

## EL BAZAR Y LOS ALGORITMOS. UNA TIPOLOGÍA DE LA COMPETENCIA DE LAS MÉTRICAS DE LA INFORMACIÓN EN LA WEB

DOMINIQUE CARDON

Si para algunos la web de los documentos parece una leonera, la web de los *big data* tiene todos los ingredientes de un genuino bazar.<sup>1</sup> La extensión de internet a todas las esferas de la vida social genera una auténtica avalancha de huellas que pretenden convertirse en informaciones. En *La riqueza de las redes*, Yochai Benkler denomina «objeción de Babel» a la idea de que la proliferación de los datos transforma la web en un inmenso bazar en el que resulta imposible seleccionar la información de calidad. Para refutar este argumento, muestra que existe un orden propio de la web que se caracteriza por el hecho de que la selección de las informaciones ha dejado de ser monopolio de los *gatekeepers*, periodistas y editores, para ser confiada al conjunto de internautas anunciantes que, citándose los unos a los otros en los nichos conversacionales, hacen visibles las informaciones de calidad a los algoritmos que después van a ordenarlas para mostrarlas en los buscadores. Así, el orden de la web aparece como una coproducción de los internautas y las herramientas computacionales de la web, los algoritmos (Benkler, 2009). No obstante, esta refutación de la objeción de Babel sitúa el tema de la selección de la información en la zona de alta visibilidad de la web que se asemeja a un espacio público. Sostiene que la inteligencia colectiva de los internautas, entendida como una especie de razón práctica inmanente a las múltiples artimañas individuales presentes en la web, consigue disciplinar la torre de Babel. Las transformaciones actuales de la web, y sobre todo la proliferación de datos que hasta la fecha no eran considerados depositarios de un valor público, nos obliga a reconsiderar este argumento. La web ya no es un simple espacio documental. Las formas expresivas aparecidas en las redes sociales, la puesta en red de datos públicos reivindicados por el movimiento Open data, la proliferación de da-

tos procedentes de las huellas digitalizadas de los individuos, en el campo del comercio, la salud o los viajes han contribuido a ampliar significativamente «el bazar de la web». Este ya no se presenta únicamente como un espacio público que ha de dar visibilidad a los documentos cuya calidad ha sido estimada por los internautas.<sup>2</sup>

Capturados, almacenados, agregados, calculados, los datos entran cada vez más a formar parte de dispositivos de mercado, de vigilancia o de recomendación. De una utilización *ex post* reservada a los profesionales, las medidas se convierten en indicadores *ex ante*, se inscriben en las interfaces y se introducen en los usos más cotidianos. Por ahora, la creación de un *mundo de datos* es a menudo considerada, con entusiasmo o espanto, como una nueva potencia susceptible, para unos, de reinventar los mercados y la organización, de hacer que la democracia sea más transparente, de facilitar las interacciones entre las cosas y el medio ambiente o de elaborar los conocimientos predictivos, mientras que otros se alarman de los usos comerciales de los ficheros, de las amenazas contra la vida pública, de la dictadura de la hipervisibilidad o de la colonización del mundo llevada a cabo por un espíritu de cálculo y evaluación. No se interroga a los datos por su materialidad, por lo que son y lo que hacen, sino más bien para escenificar las esperanzas y temores que suscitan.

Ahora bien, mientras que el debate de algunos enloquece en torno a los riesgos y las oportunidades de los *big data*, otros dedican todas sus atenciones a ajustar los algoritmos destinados a «hacer que hablen» los datos. Mantienen una relación mucho más «realista» con respecto al hecho de que la multiplicidad de las huellas digitales a priori no encierra ningún significado y se desloman en construir nuevos artefactos computacionales destinados a darles significado a los *mundos de datos*. Gestionan la heterogeneidad creciente de las huellas, ponen a prueba modelos, movilizan nuevos saberes estadísticos, diseñan nuevas formas de visualización. Son el tipo de operaciones que nos gustaría clarificar en este ensayo, relacionando la proliferación de los datos publicados en la web con la diversidad de los principios introducidos con objeto de ordenarlos en dispositivos cada vez más dispersos. Para ello, hemos decidido dibujar un mapa somero del nuevo bazar de la web de datos, para

localizar las principales líneas directrices de los nuevos dispositivos que intentan disciplinar la extravagante aceleración de la digitalización de la vida social.

#### CINCO PRINCIPIOS DE CLASIFICACIÓN DE LAS INFORMACIONES EN LA WEB

Desde los primeros pasos de la web, la clasificación de las informaciones que aparecen en ella constituyó uno de los principales motores de su desarrollo. Y hasta finales de los años noventa, aún se podían contraponer de manera sencilla tres estrategias diferentes para ordenar las informaciones y facilitar la navegación de los primeros internautas.

#### *Editorialización, autoridad y audiencia*

La primera estrategia consiste en pedirles a las personas que efectúen una *editorialización* de la web confiándoles la tarea de elegir, catalogar y clasificar los mejores sitios web. La elaboración del anuario de sitios web de Yahoo (Batelle, 2005), llevada a cabo con una seriedad de documentalista y una pizca de extravagancia *geek* por un ejército de pequeñas manos, fue el emblema de una ambición que aún no era irrazonable: realizar una clasificación humana de toda la web. El anuario participativo del *Open Directory Project* (Dmoz), la redacción de artículos en Wikipedia, que hacen referencia a las fuentes legítimas de la web, o las selecciones de los sitios web ofrecidos en los grandes portales de información han prolongado hasta hoy el riesgo que supone el confiar al juicio humano el privilegio de llevar a cabo una selección razonada y asumida de lo mejor de la web. No obstante, el crecimiento exponencial de la red ha restado rápidamente cualquier atisbo de realidad al proyecto desmesurado de una ordenación exhaustiva de los sitios web de internet por manos humanas (Weinberger, 2007; Shirky, 2005). A partir de ahora, la editorialización de la web no puede concebirse como una empresa local, muestreada y contextual. Se ejerce sobre

un perímetro identificado a partir de criterios reivindicados y apunta a la promoción, en la superficie de un océano de informaciones insondables, de una selección electiva de los editores. Mientras que la abundancia de las informaciones debilitaba la estrategia editorial, la ampliación del número de editores ha hecho cada vez más cuestionable el privilegio que algunos, los profesionales de la edición, la documentación y el periodismo, se reservaban para seleccionar y jerarquizar las informaciones para los demás. Si hoy en día la editorialización de la web es local, es sobre todo más abierta, múltiple y polifónica.

Paralelamente, se puso en marcha otra estrategia de clasificación desde los inicios de internet mediante la competición que surgió entre los primeros buscadores (Batelle, 2005; Levy, 2011). Se trataba, esta vez, de confiar la jerarquización de los sitios a un *algoritmo* buscando la fórmula de clasificación que respondiera con la mayor pertinencia a las peticiones de los internautas. Sobrepasando a los primeros buscadores con base léxica, Google aportó, a finales de los años noventa, una solución técnica audaz que abrazaba la reivindicación de «*inteligencia colectiva*» de la internet de los pioneros: la mejor jerarquía de la información se basa en el gráfico de la web y mide la autoridad de las páginas web en base al peso que les otorgan las otras páginas web dirigiendo *links* hacia ellas. Todo contribuye a garantizar la hegemonía del PageRank en internet: una aspiración de exhaustividad que los seres humanos habían tenido que abandonar, la pertinencia de clasificaciones que reproducen principios meritocráticos difícilmente criticables, una definición de la calidad basada en la masa de datos calculables más que en su significado y la artimaña de un modelo publicitario de una eficacia sin igual. Google extiende su imperio en la web definiendo las reglas del juego de un ecosistema al que deberá ceder todo productor de información deseoso de garantizar su visibilidad en la red (Levy, 2010). Mientras que la web es, en lo esencial, de naturaleza documental, la lógica que preside su clasificación es la *autoridad* que el PageRank mide como el resultado de un voto ponderado de los internautas anunciantes.

El cotejo de los modelos de *editorialización* y de *autoridad algorítmica* escenario de un conflicto sobreactuado entre los seres hu-

manos y las máquinas, duró demasiado poco. Al ampliar aún más el volumen de sus datos, el número de usuarios y la diversidad de sus servicios, la web de los fundadores ha derivado en lógicas múltiples y contradictorias (Zittrain, 2008). La masificación de los usos ha transformado profundamente la web egocéntrica de los pioneros. Los principios que dieron forma a la web postulaban claramente que aquellos que publicaban también eran los que clasificaban las publicaciones; como simples lectores, los internautas pasivos quedaban excluidos. Este elitismo podía negar que fuera aristocrático debido al cuidado que prestaba a la reivindicación de ampliar a todos el acceso a la publicación. Pero el mesianismo participativo de los pioneros quedó cada vez más en entredicho debido a la realidad de las prácticas de los nuevos internautas (Rebillard, 2007). El creciente desfase entre el número de internautas participativos y pasivos, ha abierto la vía a nuevas métricas que miden el tipo de lectura de los que, en un número cada vez mayor, navegan en la web sin publicar. Un tercer tipo de medición, la *audiencia*, ha contrapuesto con gran celeridad la clasificación de los pasivos a la de los activos, sin contar los enlaces publicados, tal y como lo hace el PageRank, sino más bien los actos de navegación de los internautas, sobre todo mediante la medición de las «páginas visitadas» y del «visitante único» (Beauvisage, 2007; Jouët, 2004). Las métricas de audiencia importan al universo de internet las técnicas de conocimiento del público desarrolladas por los medios tradicionales para satisfacer la aplicación de un modelo económico publicitario. No obstante, inmerso en el mundo plástico e inmediato de las interfaces digitales, esta medida *ex post* que sanciona a toro pasado los contenidos que ha sabido reunir el público más numeroso, se ha convertido en un indicador prescriptivo que permite clasificar *en tiempo real* las informaciones en función del número de internautas que hayan hecho clic sobre uno u otro enlace. Lo que, en los medios de flujo, medía la recepción se ha convertido en internet en un determinante calculable de la editorialización de los enlaces, sobre todo bajo la forma de *hit parade*. La aparición de indicadores de audiencia en el mundo de internet muestra el interés por la visibilidad de aquellos que publican, pero sobre todo traduce el objetivo comercial de medir la notoriedad de las informaciones, ya no en

base a los pares, sino en base al público. De esta manera, la generalización de las mediciones de audiencia acompaña el momento crucial de la masificación de los usos de la web y reproduce, transformándolo, el debate entre la calidad, evaluada por los pares, y el éxito, medido por el mayor número, que tuvo lugar en la evaluación de los públicos en los medios tradicionales (Méadel, 2010).

### *Afinidad y velocidad*

La *editorialización*, la *autoridad* y la *audiencia* podrían haber sido las tres formas dominantes de la jerarquización de la información si la web, figura por fin realizada de la biblioteca universal, hubiera conservado su naturaleza esencialmente documental. Pero una transformación decisiva ha acompañado a la masificación de los públicos de la web: al lado y al margen de la web se ha desarrollado un inmenso espacio conversacional impulsado por la dinámica de las redes sociales de internet (Vanbremeersch, 2009). Una parte de aquellos que leían la web sin escribir en ella, han encontrado una manera de participar a un coste mínimo intercambiando frases sobre su estado, chistes, debates y enlaces comentados a partir de su página personal en las redes sociales. La llegada masiva de la conversación social a la web ha traído consigo numerosas consecuencias. En primer lugar pueden participar aquellos lectores que no hubieran escrito en una web formada únicamente por documentos públicos. Además, se despliega más fácilmente cuando los internautas pueden expresarse en una escena de visibilidad en «claroscuro», privada y pública al mismo tiempo, diseñándola a partir de su red social de conocidos ampliada a aquellas personas con las que comparten intereses comunes (Cardon, 2008). Y, por último, pone de manifiesto una demanda de personalización del acceso a las informaciones que pretende sustraerse a las clasificaciones centrales, de aplastante conformismo de tanto ser compartidas por todos, de las métricas de autoridad y de audiencia. Una de las contribuciones más originales de la Web 2.0, ha sido la de mostrar que más que una búsqueda temática de la información, la red social de

los individuos constituía un filtro de mayor calidad que les propondría contenidos de su interés. A partir de entonces, son las métricas afinitarias las que se han desarrollado para ordenar las informaciones en función de las preferencias personalizadas de los individuos. El *newsfeed* de Facebook, el *timeline* de Twitter y las mil y una herramientas que permiten agregar, ordenar y publicar contenidos en función de la estructura relacional declarada por los internautas en su red de contactos y afiliaciones, todo eso son piezas dispersas de estas nuevas métricas de afinidad que erosionan la autoridad central del PageRank (Pariser, 2011).

La profunda sacudida que hoy provocan las métricas afinitarias en la organización de la web también ha contribuido a resaltar otro principio de ordenamiento de las informaciones: su novedad. Al ampliar su oferta de servicios sin cuartel, la web se ha convertido en un medio en tiempo real que acelera el ritmo de difusión de las informaciones. A partir de ahora, la visibilidad de las informaciones ya no depende únicamente del tiempo profundo registrado por la autoridad meritocrática del PageRank, sino de la capacidad vírica de crear una ventana de atracción temporal durante la cual una información será masivamente transmitida por los internautas. Con Twitter, la temporalidad de la información, y especialmente su velocidad de contagio, se ha convertido en un nuevo principio de clasificación de las informaciones (Wasik, 2009).

### *Principios en competición*

La rápida aproximación de estas cinco lógicas de clasificación de la información, la *editorialización*, la *autoridad*, la *audiencia*, la *afinidad* y la *velocidad*, muestra bastante bien la pluralidad de los principios de ordenamiento desplegados en la web. A partir de ahora se trata de un objeto mucho más complejo y multiforme que la web documental de los pioneros con sus anuarios y sus buscadores. En la actualidad, un torrente de datos se vuelca en internet y transforma su naturaleza. Cada día se contabilizan 900.000 posts de blogs, 50 millones de tweets, más de 60 millones de estados en Facebook y 210 millones de e-mails (Pariser, 2011). Al expresiden-

VIP  
Móvil  
New o  
que Cardon  
y audiencia

te y director general de Google, Eric Schmidt, le gusta decir que si tuviéramos que digitalizar todas las noticias y los escritos desde los albores de la humanidad hasta 2003, serían necesarios cinco mil millones de gigabits para almacenarlos. Hoy en día, generamos este volumen de informaciones digitales en dos días.<sup>3</sup> Por extraordinariamente masiva que pueda ser, esta explosión de datos no es un simple cambio cuantitativo, incluso un paso a una escala mayor. Es, ante todo, la consecuencia de la desaparición del filtro que aplicaban los *gatekeepers* para separar las informaciones públicas de las que no lo eran (Cardon, 2010). El espacio informacional tradicional era binario. Los periodistas y editores se ponían de acuerdo sobre lo que tenía que ser público y, por consiguiente, disfrutaban de una amplia visibilidad. Todo el resto, informaciones no verificadas, manuscritos improbables, declaraciones arriesgadas, diarios íntimos, blocs de notas o huellas cotidianas, permanecía en la sombra. Internet sustituye este recorte binario por una graduación sin costuras: todo está potencialmente accesible, pero no goza de la misma visibilidad. El filtro de las informaciones no ha desaparecido; sin embargo, su economía se ha transformado. A priori ya no separa las contadas informaciones seleccionadas por los profesionales del espacio público. Sin embargo, a posteriori ordena un inmenso océano de informaciones distribuyéndolas en base a la escala de la visibilidad calculada por el PageRank y cuyo principal trofeo es el acceso a la primera página de Google. A partir de ahora, ante este torrente de datos heterogéneos, el nuevo poder de clasificación pertenece a los algoritmos (Shirky, 2009). Esta revolución en el proceso de clasificación de las informaciones, el paso del filtro a la escala, de los *gatekeepers* a los algoritmos, aún está poco o mal visto por los internautas.<sup>4</sup> La web es el teatro de una despiadada competición por la visibilidad, donde «ser visto» significa cada vez más ser visto por los algoritmos. Asimismo, las reglas de este nuevo juego, inscritas en los parámetros de las métricas de la web, se imprimen cada vez más en las prácticas y las costumbres de los internautas (Balnaves, 2011; Galloway, 2004).

Si el PageRank de Google constituye la espina dorsal de este nuevo sistema informacional, su aplastante centralidad solamente consigue componer un orden de primer nivel con las informacio-

nes de internet. El PageRank clasifica las respuestas a las preguntas. Responde a intenciones de búsqueda explícitas que ponen de manifiesto prácticas fuertemente marcadas por la naturaleza documental de la web. La diversificación de los usos, y sobre todo el giro conversacional, ha guiado a los internautas hacia otros modos de navegación y de exposición a las informaciones que no van por el camino de los buscadores. Asimismo, una increíble proliferación de herramientas de clasificación ha hecho su aparición en las diferentes plataformas de la web, para producir clasificaciones locales arquitecturando el espacio informacional que el internauta está acostumbrado a frecuentar. A su manera, el *newsfeed* de Facebook, el *hashtag* de Twitter, los *puntos de karma* de Slashdot, los *best friends* de MySpace, las *recomendaciones* de Amazon, los *grupos «50fav»* de Flickr, las *opiniones de hoteles* de TripAdvisor, los *tags* de Delicious, los *vídeos más vistos* de YouTube, los *mejores comentarios* de artículos de Rue89, los *alcaldes* de Foursquare, etc., constituyen otras tantas técnicas de clasificación destinadas a hacer que ciertas informaciones sean más visibles que otras para los internautas. Ninguno de estos artefactos encierra de manera «pura» uno u otro de los cinco principios en los que se apoya la historia de las métricas de la web. Al multiplicarse, han hecho que los principios de organización que permiten darle un significado a las razones por las que una información está mejor clasificada que otra, se hayan complicado y entremezclado. Esta ilegibilidad contribuye al desconocimiento por los internautas de la estructuración de las informaciones que dan forma y dirigen sus navegaciones.

#### LA EXTENSIÓN DEL PERÍMETRO DE LOS DATOS

La web amplía su poder mediante su capacidad para capturar y hacer que el mundo sea manipulable por el mayor número posible de personas. La dinámica que preside su extensión se alimenta de un doble movimiento que apunta, por un lado, a *informar al mundo* y, por otro lado, a *dotarlo de significado*. Antes de que la operación de digitalización facilitara la portabilidad, la circulación y la transformación de las informaciones digitales, la condición exigida

para favorecer este doble movimiento era que los datos fueran *públicos*. En todas partes, empiezan a ver la luz reivindicaciones que invitan a compartir un volumen cada vez mayor de informaciones. Por un lado, los internautas, sobre todo en las redes sociales, se ven cada vez más expuestos a compartir informaciones de todo tipo sobre su identidad. Por otro lado, el movimiento de los «datos abiertos» (*open data*) invita a las instituciones y empresas a publicar los datos estadísticos «brutos» que guardaban celosamente para uso personal de sus propios intérpretes. Las comunidades de la web han empezado, en numerosos campos, a describir con todo lujo de detalles el mundo que les apasiona y les interesa, como los cartógrafos de OpenStreetMap, los enciclopedistas aficionados de Wikipedia o los ciudadanos vigilantes de Ushahidi. Por último, la renovación de las técnicas de *data mining*, iniciada por el Social Customer Relationship Management (Social CRM), invita a explorar las huellas que los clientes de las empresas dejan en la red.

### *La puesta en información del mundo*

Lo que motiva el crecimiento continuo de la web es la voraz extensión del perímetro de los datos que recolecta sobre el mundo, procedimiento cuya *ambición* es la «digitalización de la vida misma» (Rouvroy, 2009). La web se amplía transformando en objeto digital cosas, actividades, flujos o estados que antes permanecían en la sombra. Silenciosamente ancladas en su contexto, estas *huellas* no tenían una razón especial para circular en un espacio de información más amplio. Conservan las *huellas* que cualquier acción deposita en su contexto.<sup>5</sup> Idiosincrásicas, nada o poco intencionales, aisladas, carecen de sentido fuera de su contexto y están tan profundamente vinculadas a él que no se distinguen del flujo de actividad de los individuos. Otorgándoles una existencia digital, la web hace que *estados* del mundo que hasta ahora habían pasado desapercibidos, se tornen perceptibles, móviles y numerables. Este insaciable movimiento de expansión, cuyo punto de fuga es el desarrollo de una internet de los objetos, procede de un proceso de racionalización que lo que pretende es capturar estos estados del

mundo «con poco ruido» para transformarlos en informaciones digitales. Se trata de transformar huellas en indicios, de señalar con huellas (Merzeau, 2009). La transmutación digital de estas huellas se descompone en una serie de operaciones. Primero se suelta —a menudo, dañándolas— la referencia que ataba estas huellas a su contexto para darles un valor indiciario. A continuación, se les superpone una existencia digital que facilita su prensión y su puesta en circulación en un espacio que permita compartirlas, asociarlas y compararlas con otras. Y, por último, se las transforma en soportes de evaluación creando posibles conexiones para actividades de notación, comentario y clasificación. Entonces, las huellas indiciarias pueden convertirse en señales informacionales cuando se comparten en público y, de ese modo, permiten iniciar operaciones interpretativas que les otorguen un significado.<sup>6</sup>

Si la digitalización registra un número cada vez mayor de estados del mundo, resulta incuestionable que *acelera*, en un movimiento propiamente endógeno, la producción de estados del mundo destinados a su registro digital. Un ejemplo permite ilustrar este proceso continuo y sin sujeto de captura del mundo. Se ha observado con frecuencia que el desarrollo de la fotografía digital ha ampliado el espacio fotografiable. El universo capturado por la fotografía digital tenía un perímetro limitado. Los lugares turísticos, los acontecimientos importantes de la vida social, como las bodas o los cumpleaños, y la representación de los allegados en contextos ritualizados constituían el principal repertorio de la fotografía *amateur* (Bourdieu, 1965). El desarrollo de lo digital, y la generalización de la función fotográfica en cualquier tipo de dispositivo portátil, ha ampliado el mundo a fotografiar a objetos, situaciones o lugares inesperados, improbables y estafalarios. Incluso para algunas personas, esta captura meticulosa del mundo ha sido objeto de una implicación apasionada, obsesiva o poética, tal y como muestran las inmensas colecciones que aparecen en las páginas de intercambio de fotografías que recogen negativos de comidas de aviones, chapas de botellas de cerveza, cámaras de vigilancia, aceras, animales domésticos o autorretratos sacados todos los días a la misma hora.<sup>7</sup>

Analizados individualmente, estos registros maniacos de la vida cotidiana de los individuos significan bien poco; sin embargo,

puestos en común, pueden representar espacios documentales inesperados e iniciar las coordinaciones oportunas que constituyen el principio de la formación de comunidades en la web (Beuscart y otros, 2009; Aguiton y Cardon, 2007). Observamos un proceso similar para numerosas huellas de la vida cotidiana. No se nos había ocurrido contar el número de amigos, Facebook nos ha enseñado a hacerlo. No medíamos el número de kilómetros recorridos en un día, Fitbit lo hace por nosotros. No contábamos el número de actos militantes (reuniones, octavillas distribuidas, puerta a puerta o llamadas telefónicas), MyBarackObama.com lo ha convertido en un motivo de orgullo a lucir delante de nuestros allegados. Mientras que los trayectos cotidianos, los esfuerzos deportivos o el consumo eléctrico no dejaban huellas explotables, eran olvidados al instante y no se convertían en el soporte de ningún conocimiento (o como mucho de un conocimiento simplemente local y de uso estrictamente personal), se han convertido hoy en día en capturas explotables que pueden, en ciertas circunstancias, servir para producir información, red, métricas o conversaciones. La digitalización de las huellas participa en un movimiento de influencia sobre el mundo que aumenta las posibilidades de dar sentido a múltiples huellas cuando estas se publican o se comparten.

No obstante, el significado de estas nuevas huellas está lejos de ser adquirido desde un principio. Incluso, cabría preguntarse para qué sirven estas huellas ruidosas, extrañas y aparentemente sin objeto. Otorgarles el estatus de «información» parece sobrestimar en gran medida su capacidad indiciaria. A menudo, la web captura nuevas huellas sin una finalidad explícita. Deja al libre juego de la cooperación entre internautas decidir lo que estos datos, una vez publicados, puedan significar.<sup>8</sup> Los promotores de los datos abiertos dicen a las instituciones que si no pueden decirles de antemano para qué usos han de publicarse los datos, siempre encontrarán internautas que, a posteriori, sabrán hacerles hablar.<sup>9</sup> La inversión temporal del *ex post* hacia el *ex ante* es una característica central de la economía informacional de internet. Los datos son públicos por defecto, mientras que estábamos acostumbrados a considerarlos datos privados por defecto. La web publica y después filtra, cuando los medios tradicionales filtraban antes de publicar. Esta inver-

sión también se aplica al sentido del trabajo estadístico cuando pasa de una situación de escasez a un contexto de abundancia de datos. Al tiempo que la estadística ha concebido durante mucho tiempo la producción de datos como una manera de responder a hipótesis formadas de antemano, hoy en día, las hipótesis se manifiestan como consecuencias emergentes del trabajo de reconocimiento efectuado en el seno de gigantescas masas de datos «idiotas». Las nuevas técnicas de tratamiento estadístico del *data mining*, basado en el reconocimiento de datos y el aprendizaje, no formulan ninguna hipótesis previa sobre los datos que calculan, por muy heterogéneos que sean. Y no es poco frecuente que pretendan convertir en pertinentes las correlaciones sin causa (Rouvroy y Berns, 2010; Weinberger, 2011). Para algunos, el trabajo de interpretación de las ciencias humanas no tendría razón de ser en la era de los *big data* y podría abandonarse al reciente matrimonio entre el *behaviorismo* y los algoritmos. En «La era de los Petabytes», Chris Anderson, redactor gerente de *Wired*, escribe:

Es un mundo en el que las cantidades masivas de datos y las matemáticas aplicadas sustituyen al resto de instrumentos que podrían utilizarse. Deja de lado a todas las teorías sobre los comportamientos humanos, de la lingüística a la sociología. Olvidaros de la taxonomía, la ontología y la sicología. ¿Quién puede saber por qué la gente hace lo que hace? El hecho es que lo hacen, y que podemos trazarlo y medirlo con una fidelidad sin precedentes. Si tenemos datos suficientes, las cifras hablan por si solas.<sup>10</sup>

### *La manipulación del mundo*

En paralelo a este movimiento de información sobre el mundo, la web convierte los objetos digitales en infinitamente plásticos. Sus fronteras se vuelven inciertas. Las capturas que ofrecen a la manipulación proliferan. Se multiplican las funciones que permiten actuar sobre los objetos digitales. El contador del número de lectura, el comentario, la puntuación, el *tag*, y, de ahora en adelante, el «like» y el «+», constituyen un número tan grande de captu-

ras que, a partir de ahora, resulta posible atar a cualquier objeto digital sea cual sea. A veces, sin saber cuáles van a ser los usos dados a estas nuevas capturas, los diseñadores de servicios, a menudo eligen por defecto añadir dichas funcionalidades a su plataforma. Una foto en Flickr puede ser recomendada, etiquetada, incluida en uno o varios grupos o añadida a favoritos. Un estado de Facebook puede compartirse, comentarse, etiquetarse «Me gusta», y los comentarios que suscita pueden, a su vez, ser etiquetados como «Me gusta» y comentados, etc. Todo está preparado para hacer que los objetos de la web sean manipulables por los usuarios hasta el infinito. Esta manejabilidad digital transforma al mismo tiempo a los internautas y a los objetos que manipulan. Estos últimos pueden convertirse en el objeto de múltiples clasificaciones. Se convierten en multifacéticas y alimentan métricas con principios heterogéneos. Los internautas, por su parte, pueden hacerse con los objetos de diferentes maneras, en base a intenciones y competencias muy diferentes. A diferencia de los objetos cerrados con capturas diferentes y limitaciones del espacio público tradicional, los libros, los artículos de prensa o las películas digitales son infinitamente manipulables. La economía de la evaluación en el seno del espacio de visibilidad que representa la web sale transformada. A través de estas nuevas capturas, los objetos digitales entran en un espacio masivamente interaccional y guardan la memoria de las manipulaciones. Los nuevos espacios de interacción permiten aislar, rematar y hacer que sobresalga uno u otro objeto, que no hubiera llamado la atención si no hubiera sido tan fácilmente manipulable. La conexión en serie, con un contador o en una lista constituyen operaciones que permiten darle un significado a las fotografías, las *playlists* de música, las grabaciones de desplazamientos o las huellas de comunicación que carecen de significado si las consideramos de manera aislada. A la vez públicos y compartidos, estos nuevos objetos se convierten propiamente en «informacionales» porque están enriquecidos con las operaciones de las que son el soporte: recogida, comparación, medición, puntuación y comentario. Huellas idiosincráticas cuando se aíslan en su contexto, solamente adquieren valor a través de las cualidades que les otorga la puesta en común en los nuevos espacios de cálculo de la web.

#### ENTRAR EN LOS ALGORITMOS

¿Acaso la organización de la web del futuro será inteligible para sus usuarios? Uno de los principales retos de las transformaciones actuales de internet reside en el carácter secreto o público de los cálculos que permiten clasificar los datos digitales y actuar sobre ellos. No cabe ninguna duda de que este cálculo forma parte del futuro de la web. Es inútil oponer una resistencia a esta lógica reanimando la práctica humanista de una información apreciada manualmente por artesanos letrados, situación característica de un mundo de informaciones restringidas, o perpetuando la delegación de autoridad a profesionales, documentalistas, editores y periodistas, encargados de filtrarlos y jerarquizarlos, situación propia del régimen industrial del documento impreso. Con la web de los datos, el cálculo se ha integrado en la producción de las inscripciones, participa en la economía del «neodocumento» (Salaün, 2012) y se ha convertido en un componente reflexivo de las estrategias de los anunciantes. Quizá resulte más pertinente llevar la crítica al interior mismo de las herramientas computacionales de la web interrogando la naturaleza de las operaciones que efectúan y los principios que pueden movilizar para justificarlas. El reto al que se enfrenta la web de los *big data* es el de ganarse la confianza de los usuarios no solamente respecto a la protección de sus datos personales sino también respecto a los procedimientos que le permiten dar sentido a los nuevos flujos de información.

Las ciencias humanas se han acostumbrado a encerrar los algoritmos en una caja negra, sin discusión ni investigación. Estos calculan tontamente, miden sin pensar y aplastan la complejidad de los significados a los que tan solo la intuición humana da sentido. Demasiadas veces, la investigación se interrumpe ante los algoritmos, los idiotiza atribuyéndoles artimañas irracionales —con frecuencia la irracionalidad se confunde con la idea de que las máquinas no pueden calcular «inteligentemente»— e ignora las formas de racionalidad que se deslizan en las máquinas de cálculo oponiéndolas, con ese *pathos* tan característico, a la inefable sutileza del juicio humano. Sin embargo, toda una tradición de investigación ha mostrado el interés que representaba el observar cómo se



han delegado principios y valores a los objetos técnicos, transformándolos en *artefactos cognitivos* encargados de ejecutar los tratamientos y las elecciones, de autorizar y prohibir, de clasificar y orientar (Norman, 1991, págs. 17-38; Winner, 1980). Los parámetros de cálculo, el orden al que llegan y los resultados esperados son el objeto de decisiones. Y los que toman las decisiones desarrollan estrategias que están en deuda con el trabajo de interpretación de las ciencias humanas. Las técnicas de cálculo siempre han estado en el centro de las operaciones políticas más interesantes. Por ejemplo, el sistema de voto, con su cabina electoral y su principio «un hombre, un voto», encierra un algoritmo electoral cuya implementación suscitó miles de discusiones y resistencias (Garrigou, 1992). Todas las formas de organización de la democracia proceden de un sistema de recuento, selección y decisión cuyas reglas son el objeto de arbitrajes apasionados entre orientaciones normativas competidoras. Antes de que la llegáramos a considerar como sencilla, y de una obviedad indiscutible, la aritmética del voto fue denunciada como generadora de efectos injustos, malsanos y reductores. Hizo falta un siglo de debates para aceptar la idea de que la suma aritmética de los votos del mayor número de personas es preferible al voto ponderado de unos pocos. Y basta con observar la manera con la que las diferentes leyes electorales estructuran las formas de la competición política de cada país para medir hasta qué punto las reglas de cálculo constituyen determinantes esenciales para las culturas políticas nacionales. Asimismo, en lugar de reenviar los algoritmos al infierno del frío cálculo y de las máquinas sin alma, es necesario investigar sobre lo que calculan exactamente las métricas de la web, sobre la manera que tienen de calificar a las personas y a sus acciones, sobre el principio que subyace a las operaciones que efectúan, y sobre los comportamientos que incitan (Becker y Stalder, 2009). Hablar de la política de internet sin examinar la estructura de los algoritmos, es como hablar de la democracia sin interesarse por las leyes electorales.

No se trata aquí de enumerar en detalle las diversas familias que se despliegan hoy en la web: las técnicas de aprendizaje del *data mining*, los algoritmos de filtrado colaborativo de la recomendación, los instrumentos de *pageranking* de los buscadores, las opera-

ciones de enriquecimiento de los contenidos de la web semántica, las herramientas de representación cartográfica de la geolocalización, las métricas de reputación y de influencia de la web social, etc. Pero nos gustaría insistir en algunos aspectos críticos de la economía de los cálculos que se asienta en la actualidad en la web de los *big data*. Primero cabe subrayar el altísimo nivel de desconocimiento por parte de los internautas de la naturaleza, e incluso a veces de la propia existencia de las operaciones efectuadas cuando navegan. La web filtra, clasifica, esconde, ordena, recomienda, sintetiza, sugiere... cualquier tipo de información, sin que los usuarios lleguen a calibrar realmente las operaciones efectuadas para alimentar las listas, los indicadores, los contadores, los mapas, las nubes o los botones que dirigen sus desplazamientos en la red. Aprovechando la escasa vigilancia de los internautas frente a la proliferación de las métricas, las plataformas ahorran esfuerzos de aprendizaje y, a menudo, cultivan un alto nivel de confusión con respecto a sus operaciones. No obstante, podemos establecer la hipótesis de que aguijoneada por los vigías críticos y los trabajos de las *web sciences*, se pondrá en marcha rápidamente un bucle de aprendizaje para fomentar el desarrollo de una auténtica *alfabetización* de los internautas respecto al funcionamiento de los algoritmos. A este respecto, varios puntos críticos sobre su funcionamiento merecen especial atención.

El primero tiene que ver con la propiedad privada o pública de los datos sobre los que se ejercen dichos cálculos. Mientras que el *newsfeed* de Facebook ordena las informaciones a partir de datos que pertenecen a una única plataforma, el PageRank de Google clasifica informaciones recogidas en las páginas de la web pública. Con el desarrollo de las técnicas de perfilado del *data mining*, numerosos actores de la web calculan de manera subterránea las huellas de sus clientes para personalizar sus servicios y crear bases de datos con conocimientos a los que jamás tienen acceso.<sup>11</sup> La posesión exclusiva de los datos los exonera de cualquier tipo de explicación sobre las operaciones efectuadas. La privatización de los datos de la web contribuye a la clandestinidad de los cálculos. A la inversa, la reivindicación de apertura de datos invoca el hecho de que permite a un amplio abanico de actores proceder a tratamien-

tos competitivos con aquellas informaciones accesibles. Entonces será posible una especie de competición de heurísticos para atribuir significados diferentes a los datos y ejercer un control sobre las operaciones dirigidas por los demás. A este respecto, la disponibilidad común de los datos constituye la principal condición para la confianza en los algoritmos de la web.

② El segundo punto crítico es la explicitación del perímetro de los datos que entran en el cálculo, ya que no basta con disponer de datos públicos para garantizar la confianza de los algoritmos. También es necesario conocer el espacio de las informaciones que tratan y las categorías que sirven para incluir o excluir ciertos datos. ¿Cómo cerciorarse de que efectivamente Google «crawlea» toda la web? ¿Cómo se categorizan por géneros culturales diferentes aquellos objetos que después van a ser valorizados en las recomendaciones hechas a los usuarios? ¿Acaso el número de *followers* (seguidores) es un buen índice de la reputación de un usuario de Twitter?, ¿no sería más preciso reemplazarlo por el número de *retweets*? En el mundo de las grandes bases de datos, el volumen de informaciones suple a menudo al examen de su representatividad y significatividad (Rogers, 2009, págs. 29-39). Al quedar eximido de los procedimientos de comprobación de la estadística tradicional, no es nada seguro que los modelos de aprendizaje del *data mining* no acaben llevándonos a una multiplicación de las aproximaciones y deformaciones. El examen crítico de los datos y de su modo de categorización constituye la segunda condición del establecimiento de la confianza en los algoritmos de la web.

③ El tercer punto crítico, sin duda el más importante y delicado, tiene que ver con la inteligibilidad de los cálculos que alimentan las métricas de la web. Sin llegar a reivindicar aquí una completa publicidad del funcionamiento de los algoritmos,<sup>12</sup> es importante, sin embargo, que el internauta conozca al menos sus principios generales y el tipo de operación que efectúan sobre las informaciones conseguidas, tal y como hemos intentado hacer durante la primera parte de este texto explicando los principios de editorialización, autoridad, audiencia, afinidad y velocidad. Bajo la forma de una generalidad justificable, estos principios designan el tipo de operación a la que obedece la construcción de diferentes familias de mé-

tricas.<sup>13</sup> Podríamos comprobar si en la realidad de los cálculos las diferentes métricas de la web honran los principios según los cuales son comprendidas por los internautas. Podríamos verificar también si la complejificación de las métricas de la web, especialmente fomentada por la llegada de la web social, ha contribuido a una creciente hibridación de sus principios. Los diseñadores de plataformas mezclan elementos de autoridad, afinidad y velocidad para generar clasificaciones de información cada vez más alejadas de los usuarios. Por otro lado, nada prohíbe que alteraciones de estos principios no sean instauradas por las plataformas introduciendo, en el cálculo «natural» que pretenden efectuar, informaciones controladas por su política comercial. El debate sobre la neutralidad de los algoritmos tan solo se centra en las herramientas de *search* (Pasquale, 2010; Introna y Nissenbaum, 2000). Sin duda sería necesario extenderlo a muchas otras métricas de la web. ¿Quién comprueba si los algoritmos de filtrado colaborativo de la recomendación cultural son «filtrados» por los contratos comerciales que las plataformas pueden tener con algunos editores? ¿Quién nos garantiza que el número de páginas visitadas de los vídeos no serán compradas por agencias de comunicación para fomentar una campaña de marketing? ¿Acaso los *trending topics* de Twitter reflejan exactamente el comportamiento de los usuarios o son el objeto de una moderación políticamente orientada como parecen indicar las investigaciones recientes? (Gillespie, 2011). El desarrollo de una cultura crítica de los algoritmos parece cada vez más necesaria e invita a las ciencias sociales a no dejar únicamente en manos de los matemáticos el monopolio de la atribución de significado al nuevo mundo de los datos digitales.

#### NOTAS

1. La tipología presentada en este texto es la síntesis de un trabajo colectivo efectuado en el marco del proyecto «Algoritmo y sociedad» del laboratorio SENSE de Orange Labs. Está en deuda con los debates y las ideas de Christophe Aguiton, Thomas Beauvisage, Jean-Samuel Beuscart, Patrick Launay, Kevin Mellet, Véra Nikolski, Valérie Peugeot, Anne-Sylvie Pharabod y Marie Trespeuch.

2. Aunque en la materia cualquier tipo de medición resulta imposible, ya que supondría que conociésemos el tamaño de la web, se calcula que el 95% de las audiencias se concentran en el 0,03% de los contenidos. Véase Fouetillou (2012).

3. Véase Siegler (2010). Sobre la medición del tamaño de la web, consultar la página web World Wide Web Size: <<http://www.worldwidewebsize.com/>>.

4. De una encuesta de 2005 se desprende que solamente el 38% de los internautas americanos eran capaces de distinguir la clasificación «orgánica» de las páginas web que contestaban a su petición, de los enlaces comerciales asociados a estas. Véase Fallows (2005).

5. Roland Barthes (1980) utiliza el término «huella» para designar a lo que está por debajo del índice, que no suscita interés y que escapa a cualquier significado.

6. Este mismo procedimiento se encuentra en el centro de la elaboración de las técnicas de *data mining*, que no abordamos aquí directamente para concentrarnos en el trabajo de transformación de las huellas en indicios a través del establecimiento de métricas que, la mayoría de las veces, son públicas y compartidas. Pero es evidente que estas huellas convertidas en datos alimentan también, y a menudo al mismo tiempo, los cálculos que las empresas efectúan de manera subterránea sobre los perfiles y las actividades de sus clientes. Véase, de modo más general acerca de esta cuestión, Rouvroy y Berns (2010).

7. Por ejemplo: <<http://clickflashwhirr.blogspot.com/>>.

8. Una de las explicaciones de esta inversión la encontramos en el hecho de que en el mundo digital, los costes de coordinación se reducen considerablemente, y son tendencialmente nulos cuando esta se encomienda a las capacidades de autoorganización de los usuarios. Asimismo, resulta bastante más fácil utilizar los datos que ya están ahí. Clay Shirky nos da un ejemplo de ello, las fotos Flickr de los desfiles en las calles de Londres. Ningún fotógrafo profesional es capaz de generar un conjunto diversificado de fotos a bajo coste, mientras que usando un *tag*, que permite agregar fotos, los usuarios pueden producir una representación de conjunto del evento a partir de producciones individuales cuyo objeto no era la realización de una obra colectiva. «The basic capabilities of tools like Flickr reverse the old order of group activity, transforming “gather, then share” into “share, then gather”.» Véase Shirky (2008).

9. Acerca de estas formas de interpretación que emergen *ex post* del tratamiento que le dan los internautas a los datos públicos, véase la noción «democracia cooperativa» en Noveck (2009).

10. Anderson, Chris, «The End of Theory. Will the Data Deluge

Makes the Scientific Method Obsolete?», *Edge* (2008) citado en Boyd y Crawford (2011). Véase también Manovich (2011).

11. Para más información sobre este tema consultar Rouvroy y Berns (2010, págs. 88-103).

12. Este punto es motivo de debate en torno al PageRank y al abordar la *search neutrality*. Existen numerosos argumentos convincentes para proteger parcialmente el secreto en torno al algoritmo de Google para proteger la web contra las estrategias oportunistas de los internautas. Véase Grimmelmann (2010), págs. 435-459.

13. Podríamos hacer que estos cinco principios correspondan con la gramática de las ciudades desarrollada por Luc Boltanski y Laurent Thévenot en *Les économies de la grandeur*, asociando la editorialización con la ciudad *inspirada*, la autoridad con la ciudad *cívica*, la audiencia con la ciudad *mercantil*, la afinidad con la ciudad *doméstica* y la velocidad con la ciudad *industrial*. Véase Boltanski y Thévenot (1991).

#### BIBLIOGRAFÍA

- Aguiton, Christophe y Cardon, Dominique (2007), «The Strength of Weak Cooperation: An attempt to Understand the Meaning of Web 2.0», *Communications & Strategies*, n° 65, 1<sup>st</sup> quarter, págs. 51-65. Disponible en: <<http://ssrn.com/abstract=1009070>>.
- Balnaves, Mark y Willson, Michele (2011), *A New Theory of Information and the Internet, Public Spheres meets Protocol*, Nueva York, Peter Lang.
- Barthes, Roland (1980), *La chambre claire. Note sur la photographie*, París, Gallimard (trad. cast.: *La cámara lúcida: nota sobre la fotografía*, Barcelona, Paidós, 2010).
- Battelle, John (2005), *The Search. How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*, Londres, Portfolio/Penguin Group (trad. cast.: *Buscar: cómo Google y sus rivales han revolucionado los mercados y transformado nuestra cultura*, Barcelona, Tendencias, 2006).
- Beauvisage, Thomas (2007), «Mesurer l'activité sur ordinateur: enjeux et méthodes», *Réseaux*, vol. 25, n°s 145-146, septiembre-diciembre, págs. 45-80.
- Becker, Konrad y Stalder, Felix (dirs.) (2009), *Deep Search. The politics of Search Beyond Google*, Innsbruck, StudienVerlag.
- Benkler, Yochai (2009), *La richesse des réseaux. Marchés et libertés à l'heure du partage social*, Lyon, Presses Universitaires de Lyon.

- Beuscart, Jean-Samuel; Cardon, Dominique; Prieur, Christophe y Pisard, Nicolas (2009), «Pourquoi partager mes photos de vacances avec des inconnus? Les usages de Flickr», *Réseaux*, n° 154, págs. 91-129.
- Boltanski, Luc y Thévenot, Laurent (1991), *De la justification. Les économies de la grandeur*, París, Gallimard.
- Bourdieu, Pierre (dir.) (1965), *Un art moyen. Essai sur les usages sociaux de la photographie*, París, Minuit (trad. cast.: *Un arte medio: ensayo sobre los usos sociales de la fotografía*, Barcelona, Gustavo Gili, 2003).
- Boyd, Danah y Crawford, Kate (2011), «Six Provocations for Big Data», *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, septiembre de 2011. Disponible en SSRN: <<http://ssrn.com/abstract=1926431>>. (Traducción francesa sur Internet. actu: <<http://www.internetactu.net/2011/09/23/big-data-la-necessite-d%E2%80%99un-debat/>>).
- Cardon, Dominique (2008), «Le design de la visibilité. Un essai de cartographie du web 2.0», *Réseaux*, n° 152, págs. 93-137.
- , (2010), *La démocratie Internet. Promesses et limites*, París, Seuil.
- , Heurtin, Jean-Philippe y Lemieux, Cyril (1995), «Parler en public», *Politix*, n° 31, 3<sup>r</sup> trimestre, págs. 5-19.
- , y Levrel, Julien (2009), «La vigilance participative. Une interprétation de la gouvernance de Wikipédia», *Réseaux*, n° 154, págs. 51-89.
- Donath, Judith (2007), «Signals in social supernets», *Journal of Computer-Mediated Communication*, vol. 13, n° 1.
- Fallows, Deborah, «Search Engine Users», *Pew Internet & American Life Project*, 23 de enero de 2005, <<http://www.pewinternet.org/Reports/2005/Search-Engine-Users.aspx>>.
- Fouetillou, Guilhem «Les mécaniques de l'influence en période de campagne présidentielle», *SocialMediaWeek*, 16 de febrero de 2012, París, Sciences Po.
- Galloway, Alexander R. (2004), *Protocol. How Control Exists after Decentralization*, Cambridge, The MIT Press.
- Garrigou, Alain (1992), *Le vote et la vertu, comment les Français sont devenus électeurs*, París, Sciences Po.
- Gillespie, Tarleton, «Can an algorithm be wrong? Twitter Trends, the specter of censorship, and our faith in the algorithms around us», *Culture Digitally*, 19 de octubre de 2011, <<http://culturedigitally.org/2011/10/can-an-algorithm-be-wrong/>>.
- Grimmelmann, James (2010), «Some Skepticism About Search neutrality», en Szoka, Berin y Marcus, Adam (comps.), *The Next Digital Decade. Essays on the Future of the Internet*, Washington, TechFreedom.
- Introna, Lucas D. y Nissenbaum, Helen (2000), «Shaping the Web: Why the politics of search engines matters», *The Information Society*, vol. 16, págs. 169-185.
- Jouët, Josiane (2004), «Les dispositifs de construction de l'internaute par les mesures d'audience», *Le Temps des Médias*, n° 3, vol. 2, págs.160-174.
- Lemieux, Cyril (2009), *Le devoir et la grâce. Pour une analyse grammaticale de l'action*, París, Economica.
- Leskovec, Jure; Huttenlocher, Daniel y Kleinberg, Jon (2010), «Governance in Social Media: A Case Study of the Wikipedia promotion process», *AAAI International Conference on Weblogs and Social Media (ICWSM)*.
- Levy, Steven, «How Google's Algorithm Rules the Web», *Wired*, 22 de febrero de 2010, <[http://www.wired.com/magazine/2010/02/ff\\_google\\_algorithm/all/1](http://www.wired.com/magazine/2010/02/ff_google_algorithm/all/1)>.
- , (2011), *In the Plex. How Google Thinks, Works and Shapes our Lives*, Nueva York, Simon & Schuster.
- Manovich, Lev (2011), «Trending: The Promises and the Challenges of Big Social data», en Gold, M. K. (comp.), *Debates in the Digital Humanities*, Mineápolis, University of Minnesota Press.
- Méadel, Cécile (2010), *Quantifier le public. Histoire des mesures d'audience à la radio et à la télévision*, París, Economica.
- Merzeau, Louise (2009), «Du signe à la trace: l'information sur mesure», *Hermès*, n° 53, págs. 23-31.
- Norman, David A. (1991), «Cognitive Artifacts», en Carroll, J. M. (comp.), *Designing Interaction: Psychology of the Human-Computer Interface*, Nueva York, Cambridge University Press.
- Noveck, Beth Simone (2009), *Wiki Government. How technology can make government better, democracy stronger, and citizens more powerful*, Washington, Brookings Institution Press.
- Pariser, Eli (2011), *The Filter Bubble. What the Internet is Hiding from You*, Nueva York, The Penguin Press.
- Pasquale, Franck (2010), «Dominant Search Engines: An Essential Cultural & Political Facility», en Szoka, Berin y Marcus, Adam (comps.), *The Next Digital Decade. Essays on the Future of the Internet*, Washington, TechFreedom, págs. 401-417.
- Rebillard, Franck (2007), *Le Web 2.0 en perspective. Une analyse socio-économique de l'Internet*, París, L'Harmattan.
- Rogers, Richard (2009), «Post-Demographic Machines», en Dekker, Annet y Wolfsberger Annette (comps.), *Walled Garden*, Ámsterdam, Virtual Platform.
- Rosa, Harmut (2010), *Accélération. Une critique sociale du temps*, París, La découverte.

- Rouvroy, Antoinette (2009), «Détecter et prévenir: de la digitalization des corps et de la docilité des normes», en Lebeer, Guy y Moriau, Jacques (dirs.), *(Se) gouverner. Entre souci de soi et action publique*, P.I.E. Peter Lang.
- Rouvroy, Antoinette y Berns, Thomas (2010), «Le nouveau pouvoir statistique. Ou quand le contrôle s'exerce sur un réel normé, docile et sans événement car constitué de corps numériques», *Multitudes*, n° 40, invierno de 2010, págs. 88-103.
- Salaün, Jean-Michel (2012), *Vu, Su, Lu. Les architectes de l'information face à l'oligopole du web*, París, La découverte.
- Shirky, Clay (2005), «Ontology is Overrated: Categories, Links and Tags», *Clay Shirky's Writings About the Internet*, <[http://shirky.com/writings/ontology\\_overrated.html](http://shirky.com/writings/ontology_overrated.html)>.
- , (2008), *Here Comes Everybody. The Power of Organizing without Organizations*, Nueva York, The Penguin Press.
- , (2009), «A Speculative Post on the Idea of Algorithmic Authority», *Clay Shirky Blog*, 15 de noviembre, <<http://www.shirky.com/weblog/2009/11/a-speculative-post-on-the-idea-of-algorithmic-authority/>>.
- Siegler, M. G., «Eric Schmidt: Every 2 Days We Create as Much Information as We Did up to 2003», *TechCrunch Blog*, 4 de agosto de 2010, <<http://techcrunch.com/2010/08/04/schmidt-data/>>.
- Vaidhyanathan, Siva (2011), *The Googlization of Everything (and why we should worry)*, Berkeley, University of California Press.
- Vanbremeersch, Nicolas (2009), *De la démocratie numérique*, París, Seuil.
- Wasik, Bill (2009), *And Then There This. How Stories Live and Die in Viral Culture*, Viking Book.
- Weinberger, David (2007), *Everything is Miscellaneous. The Power of the new Digital Disorder*, Nueva York, Times Books.
- , (2011), *Too Big to know. Rethinking knowledge now that the facts aren't the facts, experts are everywhere, and the smartest person in the room is the room*, Nueva York, Basic Books.
- Winner, Langdon (1980), «Do Artifacts Have Politics?», *Daedalus*, vol. 109, n° 1.
- Zittrain, Jonathan (2008), *The Future of Internet. And How to Stop it*, Nueva York, Yale University Press.