

METODOLOGÍA JURÍDICA DIGITAL: CONCEPTUALIZACIÓN Y PROBLEMÁTICAS PARA SU CONSTRUCCIÓN

Ana Luz Ruelas Monjardín¹

Fecha de publicación: 01/01/2019

Sumario: Impactos de las TIC en el derecho. Desarrollos tecnológicos y lo jurídico. Etnografía digital y etnografía tradicional. La evidencia electrónica. La entrevista a abogados u operadores del derecho. A manera de colofón. Bibliografía.

Resumen: El artículo discute aspectos fundamentales de la metodología jurídica digital que emergen por la recurrencia de fenómenos socio jurídicos que son originados, marcados e impactados por su ocurrencia en el ambiente electrónico, en las redes sociales o Internet. Consideramos que la investigación jurídica en México está emplazada a adoptar las herramientas digitales desarrolladas a partir de las ciencias sociales y humanidades, o bien las que han sido adaptadas concretamente al ámbito jurídico-legal. Aquí se ponen a consideración algunos conceptos atinentes a la investigación jurídica digital como la inteligencia artificial, los algoritmos, las huellas digitales, las bases de datos masivas (*big data*) e Internet de las Cosas. En particular explicamos dos técnicas de investigación que se pueden adaptar a la investigación jurídica: la entrevista y la encuesta en línea.

¹ Profesora de la Universidad Autónoma de Sinaloa. Investigadora Nacional nivel III de México. Este artículo fue redactado durante la estancia de investigación (agosto de 2017 a julio 2018) en la Universidad de Arizona, en el Department of Mexican American Studies, y bajo los auspicios del Conacyt y de la Universidad Autónoma de Sinaloa.

aruelas@uas.edu.mx

Palabras clave: TICs y Derecho, metodología digital, huellas digitales, datos masivos.

DIGITAL LEGAL METHODOLOGY: TOWARDS THE CONSTRUCTION OF A CONCEPT

This article discusses fundamental aspects of digital legal methodology that emerge from the recurrence of socio-legal phenomena that are originated, marked and impacted by their occurrence in the electronic environment, on social networks or the Internet. We believe that legal research in Mexico is set to adopt the digital tools developed from the social sciences and humanities, or those that have been specifically adapted to the legal-legal environment. We review some concepts related to digital legal research such as artificial intelligence, algorithms, fingerprints, big data bases and the Internet of Things are put to consideration. We explain two research techniques that can be adapted to legal research: the interview and the online survey.

Introducción

La metodología de estudio de prácticamente todas las disciplinas científicas se ha visto impactada por la omnipresencia de las tecnologías de información y comunicación (TIC). Son un cisma para la ciencia, para la docencia, para la práctica de la medicina, para la comercialización, para el periodismo, para la industria del turismo, para la seguridad nacional, para el entretenimiento, la comunicación misma, las relaciones interpersonales y, desde luego, para la investigación científica y/o académica, y un largo etcétera.

En este contexto, una preocupación que persiste por el crecimiento desenfrenado del mundo digital es el poderío que tienen los corporativos de TIC en la mercantilización de la información que siembran los usuarios en las distintas plataformas por lo cual debemos entender los conceptos que se han extendido entre algunas disciplinas científicas y que también están impactando al derecho. Aunque se ensayan mecanismos para controlar el uso irrestricto de la información, hasta hoy el freno sobre ellos está en la ley. De hecho, no se ha inventado otra defensa más efectiva que recurrir a la ley, junto a la participación ciudadana.

El descrédito por el que atraviesa el corporativo de Facebook desde 2017, proviene de tres hechos: el incumplimiento de reglas sobre el manejo de bases de datos, la violación de datos personales, o bien la competencia tecnológica. Lo oscuro de la operación de los navegadores y sus aplicaciones, de recomendaciones, de escharbar en los perfiles, de recolectar datos, de empotrarse donde sea a través de los algoritmos, se contrarresta con la transparencia, la difusión, la acción social y el ejercicio de los derechos.

Fue hasta 2017-2018 cuando se despertó un amplio interés popular por saber cómo funcionan las redes digitales, cómo almacenan la información de sus usuarios y para qué la usan. El exrector de la UNAM, Juan Ramón de la Fuente, escribió en marzo de 2018 (De la Fuente, 2018) sobre inteligencia artificial, aludiendo al desconocimiento de nombres de empresas, domicilios, falta de contacto con ellas, etc., pues “nadie sabe a ciencia cierta —dijo— todo lo que ofrecen, donde se ubican, y como es que saben tanto acerca de cada uno de nosotros”, señaló. Los especialistas lo han analizado desde hace años, pero su difusión solo se dio en círculos estrechos, en sus publicaciones,

reuniones o congresos. Por ello, es importante conocer cómo se gestan los datos, qué se hace con ellos y sus impactos sociales.

Para referirnos a los impactos de las TIC en la investigación jurídica, utilizaremos los términos de Internet, metodología digital o etnografía digital, de manera indistinta; precisando brevemente en primer lugar la diferencia sustancial entre Internet y la web.

La web, como una parte de Internet, que contiene inconmensurables volúmenes de información, no es el único camino para compartir información, también se utilizan el correo electrónico, la mensajería instantánea, los grupos de noticias y la transferencia de archivos.

Desde el punto de vista de la etnografía digital, Internet se toma como “el lugar” o intermediario entre los sujetos de investigación, ya sea mediante observaciones, entrevistas o encuestas, y añade efectos particulares a la técnica, aunque se considera una tecnología “menos visible”, como mediador de la vida social. A medida que crece, según Hine (2015), se introduce o empotra en el día a día, a través de las diferentes plataformas, reta los enfoques que se tienen sobre construcción de la identidad, relaciones, cultura y la estructura social. Así, literalmente satura desde hace dos décadas “cada parte de nuestras vidas cívica, profesional y social” (Markham & Baym, 2016).

Impactos de las TIC en el derecho

Las TICs impactan al derecho en diferentes facetas: en la impartición de justicia, para el resguardo de la jurisprudencia y actuaciones judiciales, para la evidencia o prueba procesal, para la transparencia de la institución judicial y las instituciones públicas en general, para la educación legal y, por supuesto, para la investigación jurídica en sus dos variantes metodológicas más importantes: la investigación empírica-cualitativa y la cuantitativa.

De hecho, varios textos jurídicos, o de introducción al derecho, hacen alusión al sacudimiento que han provocado las tic en los diferentes campos o actividades relacionadas con el derecho, tanto para transformar y copiar vastas cantidades de información a formatos electrónicos digitales; almacenar los textos legales históricos y actuales, así como acceder con extrema rapidez a la legislación que se requiera en los tribunales o por los aplicadores del derecho o para docencia e investigación. Lo que anteriormente existía en microfilms o fichas de papel resguardadas físicamente ahora se han descartado para hacer fáciles los sistemas de búsqueda en bases de datos organizadas y ordenadas para diversos efectos. Incluso con el cómputo móvil, laptops o teléfonos celulares, la labor de

litigantes, jueces y profesores se optimiza y, a la vez, es un reto para estos últimos, pues los estudiantes pueden poner en duda fácilmente información manejada por sus profesores. Los jueces prácticamente pueden tener una biblioteca portátil de leyes, tratados, tesis y jurisprudencia.

La tecnología ha desafiado los esquemas rígidos que caracterizan las mentalidades y sistemas jurídicos arraigados por milenios. Tanto los de la ley común (*common law*) como los de la ley civil, “con litigantes renuentes a cambiar los patrones de operación y conductas establecidos, y que no siempre ven hacia adelante”, como dicen Calvi y Coleman (2017) se han visto sacudidos por la era del cómputo y la ubicuidad que brindan computadoras personales y portátiles, teléfonos celulares, tabletas. Hoy los abogados utilizan correos electrónicos, mensajería instantánea o textos, construyen sus páginas personales y, en las redes sociales, anuncian servicios o asesoran clientes. Pero también las salas de los juzgados y tribunales están equipadas con circuitos cerrados y conexiones en red para grabar y transmitir audiencias, o bien para acceder a expedientes en línea. La tecnología electrónica es un elemento omnipresente en las salas de juicios orales.

Las actuaciones judiciales se favorecen por los sistemas de notificación y publicación de acuerdos en línea y sobre todo se ha abierto la esperanza de desechar los voluminosos expedientes de papel cuyo almacenamiento es un problema administrativo descomunal. Además, los litigantes pueden dar seguimiento puntual a sus casos sin tener que acudir a los estrados. La Suprema Corte de Justicia de la Nación de México acertadamente dispone desde mayo de 2005 un canal de televisión restringida que actualmente se transmite en alta definición a través de todos los concesionarios como Canal Judicial.² Dispone en línea de un conjunto de información sobre sentencias y datos de expedientes, estadística judicial, el *Semanario Judicial de la Federación*, una biblioteca digital, la consulta de información electoral, el sistema de consulta de información en materia de comercio exterior, la consulta de información sobre el sistema penal acusatorio, sentencias relevantes del pleno, un catálogo de expedientes históricos (1898-1928), la lista de acuerdos de juzgados y tribunales, entre otros³.

² A la fecha amplió su programación para transmitir además de las sesiones del pleno de la Suprema Corte, las sesiones del Tribunal Electoral del Poder Judicial, y también “series, películas y programas nacionales e internacionales relacionados con el derecho, la cultura, la ciencia y las disciplinas sociales”. <http://bit.ly/2FD4RSp>, Consultado el 26 de febrero de 2018.

³ Véase <http://bit.ly/2EWYIPL>, accesado el 26 de febrero de 2018.

El uso de software e inteligencia artificial para el litigio y para llevar a cabo investigación jurídica en México es tardía. Ello se explica en parte por la resistencia al cambio tecnológico del gremio. Según Tarrin Wills (2017), de la Universidad de Sidney, Australia, adoptar una tecnología dominante para realizar investigación tarda entre cinco y quince años, de acuerdo a “la historia de las humanidades digitales en los campos tradicionales”. Pero también, existen restricciones financieras para adquirir la tecnología o tener acceso a ella. Hoy con los diseños más amigables del software y con interfaces muy manejables, los investigadores jóvenes y los consagrados pueden manejar con más soltura la tecnología sin un entrenamiento previo.

Inteligencia artificial y abogados

La inteligencia artificial se puede aplicar en softwares y repositorios especializados, ya sea como facilitadora en el manejo de archivos, o bien en la clasificación de documentos, aunque hasta hoy se usa solamente en los grandes despachos de abogados que cuentan recursos para ello. En Estados Unidos, se utiliza para realizar actividades propias de ese poderoso gremio, como es la revisión de miles y miles de resúmenes legales, archivos de casos para presentar a los tribunales los antecedentes aplicables. Ello ha desplazado la mano de obra dedicada a esos menesteres, mediante algoritmos de aprendizaje automático, contruidos para analizar enormes montos de información, que en “cuestión de segundos [los algoritmos] revisan documentos que normalmente lleva unas 360 000 horas completar a asistentes legales humanos”⁴.

CaseMine, prestadora de servicios de tecnología legal en India, puede, por ejemplo, sugerir “cambios a un informe legal para que la redacción adquiera un tono duro o ´autoritario””, o bien proponer consultar “documentos adicionales que ayuden a fortalecer los argumentos” (Winick, 2017). También Kira Systems desarrolla plataformas con tecnología de aprendizaje automático para identificar y extraer información de contratos de arrendamiento” y ofrece cincuenta modelos de éstos. Además, si se requiere revisar información adicional o inusual, “el sistema identifica virtualmente cualquier cláusula deseada”, reduciendo drásticamente el tiempo dedicado a esas tareas. También maneja virtualmente carteras de clientes proporcionando información sobre proyectos de manera personalizada.

Un obstáculo que deben superar los desarrolladores de software, de algoritmos y plataformas para servicios legales es la difícil accesibilidad a

⁴ JPMorgan anunció a principios de 2017 que utilizaba el software llamado Contract Intelligence (Coin) para esos efectos.

datos⁵ e informes que generan las instituciones públicas gubernamentales, en las que se requiere alta inversión para digitalizar los que aún están en documentos de papel e impulsar que las nuevas actuaciones en todos los niveles de gobierno se almacenen digitalmente, con estándares de confiabilidad, seguridad y calidad.

Aunque los abogados no se distinguen por adoptar inmediatamente la tecnología, Ziegler⁶ anticipa un ineludible interés: “Creo que el hecho de que los clientes saben que la tecnología puede llevar a cabo muchas de las tareas repetitivas los hará cada vez más reacios a que los abogados se encarguen de ese trabajo” (Winick, 2017). En Estados Unidos, durante 2016, había más de 1 300 000 abogados con licencia para litigar y 200 000 técnicos paralegales. Y se estimaba que el 22 por ciento del trabajo de un abogado y el 35 por ciento de los paralegales podían automatizarse, por lo que los ajustes en las carreras de Leyes, no debían esperar mucho.

Desarrollos tecnológicos y lo jurídico

Internet originalmente se tomó como una red de computadoras para compartir trabajo, pero pronto cambió para incorporar “diferentes tecnologías, capacidades, usos y espacios sociales (Markham & Stavrova, 2016). Hoy, uno dice estoy en línea o estoy en Internet para designar una actividad que puede significar mensajear, ver videos o escuchar música. O puede uno decir también que hizo una encuesta por Internet. Sin embargo, esto conlleva compartir varias facilidades digitales, de redes de comunicación, de desarrollos o innovaciones en software, de comprar servicios de encuestas en línea, etc., a la vez que implica que dejamos huellas digitales que se almacenan en grandes bases de datos y luego se comercializan, y nos catalogan o forman perfiles de usuarios.

A partir de Internet y la web, se han hecho desarrollos tecnológicos fabulosos que parecen no tener límites. Aquí nos vamos a referir a una serie de aplicaciones recientes que vienen transformando la forma como se conservan los archivos, como se comparte el trabajo, como se realiza trabajo de campo y sus implicaciones para el derecho.

La nube y la regulación

⁵ Legal Robot manifestó que, para “entrenar su programa, un equipo de desarrolladores construyó su propia base de datos de términos y condiciones mediante la recopilación de ejemplos de las principales páginas web. Pero eso no fue suficiente: la empresa también necesitó cerrar acuerdos con bufetes de abogados para obtener acceso a sus repositorios privados” (Winick, 2017).

⁶ Adam Ziegler es el director del Laboratorio para la innovación de la Escuela de Leyes de Harvard.

Entre los desarrollos tecnológicos en línea más recientes, que han tenido un impacto tremendo en la administración de justicia, está la conocida nube, donde tecnologías y derecho pueden interconectarse en aspectos,⁷ como:

- a) Regulación de aspectos técnicos y acuerdos de prestación de servicios.
- b) Privacidad y seguridad en la prestación de servicios en la nube.
- c) En manejo de Internet de las cosas para efectos de auditorías, recepción de amenazas o riesgos, fincar responsabilidades y otros asuntos que detonan lo legal.
- d) Aplicaciones del cómputo en la nube en sectores regulados.
- e) Regulación de sistemas y servicios de infraestructura de cómputo en la nube.
- f) Interacción entre la nube y las tecnologías de Internet de las cosas, *big data* y mecanismos de aprendizaje.
- g) Aspectos de descentralización, contenidos, etcétera, del cómputo en la nube (*cloudlets, droplets*).
- h) Compatibilidad de regulaciones sobre la implementación técnica.
- i) Violencia, proliferación de basura (*spam*), *phishing*⁸, ataques cibernéticos en la nube e Internet de las cosas.
- j) Encriptación y seguridad.
- k) Vigilancia a través de la nube.
- l) Derechos humanos, discriminación, privacidad y ejercicio de poder en la nube.
- m) Interacción en la nube entre consumidores y prestadores de servicios (empleo).
- n) Decisiones tomadas a través de algoritmos y sistemas automatizados.

En materia de política regulatoria adquieren relevancia los desarrollos tecnológicos también relacionados con aspectos jurídicos. Así surgen

⁷ IEEE 3rd International Workshop on legal and technical issues in Cloud Computing and the Internet of Things 2017, de <http://www.claw-workshop.org>, Recuperado el 11 de febrero de 2018.

⁸ Es una práctica truculenta de enviar mensajes a correos electrónicos usando nombres de instituciones reputadas para inducir a las personas a revelar información personal como claves para acceder a tarjetas de crédito.

preocupaciones sobre cómo se ha dado la convergencia entre medios de comunicación tradicionales como la radiodifusión y la prensa con los medios digitales que retan los marcos normativos; o bien cómo los grandes intermediarios digitales, como Google, Yahoo, Facebook y Amazon, son actores centrales en asuntos de privacidad, seguridad, manejo de datos personales, etcétera.

De particular interés es la seguridad digital con implicaciones jurídicas por el manejo individual de las redes sociales y su regulación. Nos referimos a las noticias falsas que influyeron en la elección presidencial de Estados Unidos en 2016-2017. Sus implicaciones son de dos tipos: para la cultura de los usuarios de Internet en el manejo de su privacidad y para la regulación jurídica. En materia legal se detonaron procedimientos trascendentes, pues ante el diferendo de tal envergadura, el recurso jurídico puede dar certeza y confiabilidad, interviniendo en 1) aspectos de la violación de las reglas de privacidad de Facebook; 2) la contratación de firmas forenses digitales para determinar si los datos recopilados⁹ de las personas y usados con propósitos de influenciarlos se habían borrado o no; 3) la investigación por parte de la Comisión Federal de Comercio de Estados Unidos sobre si Facebook había quebrantado un acuerdo de consentimiento de 2011 sobre privacidad de datos de usuarios. Además, la fiscalía general de Massachusetts abrió una investigación al respecto, mientras el Comité de Comercio del Senado solicitó a Facebook una explicación de cómo se recopilaron los datos.¹⁰

Cabe señalar que el origen de este caso y su investigación criminal provienen del Departamento de Justicia, la Cámara de Representantes y el Oficina Federal de Investigación sobre lo ocurrido en las elecciones para presidente de Estados Unidos en 2016¹¹.

⁹ En 2015 por el profesor Aleksandr Kogan de la Universidad de Cambridge y entregados a Cambridge Analytica.

¹⁰ El 10 y 11 de abril de 2018, Mark Zuckerberg de Facebook, compareció ante el Congreso estadounidense. También, el Comité de Asuntos Digitales, Cultura, Medios de Comunicación y Deportes de la Cámara de los Comunes británica solicitaron también a Zuckerberg que compareciera ante el Parlamento para evaluar si Facebook había comprometido la privacidad de sus usuarios.

¹¹ La Oficina Federal de Investigaciones (Federal Bureau of Investigation o FBI), la Agencia Central de Inteligencia (Central Intelligence Agency, CIA) y la Agencia Nacional de Seguridad, “concluyeron que el gobierno ruso fue responsable de hackear al Partido Demócrata mediante el envío de miles de correos electrónicos y otros documentos para dañar la campaña de Clinton”. Ello, hasta hoy, ha afectado a la administración pública del presidente Donald Trump por los despidos y responsabilidades asignadas a funcionarios y una inestabilidad permanente en el equipo de gobierno de ese país.

Y, aunque no son una rama de los poderes públicos, la labor de los periódicos fue crucial para activar esos procedimientos. *The New York Times* y Canal 4 del Reino Unido realizaron investigaciones dadas a conocer en febrero de 2018.¹²

El uso de los robots para desestabilizar naciones enteras es preocupante y refleja la evolución perversa en el manejo de las redes digitales a escala transnacional. Un robot de Internet (*web bot* o simplemente *bot*) es una aplicación de software que lleva a cabo tareas automáticamente (*scripts*) en la red que se despliegan de manera estructural a velocidades mucho más altas que las que puede desarrollar una persona. Se usan principalmente mediante los rastreadores web (*web crawler*), en los que un guion automatizado recupera, analiza y ordena información localizada en los servidores web. De hecho, más de la mitad del tráfico en la web la llevan a cabo los robots¹³.

Ahora abordaremos los conceptos indispensables de la metodología social en línea, para posteriormente analizar las técnicas de investigación más importantes para el derecho y que se han adaptado al ambiente digital.

Metodología social digital

Para la metodología social digital son válidas las técnicas de la metodología social tradicional como la observación participante, la entrevista, la encuesta, la etnografía, las bases de datos, con el distintivo de que los fenómenos o conductas investigadas son mediadas por las tecnologías de comunicación digital. En un recuento de 2017 sobre la metodología digital, el autor Keith Hampton (2017) se hace cuatro preguntas para distinguirlas de la metodología tradicional:

1. Si la metodología basada en tecnologías digitales ¿se refiere solamente a la observación y a las técnicas de investigación tradicionales en línea?
2. ¿Cómo medir la exposición a las tecnologías en línea?
3. ¿Si las encuestas pueden ser reemplazadas por las bases de datos masivos (*big data*) o huellas en los datos (*trace data*)? y

¹² Informaron de la revelación de Facebook sobre evidencias de la interferencia rusa, concretada a través de propaganda comprada, enfocada a parámetros de raza, derechos de los homosexuales, control de armas e inmigración, misma que se diseminó entre junio de 2015 y mayo de 2017 y fueron relacionadas con 470 cuentas catalogadas como falsas. Shane, Scott y Goel, Vindu, “Fake Russian Facebook accounts bought \$100,000 in political ads”, *The New York Times*, 6 de septiembre 2017, <https://nyti.ms/2f4Q0nB>

¹³ Google.docs, 2018, enero 3, tomado de <https://bit.ly/11QGJzj>

4. ¿Cuál es la validez de los experimentos en línea? Es decir, si son tan válidos como los realizados fuera de línea.

En el desarrollo del artículo, esclareceremos éstas y otras preguntas.

Los académicos mexicanos, podríamos utilizar la etnografía digital, tal como hicieron los primeros “experimentadores” de la vida en línea (*online*) y fuera de línea (*offline*), quienes analizaron, con participantes de carne y hueso, las diferencias que tenían sobre la percepción de la identidad de género y pertenencia a la raza en espacios análogos y físicos, así como las diferencias de interactuar en un lugar no físico o virtual (Hampton, 2017), es decir, investigaciones que eran una extensión natural de la tradición etnográfica de los antropólogos y sociólogos

Etnografía digital y etnografía tradicional

El referente de la etnografía tradicional siempre está presente en la etnografía digital. Por lo que es preciso establecer sus diferencias, siguiendo al mismo Keith Hampton.

La práctica de la etnografía tradicional se enfoca preponderantemente en la investigación de campo, donde el investigador observa y participa de manera abierta o encubiertamente, tomando información de múltiples fuentes tanto de individuos como de instituciones¹⁴. Se busca conocer experiencias y entender el fenómeno social en sus manifestaciones *in situ* (Madden, 2010). Una característica propia de la etnografía es la cercanía y prolongación del contacto entre el investigador y su objeto de investigación. Esta es la diferencia principal entre ambas.

Por su parte, la etnografía en línea (o ciberetnografía, etnografía en red, etnografía virtual) refieren a una adaptación de la etnografía tradicional al medio virtual. De hecho, el factor digital, además de figurar como un intermediario o mediador entre el objeto observado y el observador, le adhiere connotaciones que surgen por el hecho de la intermediación digital.

¹⁴ Para autores de textos escolares sobre el sistema jurídico estadounidense como Calvi y Coleman, las fuentes primarias de una investigación jurídica son la constitución nacional y la de los estados, las reformas legislativas, las decisiones escritas tomadas en las cortes, las regulaciones administrativas, las reglas procedimentales, y cualquier estatuto constitucional que provenga del *common law* inglés. Se advierte que todas las jurisdicciones de Estados Unidos, excepto la del estado de Luisiana, que tiene una fuerte ascendencia de la ley civil, han adoptado el *common law* como la base de su jurisprudencia. Ésta es la tradición jurídica desarrollada en Inglaterra, donde la ley se constituye a base de las decisiones tomadas por los jueces en ausencia de una ley escrita. Tales decisiones sirven como precedentes convirtiéndose en “comunes” para todo el Reino Unido (Calvi y Coleman, 2017, p. 387).

Así, el espacio virtual, influye o modifica la percepción del observado o entrevistado; percepciones a las que el investigador no puede abstraerse.

Realizar investigación etnográfica y/o cualitativa es una necesidad inaplazable para los académicos del derecho para así dar cuenta integralmente de los fenómenos sociales convulsos que impactan al derecho en México. No hay que olvidar que el derecho es una disciplina humanista que se nutre de y a la vez incide en la cultura de un país, de sus cotidianidades y tendencias que pueden afianzar las instituciones públicas o bien socavarlas. Ello obliga a escrutar al derecho en su accionar con las metodologías y herramientas digitales novedosas: la encuesta en línea, los registros o datos digitales judiciales, los archivos de instituciones públicas, las evidencias generadas a través de artefactos electrónicos, la entrevista mediada electrónicamente, entre otras.

Las bases de datos masivos (big data) como objeto de la etnografía digital

A partir de la explosión de datos masivos, los archivos documentales han adquirido preeminencia en la investigación reciente. Para el análisis etnográfico de mensajes electrónicos, deben hacerse registros diarios o de ciertos momentos significativos en los que se escriben. Para Evans y Aceves (2016), sin embargo, “cuando se adopta el enfoque hacia el documento, texto, video o fotografía, o sea que no es revisado en tiempo real, cesa el enfoque etnográfico y se puede adoptar el análisis del discurso, la conversación o contenido, que son métodos con su propia tradición y principios”.

Cualquier método digital debe tomar en cuenta que las conductas observadas, los entrevistados o los encuestados son sujetos de interacciones duales: en línea y fuera de línea.

Las bases de datos masivas han provocado un sisma en la investigación científica. Son el equivalente a los archivos o fuentes para la investigación e implican un reto tremendo para la investigación cuantitativa, por la generación de enormes cantidades de datos por las actividades en línea: los mensajes escritos y grabados, los videos, los comentarios, los likes, tuits, las fotografías, los memes, etc. Es decir, nos encontramos frente a datos masivos e información que pareciera ser inconmensurable, lo que obliga a segmentarla y trabajar a veces con hipótesis con mayor grado de generalización.

Un aspecto a cuidar ante las bases de datos masivas es la validez de la información. Para ello existen aplicaciones en la web que proporcionan a la

investigación social medios de verificación confiables para detectar, almacenar y codificar información para los académicos que a veces requieren alguna intervención personal. Por ejemplo, para armar una base de datos extraída de la *World Wide Web*, de Facebook, de Twitter, Instagram, Youtube, de la nube, etc., se pueden contratar los servicios de compañías especializadas en validar información, ya que son profesionales que explotan los datos, escurbando en la *World Wide Web* cualquier tipo de información que haya quedado sembrada en diversos formatos que se convierten luego en fuentes de información¹⁵.

A quienes hacen esto también se les denomina escrutadores de la red (*web scraping*), es decir, se trata de una técnica para extraer automáticamente datos de los sitios web usando un software apropiado para esta labor. Los hay, generalmente para que no se necesite realizar la codificación. Estas herramientas pueden leer y extraer datos de las páginas web con mayor rapidez que cualquier persona, por lo que se requiere tener cuidado para que no afecten el desempeño de los servidores o de la banda ancha. También cada sitio web debe respetar los derechos de autor y consignar los términos de uso (sobre cómo se pueden acceder y utilizar los datos), y respetar la privacidad, que es un asunto polémico en la actualidad.

En una publicación de advertencia a los usuarios de habla hispana, María Rubal Thomsen (2018) enseña cómo Facebook registra cada movimiento de cada usuario, “guarda fotos, comentarios, *likes*, clics y cualquier cosa que se escriba o se comparta en la red, incluso la actividad que se lleva a cabo *offline*”. Gracias a las *cookies*, señala, se graban “información variada sobre la lista de amistades, recopila publicaciones en el muro y el historial de *posts* y anuncios sobre los que se hace clic”. Después, Facebook convierte la información en dinero. Aunque los servicios son gratis para los usuarios, la red explota económicamente la información que recopila y luego la vende a anunciantes o para investigaciones académicas. Con los datos que se almacenan automáticamente, se segmenta la información de acuerdo con: a) localización, b) género, c) nivel de formación, d) edad, e) ingresos, f) ideales políticos, g) intereses y muchas otras variables de acuerdo con cada perfil (Rubal, 2018).

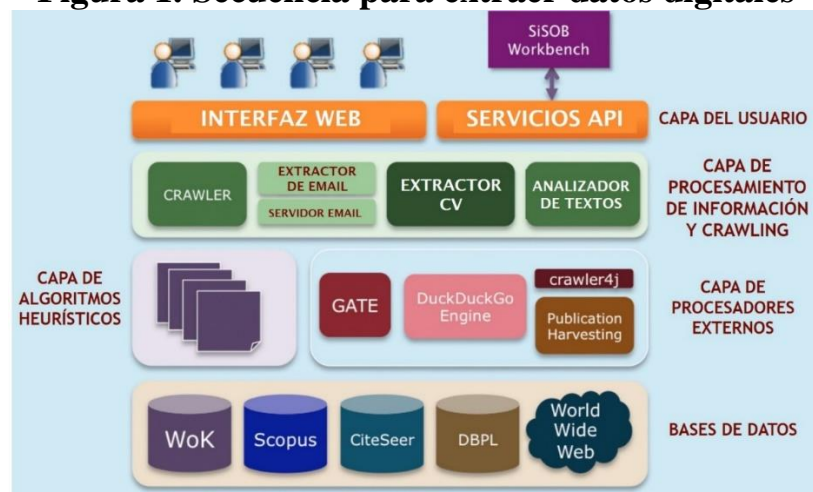
La práctica comercial de escrutar en la web ha traído consigo problemas legales por el uso de datos para competir comercialmente e incluso para

¹⁵ Por poner un ejemplo, PromptCloud se anuncia como un extractor de datos de múltiples fuentes, sin importar que contengan datos estructurados o sin estructurar. Tomado de <http://bit.ly/2oIFq9Y>, 28 de febrero de 2018.

realizar inversiones financieras. Se convierte en un asunto sensible si, por ejemplo, se realizan correlaciones de los tuits y las defunciones por infartos para las compañías de seguros médicos y de vida. Hay servicios para checar los cambios de precios de los productos, monitorear la liquidez, extraer información automáticamente sobre precios a gran escala, usar datos en tiempo real para la toma de decisiones en los negocios; controlar ventas o descuentos a productos no autorizadas, dar seguimiento a sitios web que venden productos piratas, o hacen envíos a países no autorizados (*Data Extraction Service*). También se ha convertido en una herramienta invaluable para la estadística, la búsqueda de información y el análisis científico.

Por lo pronto, se puede decir que las preocupaciones legales sobre su utilización dependen de cada contexto específico de hechos (Snell & Menaldo, 2018). Retomamos aquí un esquema construido *ad hoc* para extraer información o datos digitales (véase figura 1), donde se aprecian las fases que se siguen para tal fin (Geuna, et al., 2015). Encontramos desde las fuentes primarias de los datos (la *www*, Scopus, Wok, CiteSeer y PubMed), luego sigue la etapa heurística donde los investigadores o personas que construyen la base de datos estudian, hacen mayores búsquedas¹⁶ y adecuan la información; en la tercera etapa o capa se extraen datos de fuentes en línea específicas acordes a la naturaleza de la investigación que se trate y también se someten los datos a un análisis de textos; y la última etapa es cuando se pone a disposición en línea o para los usuarios la información procesada.

Figura 1. Secuencia para extraer datos digitales



Fuente: tomado de Geuna, Aldo, et al. (2015, p.1647)

¹⁶ El “crawler” es un motor que hace búsquedas sistematizadas en la ‘www’ para crear índices de datos.

Una guía para saber qué tipo de información podemos utilizar en las investigaciones es la tipología nominal de los usos de internet, que aunque aún está en construcción, existen adelantos para ubicar el terreno donde se pueden utilizar las herramientas digitales para la investigación: a) entretenimiento, b) comercio, c) búsqueda de información, d) socialización, e) uso del correo electrónico, f) contar con un blog o *weblog*, g) usos similares a los medios de comunicación tradicionales: noticias, deportes, eventos, etc., h) educación, i) trabajo, y, j) adicciones o vicios (pornografía, apuestas).

Pero también se pueden tipificar a los usuarios de acuerdo con sus habilidades digitales y las plataformas de preferencia (Facebook Snapchat, Twitter). O bien, de acuerdo con tres dimensiones:

1. cantidad o número de horas de dedicación en términos generales a Internet.
2. variedad de aplicaciones específicas como correo electrónico, videos, y
3. tipos de uso (Hampton, 2017, p. 176).

Las huellas en Internet. Insumos para estudiar

Las huellas en Internet son rastros que deja el uso de medios digitales y a partir de los cuales se generan datos de manera automática. La operación que permite detectar y almacenar rastros y datos son los algoritmos. De tal suerte que cualquier actividad en la red, el correo electrónico, el blog, la búsqueda en cualquier plataforma como Facebook, YouTube, Netflix, Uber, mensajería instantánea, llamadas telefónicas tradicionales, uso de GPS, mapas, etc.; registra aspectos de la personalidad del usuario: preferencias, intereses, interacciones, cualidades. Cada clic o búsqueda es un rastro.

La investigación social digital en la actualidad tiene un soporte metodológico clave en los mecanismos o sistemas de información donde se forman, se colectan y almacenan datos para ser interpretados y aplicados. Los instrumentos para realizar esas tareas pueden catalogarse como comerciales, o sea de proveedores que venden la información, o a través de plataformas disponibles abiertamente, como la de Google Analytics. Los primeros se pueden encontrar según la naturaleza de la información que se requiera. Por ejemplo, si queremos datos de audiencias de música o televisión, está Nielsen, que checa los tamaños de la audiencia, la exposición de géneros musicales, los productores, los intérpretes, etc. Históricamente se basaban en el fetiche de los números, que hoy se ha innovado hacia los *big data*.

Es fundamental entender que en Internet, como dice Nancy Baym (2013), el involucramiento deja huellas detectables. Los usuarios se conectan a través del tiempo y el espacio, sostienen grupos de discusión, crean repositorios de información, construyen redes de distribución entre ellos; y en esa medida se vuelven identificables para la industria. Así, con las tecnologías más avanzadas, la medición va más allá de número y porcentajes, se profundiza en las emociones a través de técnicas de reconocimiento fácil, monitoreo de latidos del corazón, análisis de las reacciones de la piel, encuestas, medición de las ondas cerebrales (Napoli, 2011), que eventualmente tienen efectos jurídicos para la administración de la justicia o la investigación jurídica. Es decir, se trata de darle sentido a los datos (Boyd & Crawford, 2012), para lo que se necesita capacitación, recursos humanos y software. Pero también se han creado sistemas de información personales que se combinan con los sistemas comerciales. Y al convertirse en grandes bases de datos pueden ser un punto de partida para aplicar encuestas o complementarlas.

Pero veamos ¿cómo se “siembran” las huellas digitales? O de qué fuentes se pueden obtener. Las más conocidas son los que están adherido o empotrados en las computadoras o son interfaces de programación (Hampton, 2017, p. 180);

1. El buscador o *scraping*, que es una especie de guion adherido en las computadoras para recolectar secciones de una página web y almacenar la información en bases de datos (Wesler, Smith, & Gleave, 2008).
2. Las interfaces de programación o herramienta de programación son proporcionados por una compañía de medios o de software que permite interactuar con el sistema de la computadora (*Application Programming Interface*, o API).

Como hemos dicho, lo más sensible es que a través de las huellas o trazos de datos y por el uso frecuente de algunas plataformas, se pueden construir perfiles de las personas, determinar preferencias políticas, la personalidad, estados de ánimo inferidos por el discurso, sentimientos, conductas, etcétera (Golbeck, Robles, & Turner, 2011).

En la industria de la música por ejemplo hay compañías de medición y análisis de información como Musicmetric, que agrega, almacena y analiza vastos montos de actividad sobre artistas y bandas en línea. Hay, asimismo, sistemas de información creados por personas individualmente conocidos como DIY (*do it yourself*) que se construyen para organizar datos a la medida, guardarlos y analizarlos.

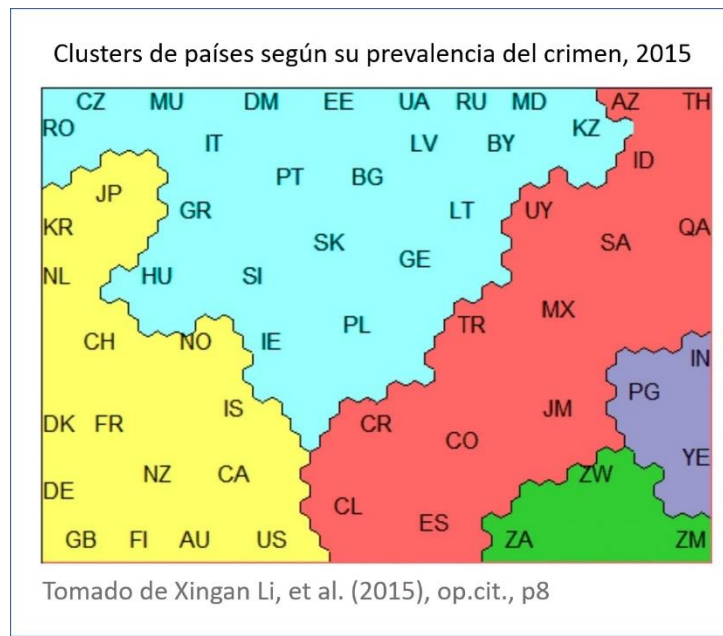
Las plataformas proporcionan herramientas para realizar mediciones que no requieren especialización y que comúnmente son visibles en las interfaces de sus sitios y se pueden activar para dar seguimiento, aprobar (con un *like*), reenviar, etc., cualquier información. No es una concesión graciosa de las grandes corporaciones de medios, sino que son sus estrategias para contabilizar, enganchar a los usuarios y hacer crecer la audiencia o seguidores.

Un ejemplo del avance en herramientas digitales para la investigación es el mapeo sobre el crimen en más de 50 países. Lo que se inició a principios de los ochenta, utilizando un software para el análisis de datos en estudios cualitativos, hoy ha avanzado hacia innovaciones de herramientas de investigación en línea a escalas continentales.

Por su importancia metodológica nos referiremos aquí a un estudio realizado por una universidad finlandesa en colaboración con investigadores de otras instituciones. Xingan Li, de la Universidad de Tallinn, Estonia, junto con Henry Joutsijoki, Jorma Laurikkala, Martin Juhola (Li, Joutsijoki, Laurikkala, & Juhola, 2015), de la Universidad de Tampere, hicieron un mapeo de gran alcance espacial sobre la situación del crimen en cincuenta y tres países utilizando las grandes bases de datos generadas en cada uno de éstos entre 2004 y 2008. Los datos fueron procesados por el programa SOM (*Self-Organizing Map*) asistido por un conjunto de métodos y estadísticas que posibilitaron mapear y comparar la incidencia de los crímenes en esos países. Lo relevante para los estudios sociojurídicos es la jerarquización de los países por la incidencia del crimen, un fenómeno crítico internacionalmente. Pero los datos organizados a gran escala son la parte metodológica crucial. La información que organizaron provino de datos demográficos sobre estructura de la población (desempleo, población urbana), calidad de vida (alfabetismo, gastos en salud, mortalidad infantil, esperanza de vida, tasas de crecimiento poblacional, tasas de mortalidad) y dinámica poblacional (tasa de nacimientos, de muerte, fertilidad, migración neta, densidad poblacional). Las fuentes de información fueron las bases de datos del Programa para el Desarrollo de Naciones Unidas y las Estadísticas de Finlandia.¹⁷ Ello es ejemplo de los avances obtenidos con el manejo de datos para conocer fenómenos sociales muy sensibles a escala global.

¹⁷ Véase Li, Xingan et al., *op. cit.*, p. 10. En los cinco clusters clasificados (véase figura 2), el de color más fuerte (rojo) corresponde a valores más altos, mientras que los tonos bajos (amarillo y azul) corresponden a valores bajos. Ahí podemos apreciar que México se localiza en el color rojo fuerte, que es el cluster de trece países donde están Azerbaijan (AZ), Tailandia (TH), Indonesia (ID), Uruguay (UY), Arabia Saudita (SA), Qatar (QA), México (MX), Turquía (TR), Jamaica

Figura 2. Clusters de países según prevalencia del crimen, 2015



La evidencia electrónica

La evidencia electrónica no es sólo la generada por computadoras, sino la creciente variedad de materiales como textos, imágenes y datos ubicados en teléfonos móviles, así como las imágenes y sonidos grabados por las cámaras corporales donde quedan sembradas huellas o rastros de conductas y actitudes humanas. En los tribunales judiciales, la evidencia electrónica es ya crucial para el desenlace de los juicios. Significa un gran reto para jueces, defensores y fiscales que deben capacitarse o auxiliarse de técnicos que resuelvan los problemas que involucran.

Hasta el momento, han surgido ideas para que los documentos públicos en México cuenten con un código susceptible de ser escaneado para generar la certidumbre deseada y prescindir del testigo, el perito o el acusado para que los reconozcan o informen sobre ellos, y su desahogo sea exclusivamente por medio de esa lectura. Así se evitarían dilaciones por falta de comparecencia de servidores públicos o testigos para reconocer o informar acerca de los documentos y, por supuesto, se evitaría la desatención de las funciones de los órganos del Estado mexicano (Moreno, 2018).

Tratar estos temas requiere una comprensión tanto de doctrinas legales que se desarrollaron en un contexto y de una variedad de tecnologías de

(JM), Costa Rica (CR), Colombia (CO), Chile (CL) y España (ES); mientras que los colores amarillo y azul se encuentran Canadá, Estados Unidos junto con Japón, Nueva Zelanda, Gran Bretaña, Francia, Nueva Zelanda, del clúster tres

reconocimiento facial, grabación de sonido e imágenes, fotografías, sensores, drones y, las más recientes, que combinan todas éstas y les dan sentido o interpretación a los datos a través de algoritmos. En general, Mason y Seng (2017), junto con los otros colaboradores del libro *Electronic Evidence*, hacen un excelente trabajo explicando los tecnicismos y poniéndolos en perspectiva.

En una investigación financiera sobre el corporativo inmobiliario Homex, radicado en Culiacán, Sinaloa, llevada a cabo a partir de 2012 por el órgano regulador del mercado bursátil en Estados Unidos, la Comisión de Bolsa y Valores, se descubrió con tecnología satelital que la constructora cometió fraude o engaño, ya que las fotos desde satélites arrojaron que las viviendas reportadas, no existían, y que en realidad eran terrenos planos sin casas edificadas (Sicairos, 2018). Esa fue la evidencia para castigar a la empresa, auxiliada por los satélites.

Asimismo, equipos multidisciplinarios, en los que participan ingenieros, arquitectos, cineastas, codificadores, arqueólogos, diseñadores y psicólogos, que son especie de detectives sociales que colaboran con organizaciones pro derechos humanos como Amnistía Internacional y Human Rights Watch, realizan indagaciones forenses altamente especializadas. Se trata de Forensic Architecture, de la Universidad de Londres, que ayuda a investigar asesinatos mediante el examen de videos, las publicaciones en redes sociales, software de representación arquitectónica, fotogrametría, imágenes, bandas sonoras, etc. A través de softwares elaborados especialmente, pueden determinar o encontrar evidencias de mentiras, crímenes, violaciones de derechos humanos, etcétera.

Las huellas digitales como evidencia judicial

Las huellas digitales más conocidas son las encontradas en dispositivos móviles digitales como los teléfonos inteligentes por ser la tecnología de comunicaciones mayormente masificada, y devienen en pruebas por excelencia para asuntos judiciales. Sin embargo, su reconocimiento oficial ha estado plagado de complejidades. Los casos se repiten sobre todo por el uso de dispositivos inalámbricos inteligentes que graban, toman fotografías, videos que se transmiten en tiempo real.

Un caso histórico es el del FBI contra Apple. En 2017, la Agencia de Inteligencia de Estados Unidos le solicitó a la empresa una “llave maestra” para acceder a la información del iPhone de uno de los atacantes del tiroteo de San Bernardino del 2 de diciembre de 2015, por el cual se armó un litigio casi interminable, al cual la compañía se negó sistemáticamente a abrir.

Fuera de aspectos criminales, existen actividades que tienen que ver con prestación de servicios cuyo insumo es el conocimiento y su aplicación en línea y que dan lugar a contratos civiles, operaciones mercantiles y financieras que demandan la intervención de abogados y tribunales judiciales. Ahí encontramos un campo de conocimiento que pronto rebasará al alcanzado por la ciencia previa al advenimiento de Internet.

En muchos países, las pruebas que aporta el celular son clave para reconstruir hechos. En Argentina, la evidencia contenida en los celulares tiene validez cuando la ordena un juez. “Se hace por pedido judicial. Si el Juez no lo ordena y no envía el teléfono no se puede hacer”. El aparato se manda con un oficio a una Dirección de Inteligencia para que procedan a extraer la información (Zagalsky, 2018). Al menos desde 2009 los tribunales se ponían a prueba por el uso de evidencias electrónicas. Pero en 2011, causó revuelo que un mensaje de correos con contenido erótico no se aceptara como prueba de “que hubiera existido una relación ni mucho menos infidelidad”. De acuerdo con la Sala M de la Cámara Nacional de Apelaciones en lo Civil en Argentina, “La infidelidad virtual en tanto no pase a 3D, no llega a consumir el encuentro carnal que configuraría el adulterio” (Himitian, 2014). No obstante que se informó que más del 70 por ciento de los casos de divorcio que llegan a tribunales se inician con pruebas informáticas como mensajes de texto.¹⁸,

En Colombia, la validez probatoria de una evidencia digital debe cumplir con requerimientos documentales para su admisión en un procedimiento legal. Por la importancia, cada día hay más leyes y normativas que regulan actividades relacionadas con la informática y el uso (o mal uso) que se haga de ella. Según la Corporación de Seguridad Legal y Forense de ese país, “el 80 por ciento de las evidencias de cada día, pruebas, notificaciones en los procesos administrativos y judiciales son correos electrónicos, grabaciones digitales, mensajes de texto, páginas de Internet, videos multimedia, sim card, entre otros muchos que se encuentran en formatos digitales” (ADALID,

¹⁸ También, en la ciudad de Buenos Aires, se condenó en febrero de 2018 “a una mujer a 150 horas de trabajos sociales por haber calumniado e injuriado a un hombre en la red social Twitter” siendo la primera sentencia de este tipo que tuvo lugar. “Se acreditó que la mujer, inventó una cuenta falsa desde la que acusó a un hombre de haber cometido varios delitos, lo que motivó su denuncia meses después de que estos tuits fueran ampliamente difundidos”. Redacción La Vanguardia, “Primera condena en Argentina a una mujer por insultar en Twitter”, *La Vanguardia*, 17 de febrero 2018, <https://bit.ly/2xNkjaU>

2016), por lo que se ha legislado la admisión de la evidencia digital ante una instancia judicial.¹⁹.

En suma, la petición, aporte, decreto, práctica y valoración de la evidencia digital en el marco de un proceso judicial o administrativo es especialmente importante para confirmar la existencia de la infracción y responsabilidad de los infractores, como también lo son para la investigación jurídica especializada. Sin duda, con la tecnificación de la delincuencia, las tecnologías digitales son herramientas necesarias en el auxilio de la administración de justicia, el esclarecimiento del delito y la persecución del delincuente.

A propósito del uso de Periscope como evidencia de violación de normas jurídicas de buen gobierno, Irene Levy introdujo una discusión con elementos muy completos sobre el uso de una tecnología digital para inculpar personas. En la ciudad de México, a finales de 2015 un funcionario transmitió en vivo videos a través de Periscope de ciudadanos que cometían infracciones de tránsito o de normas urbanas. Al ser una transmisión del momento, en vivo, la autora cuestiona desde el punto de vista moral el hecho de que “la sola aparición de una imagen, un video o un audio en las redes sociales que denuncia a una persona que actuó “mal”, nos haga presumir su culpabilidad” (Levy, 2017).

Ante ello, lo que se pone en juego son la presentación de fenómenos jurídicos renovados por el uso de las tecnologías. La misma observadora se pregunta, poniendo en juego la hipótesis de colusión de diversos derechos, si

¿Es legal que la autoridad transmita en vivo videos de ciudadanos infractores?
¿Qué fin se persigue con ello: ¿sancionar, obtener pruebas, disuadir a otros de cometer conductas negativas, hacer propaganda política, transparentar su actuación? ¿Es lo mismo jurídicamente que un particular divulgue datos personales de terceros a que lo haga la autoridad? ¿Es justiciable la vulneración de la privacidad frente al valor de la transparencia —suponiendo que este fuera— en el actuar de los funcionarios o de la exhibición pública del mal ciudadano? ¿Están las autoridades, en dicho papel, ejerciendo su derecho a la libertad de expresión? ¿Se plantea una colisión de derechos válida? (Levy, 2017, pp. 8-9).

De la misma manera, así como Anonymous y WikiLeaks llevaron a pensar que Internet es un espacio de libertad casi absoluta, también, la realidad

¹⁹ E instan a “los profesionales en Derecho, así como los entes de investigación mantenerse a la vanguardia del conocimiento, de la aplicación de la legislación y de las técnicas para el correcto manejo de la evidencia digital y la admisibilidad legal de la evidencia digital en situaciones judiciales” (ADALID, 2016).

virtual, plantea serias polémicas legales de garantías y derechos con nuevas connotaciones por reivindicar. Los millones de personas a cuya costa se generan los datos masivos que a su vez dan poder económico y financiero a las corporaciones de negocios virtuales pueden tarde o temprano dar lugar a acciones legales de mayor alcance que exija a los corporativos una recompensa por los datos utilizados, la privacidad vulnerada, los contenidos generados a través de Internet, las redes, los blogs y las múltiples plataformas que son alimentadas por los usuarios.

Internet de las cosas y las leyes como recurso secundario

La ubicuidad y movilidad de las tecnologías de información y comunicación o Internet se ha impulsado por el afán de comercialización desmedido. La generación de datos es de tal magnitud que podemos decir que “El mundo físico en sí mismo se está convirtiendo en un tipo de sistema de información”, de ahí la denominación genérica de Internet de las cosas. A cada objeto de cualquier dimensión o material, desde carreteras hasta marcapasos cardiacos, se le empotran sensores y se les conecta a redes alámbricas e inalámbricas para monitorear movimientos de órganos humanos o de flujo de automóviles, a menudo utilizando el mismo Protocolo de Internet (IP) que conecta Internet. Estas redes generan enormes volúmenes de datos que almacenados en computadoras se transforman en las grandes bases de datos para su tratamiento.

Cuando los objetos pueden percibir el entorno y comunicarse, dice McConville y Chui (2007), se convierten en herramientas para comprender la complejidad y responder a ella rápidamente. Lo revolucionario de todo esto es que los sistemas que activan la información física se empezaron a implementar incluso sin intervención humana aparente.

La Internet de las cosas surge por la incrustación de sensores en los objetos de manera estratégica:

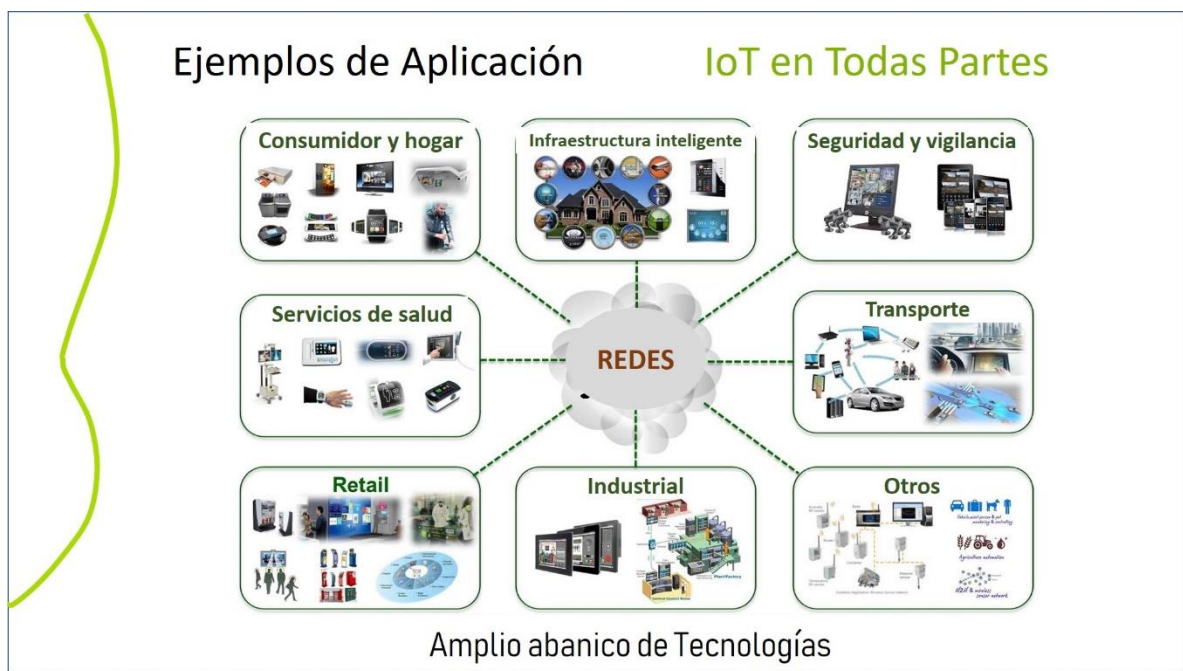
- a) en los objetos o artefactos portados por las personas,
- b) en los medios de transporte convencionales,
- c) en cada espacio que pueda ser enlazado a una red de comunicación.

Los objetos o las cosas dan pie a la generación de datos a escala masiva que se almacenan y posibilitan la conversión a información valiosa para distintos efectos. Es decir, no es necesario la acción intencionada de las personas como, por ejemplo, realizar movimientos bancarios, pagos de servicios en línea, poner mensajes en Facebook, enviar mensajería móvil, compras en tiendas comerciales o en línea, transferencias bancarias móviles, etc., sino

que de manera involuntaria con el simple accionar de un interruptor de aparatos domésticos se van registrando secuencias de huellas digitales.

La generación de datos automáticos por excelencia es la que tiene lugar con la manipulación de artefactos computarizados y conectados a redes digitales como las computadoras, laptops, teléfonos inteligentes, televisores, impresoras, donde las personas van dejando huellas de cada búsqueda, de cada imagen que ven, de cada música que escuchan, de cada clic que dan a imágenes que envían los proveedores o navegadores de Internet y que dan lugar grandes bases de datos ordenadas mediante los algoritmos sobre las preferencias, sentimientos, deseos, intenciones, proclividades, como hemos dicho.

Figura 3. ¿Cómo se genera Internet de las cosas?



Fuente: Tomada de Pelechano, Vicente (2014) el 1 de septiembre de 2018.

En una alusión inusitada por el alud comercial de la Internet de las cosas y el negocio global en que se ha convertido para las grandes plataformas, el jefe de la oficina de tecnología de IBM (International Business Machines Corporation), Bruce Scheneier, preocupado por la caída de Internet en 2017, encontró que una solución a esas vicisitudes cada vez más riesgosas está en la ley, es decir, "revertir la tendencia a conectar todo a Internet", que parece, por otra parte, ir en contra de la vorágine de su crecimiento (Schneier, 2018).

Sin embargo, esa preocupación parece imposible de realizarse. Veamos hacia dónde ha avanzado la Internet de las cosas o Internet de todo: está incrustada en nuestra cotidianidad, pues cada cosa es una computadora (Schneier, 2018). Es decir, ya no tenemos cosas con cómputo incrustado a ellas. Tenemos computadoras con cosas incrustadas a ellas.

- El refrigerador moderno es una computadora que mantiene las cosas frías.
- El horno de cocinar es una computadora que hace las cosas calientes.
- El cajero bancario automático es una computadora con dinero adentro.
- El automóvil “ya no es un dispositivo mecánico con algunas computadoras adentro; es una computadora con cuatro ruedas y un motor. En realidad, es un sistema distribuido de más de 100 computadoras con cuatro ruedas y un motor”.
- Los teléfonos son computadoras de uso general desde 2007, cuando se introdujo el iPhone.

Pero también, mediante las computadoras empotradas:

- Se rastrea la actividad física con los dispositivos habilitados para tal fin.
- El rastreador por excelencia es el teléfono inteligente con localizador porque se carga a todos lados siempre.
- Las casas tienen termostatos inteligentes, electrodomésticos inteligentes, cerraduras de puertas inteligentes e incluso encendido de luces inteligentes.
- En las oficinas los dispositivos inteligentes se conectan en red con cámaras de CCTV, y sensores que detectan los movimientos de los clientes y todo lo demás.
- Las ciudades se arman de sensores inteligentes en las calles, el alumbrado público, el riego de los jardines.
- Las carreteras tienen sensores para vigilar, medir el flujo de transporte, reportar accidentes a policías, cuerpos de rescate y hospitales. Son vigiladas en cualquier tramo por satélites.

Después de exponer algunos temas de relevancia para entender la nueva metodología de investigación digital, explicaremos dos técnicas de investigación jurídica y sus aspectos más importantes.

La entrevista a abogados u operadores del derecho

El objetivo primordial de la entrevista como técnica de investigación social es obtener información, opiniones, percepciones de una persona de manera personal o directa. La interacción cara a cara es nodal. En ella el contacto personal permite al entrevistador percibir reacciones, gestos, animosidades del sujeto, y hay oportunidad de profundizar en asuntos de la entrevista. Puede hacerse una caracterización basada en la cercanía, la cual difícilmente permite otro tipo de técnica. Incluso si se trata de entrevista estructurada o semiestructurada.

Ventajas, barreras y desventajas

La entrevista como técnica de investigación jurídica es el instrumento por excelencia para obtener información sensible y muy escasa, incluso en los medios de comunicación tradicionales. Aunque se afirma que puede ser utilizada para cualquier segmento de la población, por aquello de las inequidades alfabéticas, se tiene que adaptar a cada segmento poblacional, en nuestro caso a los operadores del derecho que tiene características peculiares a tomarse en cuenta para planear la entrevista.

Incluso por la naturaleza de la academia de juristas mexicanos, con tendencias a la secrecía en asuntos legales públicos o privados, la entrevista adquiere connotación de obligatoriedad bajo el enfoque de la sociología jurídica para tratar de romper la monotonía de los textos acartonados y repetitivos, y despertar el interés por el conocimiento más allá de leyes y códigos.

Entonces las ventajas de la entrevista es que a través de ella:

- a) Se puede obtener información confiable y suficiente para los intereses de la investigación.
- b) Se puede obtener información original.
- c) Se pueden estudiar fenómenos que comúnmente no están abiertos a la observación, como son las actitudes, sentimientos, opiniones y reacciones. Como dicen quienes realizan entrevistas para seleccionar empleados: a través de la entrevista se puede conocer el pasado, el presente y el futuro de una conducta.

Barreras

El entrevistador tiene que disponer de tiempo suficiente para realizar entrevistas tanto a personas que si acceden a hablar como a las que no tienen tanta disposición. Pues estas últimas complican la entrevista o mejor dicho ponen a prueba la capacidad de quien la realiza de obtener información profunda. Para llegar al meollo o el interés del asunto, a veces resulta difícil e incluso infértil, y los asuntos duros a veces no llegan a conocerse. También el entrevistador debe esforzarse por hacer a un lado ideas preconcebidas o prejuicios, y estar abierto a recibir diversidad de opiniones y tratos.

Resulta costoso si se trata de entrevistas a personajes importantes o encumbrados por las necesidades de traslados constantes y que pueden ser infructuosos. Ante estos la comunicación digital es irrelevante.

También puede ser subjetivo. A veces el hecho de no poder tomar notas o grabar la entrevista, predisponen al entrevistado y el entrevistador tiene que hacer uso de su memoria en el momento de tomar notas posteriores. El entrevistador necesita saber manejar las emociones, sentimientos, preferencias y prejuicios del entrevistado, sobre todo en asuntos que tocan lo político.

Sugerencias para realizar entrevistas apropiadas a operadores del derecho²⁰:

- a) Crear una atmósfera amigable, ser cortés. Presentar las credenciales personales y de la investigación motivo de la entrevista.
- b) Asegurar la confidencialidad de la información.
- c) Hacer una introducción breve, casual y positiva.
- d) Procurar no desviar la plática del asunto principal.
- e) Entrar al tema lo más pronto posible, y procurar no despertar sospechas o suspicacias del entrevistado
- f) No ser ni muy efusivo ni muy severo
- g) Debe observar un interés por las opiniones del respondiente y nunca divulgar las propias
- h) Si se le pregunta su opinión, debe desviarla y retomar la pregunta de su interés.
- i) Debe mantener la dirección de la entrevista, descartar conversaciones irrelevantes y conservar al respondiente en el asunto.

²⁰ Véase capítulo 8 de Tracy, Sara J. *Qualitative research methods. Collecting evidence, crafting analysis, communicating impact*, Oxford, Wiley-Blackwell, 2013.

- j) Saber sortear evasivas o respuestas sin sentido de parte de los operadores del derecho.

El entrevistador requiere informarse previamente sobre cada entrevistado para conducir lo más apropiadamente posible la entrevista. Se debe tener presente que algunos de los operadores del derecho son personas que cuidan mucho la confidencialidad por ser funcionarios públicos o porque se desenvuelven en ámbitos similares y están convencidos de que deben tener una actitud cauta por ocuparse de muchos asuntos privados, o ser ellos mismos personas públicas que tienen que cuidar su prestigio o presencia social.

La encuesta en línea

Las encuestas en línea son ya una herramienta estandarizada para la investigación empírica social, como lo son para el mercadeo, para calificar la obtención de un servicio vía telefónica o automáticamente, o bien para obtener estadísticas oficiales.

En los países que lograron alta densidad telefónica, las encuestas telefónicas se remontan a los años sesenta, aunque en la actualidad en México aún se practican sobre todo para obtener opiniones sobre preferencias electorales; y las encuestas online para asuntos coyunturales no las han superado. Le siguieron en los ochenta las encuestas asistidas por cómputo que han evolucionado tremendamente con la comercialización o provisión de servicios de encuestamiento en línea.

No debemos olvidar que hechas vía Internet con sus variantes son la suma de una serie de avances tecnológicos en los procesos de transmisión en red (hasta llegar a la banda ancha), evolución de la web y la estandarización de los buscadores web, el desarrollo de los correos electrónicos, la interactividad y flexibilidad que brindan los lenguajes como Java, Javascript. Las encuestas en línea se empezaron a utilizar desde finales de los ochenta en Estados Unidos (Sproull & Kiesler, 1986). Su diseño ha experimentado un gran desarrollo, hasta llegar a ser comercializado como un servicio en línea o servicios de administradoras que se especializan en prestar esos servicios.

Hoy en día existe una floreciente industria para proveer servicios de encuestas. De acuerdo con el Consejo Americano de Encuestas de Investigación (*Council of American Survey Research Organizations*), el 76 por ciento de las organizaciones que las realizan ya participaban en el Data Collecton Trends Survey en 2006, y para el 32 por ciento de ellas ese tipo de

encuestas por Internet ya representaba el primer lugar de sus ingresos (Lozar Manfreda , Bosnjak, Berzelak, Haas, & Vehovar, 2008).

A la par de la técnica de la encuesta, tenemos que tomar en cuenta para la investigación digital, el conjunto de avances tecnológicos que han revolucionado no solamente esta herramienta de investigación académica, sino a industrias enteras de cualquier ramo: las huellas en Internet, los data mining (el proceso de generación de los datos), los *big data* (el almacenamiento o bases de datos gigantescas).

Consisten en muestras no probabilísticas que pueden hacerse a conveniencia y tener sesgos en la selección, aunque no apoyan la inferencia estadística. Si bien hay formas de reducir los sesgos gracias a la aplicación de muestras en red y utilizando métodos para ajustarlas (Hampton, 2017, p. 186).

En 2018, se ven patentemente los sesgos que llegan a tener los datos generados en las redes sociales o aquéllos creados por los errores de las plataformas.

Fue muy criticado en Estados Unidos, que a raíz del asesinato de 17 estudiantes de la escuela secundaria Stoneman Douglas de Parkland en Florida el 14 de febrero de 2018, los administradores de la plataforma de YouTube, cometieron graves errores en moderar los contenidos que fluían en esa red. Como dijo Louise Matsakis (2018),

queda claro que no tienen una idea real de cuál es su línea cuando se trata de moderar el contenido, pues empezaron a desactivar cuentas que difunden engaños, lo que fue visto con buenos ojos, sin embargo, los Lineamientos de la comunidad de YouTube no prohíben la teoría de la conspiración, o los videos teñidos directamente, haciendo aparecer a los moderadores como deshonestos por eliminar videos que mencionaban armas pero que no parecían violar ninguna regla (párr.).

El problema no es nuevo, precisa la analista, “durante años, los usuarios de YouTube se han quejado de políticas de moderación inconsistentes. A veces, la compañía vigila videos principalmente en respuesta a las protestas públicas, lo que hace que sus decisiones sean inherentemente caóticas y potencialmente maleables”, escribe Matsakis. “Es una tarea descomunal tratar de moderar un sitio en el que se cargan cientos de horas de video por minuto, pero es aún más difícil si no tienes reglas claras y consistentes sobre lo que está permitido” (Matsakis, 2018, párr.).

Cómo operan los encuestadores en línea

Son proveedores de servicios de encuestamiento. Ellos se encargan de distribuir cuestionarios en línea a personas donde quiera que se encuentren distribuidos geográficamente, de manera rápida y relativamente a bajo costo. Tienen la particularidad de identificar conductas susceptibles de contabilizarse y ayudan a reclutar participantes válidos. Los más conocidos en la actualidad son

Aunque existen bastantes servicios para participantes no comerciales que utilizan las industrias y organizaciones de profesionistas, los que más se frecuentan para reclutar participantes válidos son: SurveyMonkey Audience (SMA), Qualtrics y Amazon's Mechanical Turk. La importancia de estos es su flexibilidad y la posibilidad de acceso a la población requerida (Duane , Long, Loraas, Mueller-Phillips, & Vansant, 2014).

El Amazon Mechanic Turk permite conducir de principio a fin experimentos en línea acerca de “conductas”, cuya “experiencia es muy similar a la que se lleva a cabo en un ambiente de computo simulado”. Al parecer, son experimentos muy confiables sobre los que se han realizado estudios para poner a prueba su validez. Se les denomina también experimentos en línea o laboratorios virtuales que se practican desde mediados de los noventa.

Sus ventajas son el bajo costo para reclutar participantes, la manipulación automática, reducen la intrusión del interesado y aumentan la velocidad de las interacciones (Hampton, 2017, p. 185). La infraestructura electrónica en línea que facilita la experimentación se ha hecho más accesible y a bajos costos. Tal es el caso de la AMT (Amazon Mechanical Turk) (Hampton, 2017). Ellos cuentan con procedimientos de validación de datos demográficos. Por ejemplo, utilizan el domicilio de IP para verificar el país de residencia de los sujetos AMT”, encontrando en sus casos una certeza de “97 por ciento de respuestas ciertas”. En otra prueba realizada se comparó “la consistencia de las variables demográficas de sujetos de dos diferentes estudios, obteniendo un acuerdo de entre el 81 y el 98 por ciento” (Rand, 2011, p. 172).

A través del procesamiento de cuestionarios con tecnología de cómputo y la aplicación vía Internet, se agiliza una labor que demanda mucho tiempo y esfuerzos de coordinación de equipos de ayudantes, mejorando enormemente el tradicional cuestionario de papel: las respuestas obtenidas son almacenadas inmediatamente en una base de datos de cómputo, se

procesan con un software (como el SPSS²¹), y se elaboran gráficas con las que se realizan análisis estadísticos.

Estas herramientas tienen las ventajas de

- a) reducir los costos que implica contratar encuestadores y sufragar gastos de traslados y manutención si son ayudantes de investigación o estudiantes;
- b) tienen un amplio alcance independientemente de los espacios geográficos donde se localicen los encuestados;
- c) se reducen los errores al transcribir o capturar los datos;
- d) los cuestionarios computarizados usan interfaces gráficas de la *World Wide Web*, diseños atractivos, sugerencias para saltar preguntas, y otras facilidades;
- e) hoy los respondientes pueden acceder y contestar los cuestionarios interactuando con el navegador de su preferencia y utilizar las herramientas que éstos proporcionan;
- f) Pueden dar sensación de privacidad y motivar a contestar objetivamente;

A la vez tienen ciertas desventajas que deben buscar minimizarse:

- a) La más importante para nuestro contexto es que la accesibilidad a plataformas digitales en línea se reduce a ciertos segmentos de la población que en primer lugar obedecen a la división digital primaria por carencia de acceso de Internet a los hogares²²;
- b) el analfabetismo digital,
- c) la brecha generacional, a la cual nos hemos referido cuando explicamos sobre la aplicación de nuestra encuesta a investigadores mexicanos del área de ciencias jurídicas.

²¹ SPSS (*Statistical Package for the Social Sciences*), es un programa estadístico informático muy común para las ciencias sociales y aplicadas, y las empresas de mercadeo. Brinda capacidad para trabajar con voluminosas bases de datos y una sencilla interfaz para realizar análisis de estos. Hoy este software se anuncia como una plataforma de IBM SPSS que ofrece análisis estadístico Avanzado, una vasta biblioteca de algoritmos para el aprendizaje automático, análisis de textos, integración con los *big data*, entre otras aplicaciones. Se anuncia como fácil de usar, flexible y con plataformas escalable, que mejora la eficiencia y minimiza los riesgos” (IBM Analytics, <https://ibm.co/2zcKWYS>, 27 de febrero de 2018).

²² La brecha digital en México para 2016 era según la ENDUTIH del 53 por ciento de los hogares sin conexión, aunque el 59.5 por ciento de la población se declaró usuaria de Internet (INEGI, 2017).

- d) evasión fácil del encuestado, ya que no hay compromiso del encuestado por responder, y se puede escurar en el anonimato;
- e) desconocimiento sobre el manejo de las fronteras “líquidas” entre lo privado y lo público en un ambiente en línea.

El procedimiento para encuestar online, demanda preparativos de investigación similares a las encuestas tradicionales, es decir, la elaboración del cuestionario o las preguntas que ayuden a comprobar las hipótesis de trabajo, tener en cuenta el nivel socioeconómico y educativo de las personas a encuestar, prever el tiempo que requiere contestar el cuestionario, atender las normas de ética, preservar el anonimato de los respondientes,

A través de los métodos de investigación en línea, se requiere asegurar que la fuente de la información llene los criterios de calidad. Se dice que aseguran la calidad, ya que se basan en datos documentados, es decir, corresponden a información cierta de personas que tienen registros o que han sido corroborados con sistemas propios de los navegadores o plataformas como Facebook, Twitter, Google, etc. Son los denominados metadatos o parámetros, o bien los datos producidos durante el proceso de recolección (Fielding, Lee, & Blank, 2008) como son el tiempo de inicio y finalización de cada cuestionario, el tiempo que toma responderlo, la ubicación del cuestionario cuando se suspende o abandona. Esto es información valiosa para elaborar subsecuentes encuestas sobre la preparación y monitoreo del proceso que lleva la aplicación, así como estudiar los efectos de las diferentes preguntas, las reacciones que pueden despertar o las actitudes ante ellas (Lozar Manfreda, et al., 2008). Hay metadatos básicos y avanzados (Berzelak, Lozar Manfreda, Slavec, & Vehovar, 2012).

Los metadatos básicos son los que registra cada servidor que hospeda un cuestionario en la web: el domicilio IP, el software de la plataforma, veces que se ve una página, ruta que sigue el cuestionario.

Los metadatos avanzados son recolectados por la computadora que usa cada entrevistado: tiempos de respuesta a nivel de pregunta, detección de cambio de respuesta, visibilidad de cada pregunta en la página, seguimiento de eventos, etc. (Berzelak, et al., 2012).

La muestra y la calidad

Dos aspectos metodológicos cruciales para el curso y resultados de la investigación académica tienen que ver con la muestra y la calidad de información que se recaba con la encuesta en línea.

La representatividad es algo sumamente cuestionado en los métodos en línea, ya que la selección de la población está sujeta a factores de accesibilidad a Internet, las brechas digitales por generación, por nivel de educación, y la propia de cada país (que en México no alcanza al 50 por ciento de la población), entre otros. En ciertas condiciones es complicado seguir las reglas del muestreo probabilístico y se tienen que buscar atenuantes y recurrir al muestreo no probabilístico, es decir a la selección por conveniencia. Se puede decir, que son aproximaciones al estudio de la realidad que resulta válida y que tiene una representación dependiendo de la idoneidad y alcance de los integrantes de una muestra, no obstante, habrá que tener cuidado de indicar que no pueden hacerse inferencias estadísticas del mismo rigor que una muestra probabilística. En ciencias sociales no podemos fetichizar la matemática, y el análisis tendrá profundidad y pertinencia en la medida en que las premisas, la coherencia y el aparato conceptual utilizado sean los adecuados.

Se recomienda utilizar invitaciones por correo electrónico que requieren acceder a listas de correos o integrarlas a un grupo ad hoc, por ejemplo, usar algunas que estén en listas de instituciones o de grupos que comparten ciertos intereses.

En la eslovena Universidad de Liubiana, sostienen que la encuesta que utiliza correos electrónicos puede constituir una muestra probabilística, cuando las listas de dichos correos participantes elegibles estén disponibles y sean invitados individualmente, o bien que se invite a visitar de manera directa a un determinado link donde fueron seleccionados al azar (Lozar Manfreda, et al., 2008). Esto último usualmente se implementa usando ventanas *pop-up* o elementos de *flash* que contienen una invitación general a la encuesta web y que solo ven los visitantes seleccionados.

Por otra parte, el correo electrónico considerado un mecanismo imperfecto para reclutar muestras de población en línea, es recomendable combinarlo con otros métodos como el teléfono tradicional para obtener una muestra adecuada de los participantes, aunque es aceptable, como señalan Lozar Manfreda, et al. (2008), una tasa de respuesta del 10 por ciento, tal como ocurre con las encuestas telefónicas. También, el investigador puede optar por enviar mensajes personalizados a grupos de encuestados.

¿Qué eventualidades, además de las propias de la encuesta, se deben contemplar en la tasa de no respuestas a un cuestionario en línea con invitación por correo electrónico?

Además de las motivaciones personales, errores en el diseño del cuestionario, analfabetismo en computo o carencia de habilidades para moverse en Internet, preocupaciones por la privacidad o anonimato, podemos mencionar las siguientes: La negativa tácita de los encuestados o desinterés en participar; Iniciar y no concluir las respuestas; Falta de interés por seguir un *link* para completar el cuestionario; que haya errores en las direcciones de correo electrónico; que se canalicen como *spam*; y, falta de tiempo

Esto se acentúa cuando se tiene una población significativamente diferente o con perfiles muy diversos. Por ello la tasa de respuestas en las encuestas electrónicas es generalmente baja y varía considerablemente cuando se trata de encuestas comerciales.

Lo anterior, lleva a plantear la combinación de métodos de investigación social, o sea los denominados métodos mixtos de investigación, utilizando métodos cuantitativos y cualitativos, teniendo entre estos últimos a la etnografía, los grupos de enfoque, la entrevista profunda. Aunque históricamente se les ha confrontado, cada vez es más común que se combinen, en cualquier secuencia, ya sea que primero se apliquen encuestas y luego se realizan entrevistas cara a cara o viceversa. Incluso se afirma que al combinarse los cuestionarios suministrados en la web con las entrevistas profundas se puede obtener alta validez y poder explicativo sobre los datos recolectados (Lozar Manfreda , et al., 2008).

A manera de colofón

Por último, diremos que hay dos sesgos al menos que se integran al momento de diseñar los algoritmos. Uno, por la métrica de los medios sociales digitales, que se genera precisamente por el diseño del algoritmo que da primacía a algunos mensajes y usuarios sobre otros a través de recomendaciones automáticas, de los alimentadores y de los editores. Por ejemplo, un tema en Facebook recibe mas vistas o comentarios por la jerarquización que se hace de él, de tal forma que se convierte en un patrón, que se exagera por el hecho de que entre más *post* o usuarios reciben comentarios, *likes* o seguidores, el algoritmo lo hace más visible a otros usuarios, mientras opaca a otros²³. El segundo sesgo proviene de que los medios sociales son diseñados para reflejar mayormente los afectos positivos, o sea las aprobaciones... “que es parte de la economía de los *likes*”.

²³ Baym, Nancy, *op. cit.*

Finalmente diremos que la metodología jurídica digital en México tiene un camino largo que recorrer: en principio se debe incorporar a la curricula de las facultades de derecho donde aún no se inicia; incorporar a los profesores especialistas a formar a los investigadores para que la adopten y que sea en los posgrados en derecho una exigencia inaplazable. En esa medida podremos revertir el déficit de la investigación jurídica del país y su concentración asfixiante en la Ciudad de México y en unos cuantos centros de investigación.

Bibliografía

- ADALID. (2016). *Evidencias digitales, manejo legal y técnico para abogados e ingenieros*. Obtenido de ADALID Security, Legal and Forensic Corporation: <http://bit.ly/2FIvh52>
- Baym, N. (2013). Data not seen: The uses and shortcomings of social metrics. *First Monday*, 28(10), 4. Obtenido de <https://bit.ly/2xQec5E>
- Berzelak, N., Lozar Manfreda, K., Slavec, A., & Vehovar, V. (2012). *Advanced paradata in web surveys: What can they tell about the response process?* Obtenido de American Association for Public Opinion Research: <https://bit.ly/2F6p85t>
- Boyd, D., & Crawford, K. (2012). Critical questions for big data. Provocation for a cultural, technological, and scholarly phenomenon. *Information Communication and Society*, 15(5), 662-679.
- Calvi, J., & Coleman, S. (2017). *American Law and Legal Systems*. Nueva York: Routledge.
- De la Fuente, J. (19 de marzo de 2018). Inteligencia Artificial II. *El Universal*. Obtenido de <https://bit.ly/2GePOgl>
- Duane , B., Long, J., Loraas, T., Mueller-Phillips, J., & Vansant, B. (2014). Online instrument delivery and participant recruitment services: Emerging opportunities for behavioral accounting research. *Behavioral Research in Accounting*, 26(1), 1-23.
- Evans, J., & Aceves, P. (2016). Machine translation: mining text for social theory. *Annual Review of Sociology*, 42, 21-50.
- Fielding, N., Lee, R., & Blank, G. (2008). *The SAGE handbook of online research methods*. Thousand Oaks (CA): Sage.
- Geuna, A., Kataishi, R., Toselli, M., Guzmán, E., Lawson, C., Fernandez-Zubieta, A., & Barros, B. (2015). SiSOB data extraction and

- codification: A tool to analyze scientific careers. *Research Policy*, 44(9), 1645-1658.
- Golbeck, J., Robles, C., & Turner, K. (2011). Predicting Personality with Social Media. *alt.chi*, 253-261.
- Hampton, K. (2017). Studying the digital: directions and challenges for digital methods. *Annual Review of Sociology*, 43, 167-188.
- Himitian, E. (2 de abril de 2014). Infidelidad virtual: ¿Es engaño el coqueteo en las redes sociales? *La Nación*. Obtenido de <https://bit.ly/2lbkEhE>
- Hine, C. (2015). *Ethnography for the internet. Embedded, embodied and everyday*. Londres: Bloomsbury Academic.
- Levy, I. (2017). El ejercicio de la función pública: Una perspectiva desde las nuevas tecnologías, la transparencia y los derechos humanos. *Cuadernos de Transparencia*(25), 8-63.
- Li, X., Joutsijoki, H., Laurikkala, J., & Juhola, M. (2015). Crime vs. demographic factors revisited: Application of data mining methods. *Weblog*, 12(1).
- Lozar Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I., & Vehovar, V. (2008). Web survey versus other survey modes: A meta analysis comparing response rates. *International Journal of Market Research*, 79-104.
- Madden, R. (2010). *Being ethnographic: A guide to the theory and practice of ethnography*. Londres: Sage.
- Markham, A., & Baym, N. (2016). *Internet inquiry: Conversations about method*. Thousand Oaks (CA): Sage Publications.
- Markham, S., & Stavrova, S. (2016). Internet digital research. En D. Silverman, *Qualitative Research* (pág. 229). Londres: Sage.
- Mason, S., & Seng, D. (2017). *Electronic evidence* (4a ed.). Londres: Institute of Advanced Legal Studies of the University of London.
- Matsakis, L. (2 de marzo de 2018). Youtube doesn't know where its own line is. *Wired*. Obtenido de <https://bit.ly/2FjBrLa>
- McConville, M., & WingHong, C. (2007). *Research Methods for Law*. Edinburgo: Edinburg University Press.
- Moreno, M. (29 de marzo de 2018). Afinemos el concepto de prueba documental. *El Mundo del Abogado*. Obtenido de <https://bit.ly/2R9mTRg>

- Napoli, P. (2011). *Audience evolution: New technologies and the transformation of media audiences*. Nueva York: Columbia University Press.
- Pelechano, V. (2014). *Internet de las cosas: Tecnologías, aplicaciones y retos de futuro*. Obtenido de SlideShare: <https://bit.ly/2RdXgPz>
- Rand, D. (2011). The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *Journal of Theoretical Biology*, 172-179.
- Rubal, M. (3 de marzo de 2018). Descubre el archivo que Facebook tiene sobre ti. La red guarda un registro con todas las fotos, comentarios y 'me gusta' de cada usuario. *La Vanguardia*. Obtenido de <http://bit.ly/2CXwxOz>
- Schneier, B. (2018). *Security Orchestration and Incident Response*. Obtenido de Schneier on Security: <https://ibm.co/2rP8s9Y>
- Sicairos, A. (5 de marzo de 2018). Homex y la conexión Peña: secuelas de la impunidad. *Nororeste*. Obtenido de <https://bit.ly/2HiKtVV>
- Snell, J., & Menaldo, N. (8 de junio de 2018). Web scraping in an era of big data 2.0. *Bloomberg Law News*. Obtenido de <http://bit.ly/2oF1n9Q>
- Sproull, L., & Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. *Management Science*, 32(11), 1492-1512.
- Wesler, M., Smith, D., & Gleave, E. (2008). Distilling digital traces: Computational social science approaches to studying the internet. En N. Fielding, R. Lee, & G. Blank, *The SAGE Handbook of Online Research Methods* (págs. 116-140). Londres: SAGE Publications.
- Wills, T. (2017). Social Media as a research method. *Communication Research and Practice*, 2(1), 7-19.
- Winick, E. (2017). Lawyer-bots are shaking up jobs. AI is augmenting and automating the tasks currently performed by hundreds of people in the U.S. alone. *MIT Technological Review*. Obtenido de <https://bit.ly/2AOdhUI>
- Zagalsky, A. (27 de enero de 2018). Teléfonos celulares: los inesperados espías de los criminales modernos. *Tecno*. Obtenido de <http://bit.ly/2oFvaiy>