/index.php/arbor/article/viewFile/1493/1504>
- Barberá, P. y Rivero, G. 2012 "¿Un tweet, un voto? Desigualdad en la discusión política en Twitter". Ponencia para el I Congreso Internacional en Comunicación Política y Estrategias de Campaña, 6 y 7 de Julio de 2012, Madrid. Madrid: ALICE. Retrieved July 04, 2012 edición digital <http://www.alice-comunicacionpolitica.com/files/ponencias/58-F4ffff91581342177169-ponencia-1.pdf>
- Blondel, V. D., Guillaume, J. L., Lambiotte, R., y Lefebvre, E. 2008. "Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*", 2008(10), P10008. Bristol: IOP Publishing Ltd.
- Cantijoch, M. 2009. "Reinforcement and mobilization: the influence of the Internet on different types of political participation". Prepared for the seminar Citizen

---

<sup>11</sup>[http://www.cis.es/cis/open/cm/ES/11\\_barometros/index.jsp](http://www.cis.es/cis/open/cm/ES/11_barometros/index.jsp)

- Politics: Are the New Media Reshaping Political Engagement? Barcelona, May 28th-30th 2009. Barcelona: IGOP.
- Castells, M. 2009. Communication power. Cambridge: Oxford University Press
  - Cheng, A, Evans y M. 2009. "Inside Twitter An In-Depth Look Inside the Twitter World Twitter Self-Disclosed Age Demographics". Versión Online <http://www.sysomos.com/insidetwitter/>
  - Colombo, C., Galais, C. y Gallego, A. 2012. "El uso de Internet y las actitudes políticas. Datos cuantitativos y cualitativos de España" En Arbor. Ciencia, Pensamiento y Cultura, 188(756), 751-766. Berkeley: Berkeley Electronic Press. <http://arbor.revistas.csic.es/index.php/arbor/article/viewFile/1498/1509>
  - Congosto, M.L y Aragón P. 2012 "Análisis de las elecciones 20N". Ponencia para el I Congreso Internacional en Comunicación Política y Estrategias de Campaña, 6 y 7 de Julio de 2012, Madrid. Madrid: ALICE. Retrieved July 04, 2012 edición digital <http://www.alice-comunicacionpolitica.com/files/ponencias/54-F50000054541342177364-ponencia-1.pdf>
  - Conover, D. 2011. "Predicting the Political Alignment of Twitter" Users. Proceedings of 3rd IEEE Conference on Social Computing SocialCom.
  - Deltell, L. Claes, F. y Osteso, J.M. 2013 "Predicción de tendencia política por Twitter: Elecciones Andaluzas 2012". Revista Ámbitos. Edición digital <http://ambitoscomunicacion.com/2013/prediccion-de-tendencia-politica-por-twitter-elecciones-andaluzas-2012/#more-148>
  - Fernández Prados, J.S. 2012. "Ciberactivismo: conceptualización, hipótesis y medida. In Arbor. Ciencia, Pensamiento y Cultura", 188(756), 631-639. Berkeley: Berkeley Electronic Press. Retrieved September 14, 2012 <http://arbor.revistas.csic.es/index.php/arbor/article/view/1490/1499>
  - Gayo-Avello, D. 2011. "A warning against converting Twitter into the next Literary Digest". Communications of the ACM. n. 10 Octubre 2011
  - Jungherr, A, Jürgens, P y Schoen, H. 2011. "Why the Pirate Party won the german election of 2009". Social Science computer review. Social Science Computer Review. May 2012 vol. 30 no. 2 229-234
  - Larsson, O. et al. 2011. "Studying political microblogging. Twitter users in the 2010 Swedish election campaign". New media & society. New Media & Society August 2012 vol. 14 no. 5 729-747

- Metaxas, P.T., Mustafaraj, E. y Gayo-Avello, D. 2011. “How (Not) To Predict Elections. Privacy, security, risk and trust ”, 2011 ieeee third international conference on social computing (socialcom). Páginas 165 - 171
- Mislove, A, Lehmann, S., Ahn, Y.A., Onnela, J. y Rosenquist, J. N. 2011. “Understanding the Demographics of Twitter Users”. In Proceedings of the 5th International AAAI Conference on Weblogs and Social Media (*ICWSM'11*), 554-557. Barcelona
- Osteso, J.M., Claes, F. y Deltell, L. 2013 “Hacia una definición de liderazgo de opinión en Twitter”. II Jornadas de Ciberpolítica
- Peña-López, I., Congosto, M. y Aragón, P. 2013-a.”Spanish indignados and the evolution of 15m: towards networked para-institutions”. IDP2013
- Peña-López, I. 2013-b “Casual politics: del clicktivismo a los movimientos emergentes y el reconocimiento de patrones”. II Jornadas de Ciberpolítica
- Sampedro V., Sánchez-Duarte, J.M., y Campos-Domínguez, E. 2013. “Participación ciudadana en las cibercampañas electorales. Debates teóricos y una aproximación tipológica”. II Jornadas de Ciberpolítica
- Skoric, et. all. 2012. “Tweets and Votes: A Study of the 2011 Singapore General Election”. SystemScience (HICSS), 2012. páginas 2583 – 2591
- Tjong, K. S. 2012. “Predicting the 2011 Dutch Senate Election Results with Twitter. Association for Computational Linguistics Stroudsburg”, 2012 Proceedings of the Workshop on Semantic Analysis in Social Media páginas 53-60
- Toret J., Calleja A., Marín O., Aragón P., Aguilera M., y Lumbreras A. 2013. “Tecnopolítica: la potencia de las multitudes conectadas. El sistema red 15M, un nuevo paradigma de la política distribuida”. Working paper IN3 (UOC).
- Tumasjan, A, Sprenger, T. y Sandner, P. 2010. “Predicting Elections with Twitter: What 140 characters reveal about political sentiment”. ICWSM, 2010