

Extractivismo de datos y correo electrónico, orientaciones para decidir qué pastilla tomar

Marianicer Figueroa Agreda/ABACOenRED
Comunidades Técnicas de Asistencia Mutua /
Proyecto de Articulación Regional Rosa Luxemburg Stiftung
México, Centroamérica y Caribe

La opinión generalizada es que, si bien el uso de buscadores en Internet es una de las aplicaciones de servicio en la red de redes más utilizada, el "correo electrónico" es el que le sigue como actividad en línea al que se accede en una buena magnitud, para la mayoría de lxs¹ usuarixs con vida digital, lo que ha sido evidenciado en los informes emitidos años tras año de Pew Internet Report, los que en el 2010 confirmaban que el correo electrónico sigue siendo, junto con el uso de buscadores, las actividades en línea primarias en Internet. Por otra parte, conforme a las encuestas de Radicati, a nivel mundial, se envían 269 mil millones de correos electrónicos cada día y actualmente hay algo más de 3,7 mil millones de usuarixs de correo electrónico en todo el mundo.²

Este enorme flujo de comunicaciones, cuantificadas en miles de datos, cobra un valor económico incalculable en el mundo de lo que muchos ya denominan la Economía de datos, entendiendo por este el mercado que está siendo generado por empresas de telecomunicación que utilizan la minería de datos, para recopilar y gestionar grandes cantidades de datos en aras de extraer de ellos información útil, tal como lo revela el informe "Big & Fast Data: The Rise of Insight-Driven Business". Siendo considerado este un mercado emergente que necesariamente debe ajustarse a los marcos regulatorios mundiales y nacionales sobre la protección de datos personales, parece que el oro negro y liquido den esta materia surge de aquellos servicios que con su cara de gratuidad, recopilan inconsulta y silenciosamente millones de datos a espaldas del consentimiento de sus usuarixs, tal como se ventiló en el 2010, cuando un par de abogados en Texarkana, Texas, develaron en una instancia legal que llegó a convertirse posteriormente en una demanda colectiva, lo que para ellos fue descubrimiento trascendental: los anuncios en Gmail están correlacionados con las palabras clave contenidas en los correos electrónicos. Los resultados de este litigio son largos de contar, sin embargo este proceso develó que desde septiembre de 2002, que varios ingenieros de Google presentaron una serie de solicitudes de patente en la que se describía las posibilidades de acceso de una larga serie de atributos de los mensaje "internos" y "externos" de quienes utilizaran sus servicios, que en cualquier

¹ En aras de acercarme a un lenguaje de género inclusivo, en este texto utilizó la letra x para indicar la no definición binaria de género. Este posicionamiento no sólo lo hago como una forma para cuestionar el lenguaje, sino como respuesta subversiva ante las formas jerárquicas y piramidales, en las relaciones sociales (Silva Segovia, Jimena, 2015) De igual manera, trato en la medida de lo posible asumir las formas de citar de las hermanas feministas quienes visibilizan nombre y apellido de lxs autores, en aras de dejar ver explícitamente cuando los aportes son de mujeres.

² Los hallazgos revelados en el estudio de febrero de 2017 de Radicati Group , marcan un crecimiento continuo en la cantidad de correos electrónicos enviados por día; De la cifra de 2015 de 205 mil millones de correos electrónicos. Para la escala, si 205 millones de correos electrónicos es el recuento diario de correos, esto significa que casi 2.4 millones de correos electrónicos se envían cada segundo y unos 74 billones de correos electrónicos al año.

combinación servirían para extraer el significado de un correo electrónico como oportunidad para seleccionar los mejores anuncios y mensajes publicitarios que en consonancia le sirvieran como solicitud a la carta.

El camino transitado desde esa fecha por Google esta caracterizado por una puesta en marcha una carrera constante de innovaciones en su tecnología de extracción de datos con fines lucrativos, de manera que la monetización de Gmail cobrara la rentabilidad que se espera. En general el valor particular en estas acciones está en la creación de perfiles de usuarios que hacen que su orientación sea más efectiva para el mundo de la publicidad y el mercadeo. El asunto quedo muy explícitamente develado, cuando en agosto, la jueza Lucy Koh del Tribunal de Distrito de los Estados Unidos para el Distrito Norte de California ordenó la publicación de una serie de correos electrónicos intercambiado por los empleados de Google en 2009, en los que se discutía el extraordinario poder de los algoritmos de perfiles de usuarios de Gmail, lo que dejó ver que la "publicidad basada en el interés" (IBA) de AdSense pudo agrupar a los usuarios en solo 700 o más categorías según los sitios web que navegan. Pero Gmail, utilizando algoritmos de clasificación como PHIL y sin duda otros también, dividiría a los usuarios literalmente en "millones de cubos". El valor de dicha "extracción del perfil de usuarios de Gmail" para la "segmentación basada en el usuarios" en otros servicios basados en anuncios de Google como AdSense era obvio para estos empleados, quienes discutieron con entusiasmo las nuevas oportunidades de segmentación que posibilitaría. Al respecto ya en el 2014, después del cambio en la bandeja de entrada de Gmail en donde apareció la pestaña promociones, también se evidenció cómo la granularización del perfil de usuarios, permitiría enviar a los usuarios según las estimaciones de ingresos del hogar derivadas de los datos del IRS publicados por código postal como los anunciantes apuntar al estado paterno, es decir, si el usuario tiene hijos en casa, entre miles de atributos con los cuales se logra esa segmentación. Una excelente y descriptiva explicación del histórico de acciones que Google ha probado y ha activado desde el punto de vista tecnológico para eso, la pueden conseguir en el artículo de [The Natural History of Gmail Data Mining Gmail isn't really about email—it's a gigantic profiling machine](#) escrito en el 2014 por Jeff Gould

Ahora bien, más allá de la forma cómo la extracción de datos de aquello que en nombre de la gratuidad sucede con el correo electrónico de Gmail, lo que ratifica que en esa gratuidad, el negocio son los datos que aportamos para alimentar el mismo, se hace más preocupante el hecho que con los mismos algoritmos que nos describen y etiquetan como usuarios, sirve de igual manera para dos tareas distintivas y a la vez vinculante entre ambas: el control biopolítico como forma para mantener las relaciones de poder y la exclusión histórica de la subalternidad, de lo que ya hemos hablado en anteriores entregas, y el espionaje evidenciado en los informes que Snowden publicó en el 2013 acerca de las formas en que la Agencia Nacional de Seguridad (NSA) de EEUU y sus franquicias en el extranjero han dominado la tecnología de "big data" para en nombre de la seguridad de estado se apoderen de nuestros metadatos y monitorear nuestros flujos de clic .

Ante estas realidades la pregunta que surge siempre es cómo escapar de esta especie de panóptico digital, cuando buena parte de las acciones que hacemos en los diferentes roles que tenemos como ciudadanxs, se ven

sustentadas en el uso de Internet, más aún cuando paradójicamente uno de los indicadores de los objetivos de desarrollo sustentable de la UNESCO expresos en la Agenda 2030 para el Desarrollo Sostenible, aboga por el deber de los Gobiernos por ejercer un mayor liderazgo para fomentar la conectividad a Internet de sus pueblos.

Esta pregunta, compartida entre muchxs quienes sufrimos la angustia existencial de saber que al parecer mientras estemos conectadxs somos parte del gran hermano de estos tiempos, podríamos asumir como salida para seguir contando con comunicaciones por email, montarse un servidor de correo en casa o afiliarse a lxs de una comunidad tecnológica que nos asegure privacidad y la no extracción de nuestros datos, como en efecto sucedió con muchas alternativas que en estos últimos 15 años aparecieron al respecto. Pero la realidad es que en nombre del resguardo de una estructura que se presta muy bien a la centralización de los datos y como forma de protección de los intereses hegemónicos evidentemente monopólicos de las grandes transnacionales de la red, a necesitarse cada vez más elementos, en los últimos años se ha elevado la cantidad y calidad de los requerimientos exigidos por los grandes proveedores de correo, para que los mensajes que enviemos de correos alternativos no sean calificados de potencialmente peligroso por ellos. Tal como lo dice Iglesias Pablo:

“...esto nos lleva a que por facilidad, por funcionalidad, y casi por necesidad, la mayoría de usuarixs utilizamos correos que nos ofrecen “gratuitamente” empresas como Google, Microsoft, Apple,... E incluso aquellos correos que tenemos con nuestro propio dominio los sincronizamos con estas plataformas”. (2017, s/p)

Pero tal como lo canta la negra Mercedes Sosa, no todo está perdido cuando recordarnos la batalla que muchxs están dando para mantener servicios de correo en los que se respete la privacidad de sus usuarixs, tal como siguen dando la pelea cooperativas proveedores de internet como [Riseup](#), [Código Sur](#), y [Primero de Mayo/Enlace del Pueblo](#) . De igual manera esperanza la tarea persistente de muchxs de seguir trabajando para devolverle a Internet su condición descentralizada, más aún cuando las tecnologías P2P y de Blockchain cada día limpian el camino para visualizar la red distribuida a la que muchos apostamos.

Como forma de cerrar esta entrega esperanzadamente, compartimos otras alternativas a la mano con versiones gratuitas y versiones premium pagas de correo electrónico encriptados, recomendamos revisar

- [ProtonMail](#): con sede en Suiza, es uno de los servicios de correo electrónico más conscientes de la privacidad, y ofrece correos electrónicos encriptados y autodestructivos. Sin embargo, solo proporciona 500 MB de almacenamiento a lxs usuarixs que no pagan.
- [TutaNota](#) ofrece correos electrónicos encriptados con Almacenamiento de 1GB de serie, actualizable. Es de código abierto, con el código disponible en Github para su inspección. Es ofrecido desde Alemania por lo que promueven que están sujetos a la rigurosa Ley Federal de Protección de Datos de Alemania. Sin embargo, los defensores de la privacidad pueden no estar de acuerdo con que sus datos se almacenen en Alemania, después de que se supo que el Servicio Federal de Inteligencia de Alemania había colaborado con la NSA en sus programas de vigilancia.

- [Mailfence](#): ofrece desde Bélgica correos electrónicos encriptados con almacenamiento de 200 MB de correos electrónicos, 250 MB de documentos como estándar. Al igual que con muchos otros países europeos, Bélgica tiene leyes de privacidad estrictas que favorecen al consumidor en lugar de a la empresa. A diferencia de muchos otros países, hay poca evidencia que sugiera que Bélgica colaboró en los esquemas de vigilancia de la NSA

También, queda la apuesta en el horizonte de la tecnología P2P que al respecto se está levantando, por lo que ya se asoman desarrollos que invitamos a revisar en un [reporte](#) publicado por GitHub que da cuenta de la visión general de los proyectos que trabajan para dar respuestas desde este tipo de tecnología, de la cual resaltamos a [Freenet](#).

Finalmente, a sabiendas que los resultados de las encuestas que a nivel global se desarrollan, dan cuenta del incremento de la preocupación por cuidar la privacidad de los datos, que lamentablemente en esa terrible realidad en la que siendo esclavos de nuestra libertad no realizamos acciones para por lo menos minimizar la cantidad de alimento que cotidianamente damos a lo que finalmente nos controla, toca increparnos entre todos y en modo Matrix, para por lo menos saber consciente y deliberadamente qué pastilla decidiremos tomar para seguir nuestra existencia digital: la azul o la roja

Referencias

- Gould, Jeff (2014) The Natural History of Gmail Data Mining Gmail isn't really about email—it's a gigantic profiling machine . Disponible en <https://medium.com/@jeffgould/the-natural-history-of-gmail-data-mining-be115d196b10>
- Iglesias, Pablo (2017) El internet descentralizado: Estado del arte y alternativas <https://www.pabloyglesias.com/freenet-red-inproxy/>
- Narayan Krishnan (2015) Nuevo estudio global realizado por Capgemini y EMC muestra el Big Data como conductor de la disrupción del mercado. Documento en línea disponible en <https://www.capgemini.com/es-es/news/nuevo-estudio-global-realizado-por-capgemini-y-emc-muestra-el-big-data-como-conductor-de-la/#>
- Radicati Group Email Statistics Report (2017), 2017-2021 Documento en línea disponible en: <http://www.radicati.com/wp/wp-content/uploads/2017/01/Email-Statistics-Report-2017-2021-Executive-Summary.pdf>
- Silva Segovia, Jimena (2015) Con los pies en el barro: Investigar en el contexto de pueblos originarios andinos: un relato autoetnográfico/ia en Silva Segovia, Jimena (Coord) y Javier Bassi. Coords. Aportes teóricos y metodológicos. para una investigación social situada. Antofagasta Universidad Católica del Norte Colombia.