Tecnología y ejercicio de poder

Publicado en: Comunicación, redes y poder. Editora, María José Labrador.

Autores:

Dr. Oscar Jaramillo

Lucía Castellón

Escuela de Periodismo, Universidad Mayor

¿Es posible aventurar cuál es el futuro del periodismo en los próximos cuarenta años, sin caer en la ciencia ficción? ¿Podemos siquiera aventurar una hipótesis en un ecosistema de medios de comunicación altamente cambiante como el actual?

Cuando recién acabábamos de asimilar el impacto de los blogs y periodismo ciudadano, irrumpió la web 2.0 y las redes sociales. Antes que alcanzáramos a entender lo que estaba

sucediendo vino la revolución móvil. Y cuando recién nos empezábamos a medir los alcances de esta nueva ola, aparecen en el horizonte el Google Glass con todos los vestibles de la

mano, el uso de drones, los bots que escriben noticias de manera autónoma y los sensores que permiten convertir en realidad los escenarios distópicos más oscuros inspirados por 1984 de George Orwell.

Todos estos cambios que están a la vuelta de la esquina, hacen que sonriamos al ver el prototipo del periodista del futuro que proponía el académico estadounidense, John Pavilk en el año 2000. Visores de realidad aumentada, teclado en el pecho, antena unidireccional para conectarse a internet y la red GPS, cámara integrada y un computador portátil que colgaba en la espalda, hacía que el periodista del futuro pareciera un más un cyborg sacado de la película Star Treck, que un simple reportero. Sólo cinco años después, todas las funcionalidades del prototipo pudieron ser reemplazadas por un teléfono inteligente, que cabía en el bolsillo.

Es por eso que hacer futurología en el campo periodístico es casi imposible debido a la velocidad con la que los cambios tecnológicos se han sucedido en las últimas dos décadas. Cualquier apuesta sería aventurada y caería en el ámbito de la especulación, porque no sabemos cuál será el medio de comunicación del mañana, debido a la imposibilidad de anticipar el próximo invento que cambiará el ecosistema de los medios de comunicación.

Hemos llegado al punto en que no podemos prever cuál será la próxima revolución tecnológica. No obstante, lo que podemos analizar es el impacto dichos cambios han tenido sobre el periodismo y cómo eso ha influido en el ejercicio de poder. Si partimos de la base de que el rol fundamental del periodismo es informar a las audiencias para que voten de manera informada, cualquier cambio que afecte a la forma en que el periodismo investiga, puede tener consecuencia insospechadas sobre los equilibrios de poder y el ejercicio de la democracia.

Un claro ejemplo en donde la tecnología, el periodismo y la democracia entran en un delicado juego de poder, son las revelaciones del ex analista de seguridad de la CIA y la NSA, Edward Joseph Snowden, a los diarios The Guardian y The Washington Post, de que el gobierno de los Estados Unidos tiene dos programas de espionaje tecnológicos de alcance mundial (PRISM y XKeyscore), (Gellman, 2013) con los cuales vigila a gobiernos aliados y a ciudadanos estadounidenses dentro y fuera de su país.

El caso Snowden es significativo, no sólo por el alcance que tiene, sino porque demuestra que las capacidades casi orwellianas para espiar a autoridades y ciudadanos comunes y corrientes, las tiene toda la industria tecnológica, las agencias de marketing y potencialmente, los medios de comunicación.

Tal como lo explican Gellman y Poitras (2013) los avanzados programas de espionajes llevados a cabo por la NSA de los Estados Unidos, pudieron ser llevados a cabo gracias a la ayuda y, en algunos casos, al “pinchar” las redes informáticas de gigantes tecnológicos como Google y Apple. En la práctica eso significa que tanto la industria tecnológica, como los medios de comunicación, poseen las mismas capacidades de intrusión que los organismo de seguridad, al punto que estos los espían para vigilar a los ciudadanos.

No podemos asegurar cómo será el periodismo en el 2050, pero sí podemos afirmar que la tecnología le otorga al periodismo nuevas capacidades de investigación, que marcan un verdadero cambio de paradigma. Para poder poder entender estas nuevas capacidades de investigación, es necesario que nos hagamos cargo del estado del arte del periodismo. Debemos analizar brevemente el rol que tienen los bots, vestibles, drones, sensores, la minería de datos, la realidad aumentada y el fenómeno conocido como Big Data, en la conformación de un nuevo entorno periodístico se ve inserto dentro un sociedad transparente, tal como nos explican Craig y Ludloff (2011).

Estado del arte: tendencias actuales

Antes de entrar de lleno en la transparencia de la sociedad actual y el fenómeno del big data, es necesario que expliquemos el alcance y la función que tiene las nuevas herramientas que existen para recopilar, procesar y difundir información. Si entendemos las capacidades de los bots, vestibles, sensores y de los drones, podremos ilustrar de una manera más precisa, como el periodismo puede lograr capacidades de investigación comparables a las de la SNA y lo denunciado por Snowden.

Los bots (Verkamp, Gupta, 2013) son programas computacionales que imitan el comportamiento humano a través de la inteligencia artificial, lo que les permite realizar de manera automática, varias tareas que normalmente realizan los periodistas. Una de las aplicaciones más llamativas, es la aplicación Narrative Science que es capaz de entrar a una base de datos, analizarla, ver cuáles son las tendencias, interpretar los hallazgos y redactar las noticias. En la actualidad, las agencias informativas Reuter y Forbes utilizan este bot para crear noticias económicas y deportivas.

Otro ejemplo es Quakebot, desarrollado por el periodista y programador del diario norteamericano Los Angeles Times, Ken Schwencke, que a partir de las alertas emitidas por el USGS (Servicio Geológico de los Estados Unidos, por sus siglas en inglés) redacta noticias sobre terremotos y movimientos telúricos.

Otro ejemplo interesante es True Teller, un bot creado por The Washington Post que chequea de manera automática las declaraciones de los políticos, al buscar información en distintas bases de datos públicas y privadas, para comprobar la veracidad de los dichos.

Otro de los campos en los que se utilizan bastante los bots es Twitter. De hecho, el proyecto Bot or not bot, desarrollado por estudiantes de Eugene Lang The New School for Liberal Arts (Foyt, 2011), descubrió que sólo 33% de las interacciones en Twitter son realizadas por humanos. En Twitter los bots son utilizados para retwittear mensajes, publicar de manera automática noticias subidas a blogs y sitios web, publicar mensajes a través del uso de la inteligencia artificial y generar trending topics.

Cabe destacar que la publicación de mensajes de manera automática en Twitter, va desde la generación de spam, la mantención de perfiles de artistas famosos que tienen una gran cantidad de seguidores, para generar twitts y responder mensajes o crear trending topics de manera artificial.

Uno de los casos más documentados es el uso de un bot por parte del gobierno de Enrique Peña Nieto en México, para generar un trending topic que tapara o dismunuyera la importancia de los hashtag #MarchaYoSoy132, #EPNVeracruzNoTeQuiere y #MarchaAntiPeña (Robles, 2013).

Otro de los ejemplos importantes, es la cuenta de Twitter @TreasuryIO que es mantenida por un bot. Lo que hace es ingresar a la base de datos de la Secretaría del Tesoro de los Estados Unidos, sacar información, analizarla, buscar anomalías e informarlas a través de un tuit.

Como podemos observar, los bots se utilizan principalmente para investigar, redactar y publicar una noticia de manera autónoma, a través del uso de la inteligencia artificial. Están diseñados no sólo para hacer minería de datos y sacar conclusiones a través del estudio de grandes volúmenes de información, sino que además reemplazan al periodista al redactar la noticia.

La idea implícita del bot es reemplazar al periodista, a través del uso de la inteligencia artificial. Es crear la apariencia de que hay una persona que investiga y redacta los artículos. Es por eso que también se utilizan para tomar decisiones editoriales en los portales informativos. Un ejemplo de ello es Visual Revenue, un bot que tiene la capacidad para predecir cuál será historia más vista durante el día y de esa manera decidir de forma automática los titulares y las noticias recomendadas en el sidebar (la caja que está al costado derecho o izquierdo de la noticia principal).

Cabe señalar que el elemento central del bot es la inteligencia artificial, la cual puede ser usada para recopilar información, realizar minería de datos y redactar noticias. La recopilación de información es realizada de manera similar a la forma en que funcionan las arañas (web spider) de los buscadores.

La araña (spider) de un buscador es un bot que viaja (crawl, en los términos utilizados por los buscadores) por la red utilizando los enlaces (links) que hay en las páginas web. Lo que hace es entrar a un sitio web, lo lee, decide cuál es el tema que tiene, lo cataloga y envía una copia a su servidor.

A comienzos de la década pasada, las arañas decidían el tema de una página a través de la repetición de palabras y del estudio de las meta tags. Es decir, contaba cuál era la palabra que más se repetía y a partir de ello tomaba la decisión. La idea detrás de dicho razonamiento, es que si la página es sobre fútbol, la palabra que más se repetirá es fútbol. Junto a ello se tomaban en cuenta las meta tags, que son etiquetas que están dentro del código HTML, en la cual el web master señala cuál es el tema central del sitio.

No obstante, en la actualidad los bots tienen capacidad semántica, tal como lo señalan Lui y Birnbaum (2007). Eso significa que son capaces de entender el sentido de una oración, a través de distintos algoritmos, que se basan en la construcción de tesauros. Incluso, algunos bots desarrollados para estudiar las redes sociales, tal como es el caso de Infochimps, tiene la capacidad para analizar el tono de un comentario, para señalar si es positivo o negativo, frente a una marca.

La ventaja es que al poder entender el sentido de una oración, el bot puede catalogar de manera más precisa un contenido y procesarlo de manera efectiva. En el caso de Tell me more, un bot desarrollado por el investigador chileno Francisco Iacobelli (2010), permite evitar las repeticiones, para buscar fuentes alternativas.

Eso significa que un bot puede ser programado para que viaje por la web, redes sociales y bases de datos, para que busque de manera autónoma información que salga de los cánones comunes sobre un tema específico. Desde el punto de vista periodístico eso tiene grandes implicancias, porque un bot puede programarse para busque información sobre un tema o persona específica, en la web, redes sociales y bases de datos. Al pedirle que obvie la información repetida y seleccione sólo la información nueva, nos ahorra una gran cantidad de tiempo al momento de analizar los datos. De esa manera podemos asegurarnos de no perdernos en un mar de información.

Desde el punto de vista político, esto tiene grandes implicaciones porque convierte la transparencia de algo real, ya que hace casi imposible que un hecho del pasado sea escondido del escrutinio de un bot, debido a su capacidad para "gatear" por la web y bases de datos. Es necesario aclarar que las arañas (bots programados para recopilar información), tienen la capacidad para burlar medidas de seguridad que restringen el ingreso a usuarios comunes y corrientes, por lo que la protección con clave y nombre de usuario para ingresar a una base de datos, son insuficientes para frenar el ingreso de estos verdaderos entes autónomos que viajan por la red.

El cambio de paradigma es que la recopilación de información es realizada de manera autónoma por el bot, sin que el periodista necesite reportear nada. Asimismo debemos señalar que la calidad y cantidad de información que pueden recopilar estos bots, es comparable a lo que pueda realizar cualquier herramienta por la CIA o la SNA, debido a la capacidad que tiene para husmera en bases de datos públicas y privadas.

Asimismo el bot tiene la capacidad no sólo para recopilar la información y hacernos un reporte de manera automática, también puede procesar los datos, analizarlos, sacar una tendencia e interpretar los hallazgos.

Aquí es donde surgen dos conceptos que es necesario aclarar, antes de seguir avanzando: nos referimos a la minería de datos (data mining) y Big Data. Tal como afirman Payton y Claypoole (2014, 289 Kindle position) la minería de datos recoge información de manera sistemática, mientras que el Big Data implica la predicción de las tendencias sobre la base de los datos recolectados.

Un ejemplo de minería de datos que explican Payton y Claypoole (2014), es el que realizó la tienda por departamentos, Target, en 2011 en los Estados Unidos, para incrementar sus ventas. Los ejecutivos de la tienda habrían descubierto que uno los pocos momentos en la vida de una persona que está dispuesta a modificar sus hábitos de compras, es después del nacimiento de un hijo. Por lo tanto, los ejecutivos de Target pensaron que como el nacimiento de un bebé es un hecho público, muchas compañías podrían tratar de influenciar a los padres para que cambiaran sus hábitos de consumo. A raíz de ello, se enfocaron en detectar a las mujeres embarazadas, para enviarles información para que compraran productos en Target. Contrataron expertos en estadística que identificaron ciertos hechos, como la compra específica de algunas vitaminas y de pañales, delataban a una mujer cuando estaba embarazada. Al cruzar las variables (compra de vitaminas y pañales) pudieron detectar las mujeres que estaban esperando un hijo, para enviarles cupones de descuento para productos de recién nacidos.

Como podemos observar, la minería de datos lo que hace es detectar hechos a partir del cruce de variables, que ponen de manifiesto ciertos hábitos y comportamientos, por parte de las personas.

Craig y Luloff (2011, 190 Kindle position) afirman que la minería de datos puede ser utilizada para predecir comportamientos de todo tipo, incluyendo hábitos de compra, políticos y delictivos. Desde el punto de vista político y policial, el data mining ha sido utilizado para detectar comportamientos que se salen de los cánones comunes. Es decir, ubicar células terroristas, lavado de dinero o castigar a los disidentes dentro de un régimen totalitario.

Desde el punto de vista informativo, el data mining ha dado origen a lo que en la actualidad se conoce como periodismo de datos. Gray, Chambers y Bounegru (2012, 123 Kindle position) definen el periodismo de datos como el uso de la programación para automatizar la recolección y cruce de información desde fuentes de gobierno, policía o del mundo civil. Según estos autores el periodismo de datos permite realizar y encontrar conexiones entre cientos de miles de documentos, para realizar reportajes de investigación.

Un ejemplo del posible uso de la minería de datos, sería utilizar un bot para que entrara la base de datos de las resoluciones judiciales, leyera todos los documentos y nos entregara la información de todos casos en los que personeros de gobierno aparecen nombrados en su rol de autores o cómplices.

Esto es algo que se podría estar realizando de manera constante, sin que el periodista tenga que estar leyendo cada una de las resoluciones judiciales que aparecen cada día dentro de la base de datos.

Asimismo, la minería de datos permite encontrar incompatilidades y anomalías en autoridades públicas, al cruzar información entre bases de datos que provengan de fuentes distintas. Esta es la forma en que se han descubierto casos de corrupción en diversos países de América Latina, al comparar información proveniente de bases de datos comerciales, del

mercado de valores, inscripción de bienes raíces y declaraciones juradas de los bienes de un político.

El periodismo de datos es una herramienta efectiva que tienen los profesionales de la información, para convertir su profesión en un verdadero cuarto poder, con capacidades casi ilimitadas para fiscalizar a las autoridades.

Tal como lo señalan Gray, Chambers y Bournegru (2012), el análisis de los datos puede revelar la figura de una historia o proveernos una nueva cámara, que nos permita visualizar los hechos que conformen un reportaje de investigación.

Un ejemplo de ello es el Murder Mysteries Project[1] desarrollado por Tom Hargrove. Este investigador construyó una base de datos demográfica de más de 185 mil asesinatos sin resolver dentro de los Estados Unidos y les aplicó un bot para detectar la posible presencia de asesinos seriales. De esa manera permitió la detección de numerosos asesinos seriales, lo que finalmente se tradujo en captura de estos criminales, que habían pasado de manera inadvertida, debido a que las policías en los Estados Unidos, trabajan los casos de manera aislada dentro de cada condado. El bot que utilizó Tom Hargrove permitió revelar la presencia de historias, que de otra forma, hubieran sido pasadas por alto.

El uso de los bots, la minería y el periodismo de datos constituyen un verdadero cambio de paradigma dentro del periodismo, porque marcan una tendencia que debiera ir creciendo con el paso de los años, a pesar de no podamos anticipar los cambios tecnológicos que venga a futuro. Eso se debe a que éste es un cambio que está directamente relacionado con la lógica propia de la Sociedad de la Información, Sociedad Digital, Post Digital o queramos llamarla.

Si en 1997 Manuel Castells afirmaba que toda la experiencia humana era posible de ser digitalizada y, por lo tanto, de ser codificada en el lenguaje binario, en la actualidad esa es una realidad, aunque la mayor parte de la población no lo sepa.

La omnipresencia de los dispositivos digitales, como teléfonos inteligentes y tabletas, sumado al uso cada vez más extendido de las tarjetas de crédito y débito para realizar todo tipo de transacciones comerciales, junto con la digitalización de trámites burocráticos, hacen que cada una de las acciones que realizamos en nuestra vida cotidiana quede almacenada en una base de datos.

Acciones tan de uso común como dar el RUT[2] (DNI) cada vez que se realiza una compra en una farmacia para acceder a un supuesto descuento, permite que los medicamentos comprados, la dosis, el lugar y la fecha en que fueron comprados, queden almacenados en una

---

[1] Un completo resumen del proyecto puede ser leído en la dirección web http://projects.scripps-news.com/magazine/murder-mysteries/ (bajado el 11/07/14).

[2] Rol Único Tributario.

bases de datos. Gracias a la minería de datos, cualquier persona que pudiera acceder a esa información y que dispusiera de los conocimientos técnicos para programar un bot, podría visualizar las enfermedades de base de todos los clientes de esa farmacia, al confrontar las compras, con una base de datos que nos indique para qué sirve cada medicamento.

Tal como lo señalan Craig y Ludloff (2011, 154 Kindle Position) nosotros vivimos en un mundo digital. Trabajamos, socializamos, pagamos los impuestos, e incluso apostamos y perseguimos nuestros intereses sexuales de manera online. Y todo lo que realizamos deja una huella digital. Eso es lo que se conoce como Big Data.

Periodismo y Big Data

Hablar del Big Data implica un cambio de paradigma, debido a que nos sitúa en un nuevo escenario en donde la recolección y procesamiento de información son herramientas comunes, que están disponibles para las empresas tecnológicas, de marketing y los medios de comunicación.

Si hace veinte o treinta años, la recolección de información sobre las personas comunes y corrientes eran algo privativo de los organismos de seguridad de los estados, hoy en día es algo que puede ser llevado a cabo por el web máster de cualquier sitio web o fan page de Facebook.

La norma es que toda la industria tecnológica, medios de comunicación y el comercio en general, recopile información sobre las personas con fines comerciales, políticos o informativos. Al referirnos al Big Data como un fenómeno, lo que estamos haciendo es señalar que el uso de bots para recopilar información y realizar minería de datos, se convierte en una realidad de alcances planetarios, que abarca cada aspecto de la vida cotidiana de las personas, con la posibilidad de identificar a una persona en específico para ofrecerle productos de acuerdo a sus costumbres, gustos y cultura; castigarla por la crítica política que realiza a una autoridad política en las redes sociales; detectar comportamientos anómalos que puedan dar origen a un reportaje de investigación o fidelizar a un lector para entregarle sólo aquellos contenidos que le interesa.

Payton y Claypooole (2014, 23 Kindle postion) explican que el Big Data le entrega a las empresas y los gobiernos de todo el mundo la capacidad para encontrar la aguja en el pajar, al analizar y clasificar masivas cantidades de datos para encontrar patrones y correlaciones ocultas, que investigadores humanos pasarían por alto.

Tal como lo señalan Craig y Ludloff (2011, 183 Kindle Position) nunca antes se había conocido tanto sobre nosotros como hoy en día y toda esa información puede ser utilizada para predecir comportamientos de todo tipo y, en especial, los de compra, políticos o criminales.

La razón por la cual podemos afirmar que nunca antes se había sabido tanto sobre nosotros, se lo debemos a la existencia de los sensores, programas de reconocimiento facial y la capacidad para almacenar y procesar el rastro digital que deja cada aparato que utilizamos, que van desde el teléfono inteligencia, pasando por el auto, hasta terminar con la tarjeta de crédito.

Habitualmente olvidamos que los teléfonos inteligentes tienen incorporados una serie de sensores como acelerómetros, GPS, cámara y micrófonos. Lo que los usuarios ignoran, es que estos sensores recopilan información de manera silenciosa, sin que el usuario lo sepa. Tal como lo señala la agencia Reuters (2014) el gobierno de China considera al Iphone de Apple como una amenaza para seguridad nacional, debido a que registra con total exactitud los lugares y la hora en los que ha estado una persona gracias al GPS.

Para el gobierno chino, conocer los desplazamientos de gran parte de la población del país, constituye un dato altamente sensible, que incluso podría ser utilizado para revelar el estado de la situación financiera de la nación. Lo que sucede es que cada vez que el teléfono se actualiza, la información sobre los lugares y la hora en los que ha estado el usuario de Iphone es enviado a los servidores de Apple, sin que se tenga claridad para qué se utiliza dicha información.

Por otra parte, hay que agregar que los desplazamientos de una persona no es la única información que registra el teléfono inteligente, sin que el usuario se percate de ello. Es justamente aquí donde adquiere importancia el concepto de metadato. Rheingold (2012, p. 134) lo define como información sobre la información, que es registrada por un dispositivo digital cada vez que realiza una operación.

Eso significa que cada vez que entramos a un sitio web o realizamos un comentario en una red social, el computador o el teléfono celular, registran una serie de metadatos sobre esa operación. Los metadatos están asociados a la acción, por lo que no incluyen, por lo general, la información semántica. Es decir, si no referimos a un comentario realizado en un red social, el metadato no se refiere al texto del comentario, sino que a información complementaria como la hora, fecha y ubicación desde la que se realizó el comentario. La red social utilizada, junto con información de la máquina, como el sistema operativo, navegador, resolución y tipo de pantalla, tipo de y marca del dispositivo. También se puede acceder a información propia del trackeo (seguimiento de un usuario), tal como lo señala Kaushik (2010). Es decir, podemos saber los sitios y las páginas web que ha leído, cuánto tiempo ha estado en cada una de ellos, cómo llegó allí; si llegó a través de un link o de un buscador, y que término de búsqueda puso en el buscador y en qué objetos ha clickeado. En el caso de un correo electrónico, el metadato es la hora, fecha y lugar, desde donde se envió el mensaje, junto con los nombre y ubicación de los destinatarios.

Gellman afirma que para la NSA el estudio de la metadata es de mayor utilidad que los mensajes en sí mismos, ya que puede revelar la presencia y la estructura de un red de terroristas. Además, dependiendo de los métodos que se apliquen, puede exponer las condicio-

nes médicas o políticas de una persona, así como su afiliación religiosa, las negociaciones comerciales que esté llevando a cabo y las relaciones extramaritales.

Lo que hay que tener claro es los metadatos están constituidos por este rastro digital, del hablamos anteriormente. Por esa razón, todo dispositivo digital los genera. Las principales fuentes de metadatos son el uso de los computadores y la navegación por Internet, de los teléfonos inteligentes, tablets, consolas de videojuegos, televisores inteligentes, uso de cajeros automáticos, tarjetas de crédito, débito, tarjetas de pago para locomoción pública, televías (Tag's), pasaportes y cédulas de identidad con chip y automóviles. Otra fuente importante de metadatos son las redes sociales, ya que ponen de manifiesto las redes de contactos de las personas, junto con sus estados de ánimo y las relaciones sociales.

Si a ello sumamos los vestibles - como los Google Glass, pulseras y relojes inteligentes - veremos que las posibilidades de la metadata es casi ilimitada, porque registran el ritmo cardiaco de una persona, lo que podría revelar desde su estado de salud, hasta la vida sexual.

Esa es la razón por la cual el Big Data supone un cambio de paradigma en torno a las capacidades de investigación propias del periodismo. La principal modificación es que la información se recolecta de manera autónoma, sin que las personas se enteren de que ello está sucediendo. Sólo se percatarán, cuando sufran consecuencias o no sepan por qué les ofrecen tal o cual producto o les nieguen la visa a un determinado país.

Asimismo la recopilación del metadato puede ser realizada, desde el punto de vista técnico, por distintos actores, ya sea sea que intervengan en el proceso comunicativo o no. Un ejemplo de ello es lo que sucede cuando una persona interactúa con otra gracias a las redes sociales a través de un teléfono inteligente. El productor del sistema operativo (IOS o Andriod) puede registrar toda la metadata, que se produce gracias a la comunicación. También pueden grabar la misma información la empresa que creó el hardware (el teléfono inteligente en este caso), los desarrolladores de todas las aplicaciones que estén instaladas en el aparato y los web masters de los sitios web que utilice habitualmente. A ello hay que sumar las empresas de marketing a las que los sitios web y desarrolladores de aplicaciones le dan acceso a nuestros datos cuando colocamos aceptar a las condiciones de uso de una app o para poder registrarnos en un sitio. Y todo se produce sin contar los hackers, crackers y organismos de seguridad, que acceden a nuestra información a través del uso de virus y malware.

Cuando hablamos de Big Data nos referimos a una sociedad en la cual la información sobre las personas, sean públicas o privadas, es recolectada de forma intrusiva y cotidiana por distintos actores. Cuando afirmamos que la recolección de información es realizada de manera intrusiva, nos referimos a que se hace la mayor parte de las veces sin el consentimiento expreso e informado de las personas.

Desde el punto de vista ético, (Jaramillo, 2003) para recopilar la información se utilizan tanto mecanismo de monitoreo, como de intrusión. Por monitoreo entenderemos como el registro de los actos de una persona a través de la observación de cada uno de sus actos.

Eso es lo que realizan Apple, al registrar el lugar y la hora en que ha estado cada uno de los usuarios de Iphone. Cabe señalar, que para realizar el monitoreo se utilizan los sensores de los dispositivos digitales, los cuales toman nota de cada una de las acciones y desplazamientos de la persona.

En tanto, la intrusión se refiere a la obtención de datos almacenados dentro de los dispositivos del usuario, a través del uso de cookies, virus computacionales, malware o puertas traseras (back doors). La diferencia es que en la intrusión se husmea dentro de los datos almacenados en el computador, teléfono inteligente o tableta.

Pero el cambio más importante en la era del Big Data, es la generación de un gran mercado de información, al cual las empresas, gobiernos y medios de comunicación pueden acceder previo pago. Si antes era necesario que un organismo de seguridad como la Stasi de la República Democrática de Alemania estableciera amplios programas de escucha para espiar a sus ciudadanos, ahora podría hacer lo mismo y más, previo pago a las empresas que se dedican a recopilar metadatos.

La nube y las redes sociales

No obstante, la figura de lo que es el Big Data no está completa si no sumamos el metadato y la información semántica propia de la nube. Cada vez que mandamos un mensaje por Whatsapp, twitteamos, etiquetamos un contenido en Pinterest, guardamos una página web en Pocket o subimos una fotografía a Dropbox, estamos utilizando la nube.

En términos sencillos, la nube tiene como finalidad reemplazar los discos duros reales, por unos virtuales, en los cuales esté almacenada toda nuestra vida en línea, para que podamos acceder a ella cada vez que nos conectamos a internet, con independencia del dispositivo desde el cual lo realicemos.

Por lo mismo, el objetivo de la nube es que almacenemos nuestras fotografías y demás archivos de corte personal, en estos discos virtuales y no, en nuestros dispositivos. La ventaja es que al conectarnos a Internet podemos acceder a ellos, desde el computador, teléfono inteligente, tablet o vestible.

Todos los correos electrónicos que enviemos por Gmail o Yahoo, quedan almacenados en la nube, junto con todas las fotografías que subimos a Facebook, los comentarios que realizamos en Twitter y los videos que subimos a Youtube. Por regla general, todos los contenidos asociados a las redes sociales y las apps, se guardan de manera automática en la nube.

Esa es la razón por la que las presentaciones que se confeccionen con Keynote en un Ipad, por defecto son almacenadas en el dispositivo y en la nube. Lo mismo sucede con los comentarios realizados en distintas redes sociales, como Twitter, Pinterest, Whatsapp o Insta-

gram. Pero la diferencia está en que lo que se realice en dichas redes sociales, sólo queda almacenado en la nube.

Es aquí donde surge una pregunta: ¿Lo que está almacenado en la nube, sólo nos pertenece a nosotros, a los autores de los contenidos? Con independencia de lo que puedan decir los largos textos legales que que dan origen a las condiciones de uso de las redes sociales, la práctica nos indica todo lo contrario. tal como lo señalan Craig y Ludloff (2011, 1252 Kindle position) todos los archivos que conforman la nube, junto con su respectiva metadata, conforman lo que se conoce como el mercado de datos.

Estos son plataformas como Gnip[3] a las que pueden acceder las empresas, instituciones, gobiernos y medios de comunicación previo pago, para utilizar la información almacenada en la nube. Pueden obtener desde información estadística, como estudios de mercado y comportamiento en línea de los usuarios, hasta perfiles de gustos de personas específicas, con nombre y apellido.

Un ejemplo de ello es lo que hace Amazon con la lectura de los libros electrónicos o e-books. Como el servicio funciona bajo el formato de app (aplicación), todos los libros y cada una de las acciones que realizan los usuarios son almacenados en la nube. Eso quiere decir que los libros, junto con el número de la página leída, las compras, los textos marcados y subrayados, las búsquedas realizadas, el número de veces que se retorna a un libro o las obras que agregamos a la categoría de deseados antes de comprarlos, pasan a formar parte de la nube.

Por lo mismo, Amazon tiene los datos suficientes para, a través de la minería de datos, predecir mis gustos de lectura, con una precisión que puede asombrar hasta al más escéptico. La categoría de “recomendados para…” con el nombre y el apellido del usuario, demuestran un gran conocimiento de los gustos y hábitos de lectura de una persona. Cabe señalar, que una parte importante de los libros utilizados para esta investigación surgieron a partir de la recomendación de Amazon. A través del estudio de nuestros hábitos de lectura, búsqueda de información, subrayados y comentarios realizados en redes sociales, son capaces de recomendar la bibliografía de utilidad para nuestra investigación.

Desde el punto de vista periodístico, esto significa que terceras partes (como Amazon, en este caso) tienen la habilidad para recomendar textos y libros que nos pueden ser de utilidad para el reportaje que estamos realizando. Las consecuencias que tiene lo anterior son insospechadas, más si consideramos las consecuencias que tiene sobre los equilibrios de poder.

En primer lugar, nos plantea un escenario en donde la transparencia se convierte en una realidad, invisible a los ojos de los mismos periodistas y medios de comunicación. Significa que terceras personas pueden saber sobre el tema que estamos trabajando, sin que nosotros lo sepamos. Al estudiar la metadata de los libros comprados, las búsquedas realizadas por

[3] gnip.com

Internet, los archivos descargados o de las conversaciones realizadas a través de foros, chats, correos electrónicos, mensajería instantánea o video conferencia, se puede determinar con precisión los temas que estamos investigando.

Las consecuencias de esto pueden ir desde lo más banal, hasta el atentado a la libertad a expresión y pensamiento. Que una empresa nos pueda recomendar el libro preciso que nos ayudará en nuestra investigación académica o periodística, puede ser algo agradable, positivo o banal, según el cristal con el que miremos la realidad. Pero cuando la autoridad política tome acciones para evitar la publicación o para entorpecer la investigación, estamos hablando de un atentado a la libertad de expresión.

El tercer aspecto, es que en una sociedad transparente, en donde la recopilación de información se realiza de manera automática y segundo plano, sin que las personas siquiera se enteren, se genera una nueva brecha de acceso al poder. Si partimos de la base de que información es poder y ella está al alcance de todos, estaríamos en una sociedad altamente democrática.

Sin embargo, eso no es lo que sucede en la actualidad. En la era del Big Data cada acto que realizamos produce información, pero el cuello de botella se produce en la recopilación y tratamiento o minería de datos, para ser más exactos.

Al partir este artículo nos preguntamos si podíamos aventurar cuál sería el futuro del periodismo en los próximos cuarenta años. Casi de inmediato respondimos que es imposible dar una respuesta, debido a rapidez con la que se suceden los adelantos tecnológicos.

No obstante, todos los escenarios descritos, sumado a la irrupción de la realidad alterna, siguen la misma lógica descrita por Castells en 1997. Nos referimos a una sociedad en la cual el poder giran en torno a la recopilación, procesamiento y difusión de la información.

Ya sea big data, vestibles, bots, drones, minería de datos, sensores o dispositivos móviles, la información es algo que se mantiene como central. Pese a que numerosos autores consideran que es un tanto anticuado hablar de la Sociedad de la Información, la irrupción de los bots, el Big Data y la minería de datos, nos demuestran que son el corazón de los escenarios futuros.

Por lo tanto, podemos predecir cuál es el futuro del periodismo. La verdad es que sí. Es que el futuro del periodismo está ligado a los datos y más específicamente, al estudio de los metadatos y del conocimiento que de ahí extraemos.

Blibliografía

Libros:

Castells, Manuel. (1997). *La era de la información: Economía, Sociedad y Cultura.* Vol. 1. La Sociedad Red, Alianza Editorial, Madrid, España.

Craig, Terence; Ludloff, Mary (2011) *Privacy and Big Data: The players, Regulators, and Stakeholders*. Sebastopol: O'Reilly.

Foyt, Kelley (2011) *Project Bot or not*. Eugene Lang The New School for Liberal Arts.

Gray, Jonathan; Chambers, Lucy; Bounegru, Liliana (2012) *The data journalism handbook: How journalist can use data to improve the news*. Sebastopol: O'Reilly.

Gellman, Barton (2013) Introduction. En The Washington Post. *NSA Secrets: Goverment spying in the Internet age*. New York: The Washington Post.

Gellman, Barton; Poitras, Laura (2013) U.S., British Intelligence mining data form nine U.S. Internet companies in broad secret program. En The Washington Post. *NSA Secrets: Goverment spying in the Internet age*. New York: The Washington Post.

Iacobelli, Francisco; Nichols, Nathan; Birnbaum, Larry; Hammond, Kristian (2010). Finding New Information via Robust Entity Detection In Proactive Assistant Agents (PAA2010) AAAI 2010 FALL SYMPOSIUM. Arlington.

Jaramillo, Oscar. (2003) *Derecho a la Información en los portales y buscadores de la Web*. Tesis doctoral. Universidad Complutense de Madrid, Facultad de Comunicaciones, España.

Kaushik, Avinash (2010) Web analytics 2.0: The art of online accountability & science of customer centricity. Indianapolis: Wiley Publishing.

Liu, Jiahui; Birnbaum, Larry. (2007) *Measuring Semantic Similarity between Named Entities by Searching the Web Directory.* Web Intelligence, IEEE/WIC/ACM International Conference on , vol., no., pp.461,465, 2-5.

Payton, Theresa; Claypoole, Theodore (2014) *Privacy in the age of Big Data*. Rowman & Littlefield: Plymouth.

Rheingold, Howard (2012) *Net Smart: How to thrive online*. Cambridge: The MIT Press.

Verkamp, John-Paul; Gupta, Minaxi (2013) *Five incidents, one theme: Twitter, spam as a weapon to drown voices of protest*. School of Informatics and Computing, Indiana University. Indiana.

Noticias:

Reuters (sábado 12 de julio de 2014) Medios estatales chinos califican el iPhone como "amenaza para la seguridad nacional". emol.com http://www.emol.com/noticias/tecnologia/2014/07/12/669636/medios-estatales-chinos-califican-el-iphone-como-amenaza-para-la-seguridad-nacional.html Bajado el 12/07/14.