Cuadernos de
CAPACITACIÓN
ELECTORAL

USO DEL BIG DATA
Y LA INTELIGENCIA ARTIFICIAL
EN CAMPAÑAS ELECTORALES

ALBERTO RENTERÍA

EL PARTIDO DE MÉXICO
IRH INSTITUTO REYES HEROLES
5

USO DEL BIG DATA
Y LA INTELIGENCIA ARTIFICIAL
EN CAMPAÑAS ELECTORALES

COLECCIÓN
CUADERNOS DE CAPACITACIÓN ELECTORAL

TOMO 5

ALBERTO RENTERÍA

USO DEL BIG DATA
Y LA INTELIGENCIA ARTIFICIAL
EN CAMPAÑAS ELECTORALES

EL PARTIDO DE MÉXICO
IRI INSTITUTO REYES HEROLES
5 Cuadernos de CAPACITACIÓN ELECTORAL

Alejandro Moreno Cárdenas
Presidente del CEN del PRI

Alma Carolina Viggiano Austria
Secretaria General del CEN del PRI

Lázaro Jiménez Aquino
Presidente del Instituto Reyes Heroles

Primera edición: 2020
ISBN: 978-607-98995-3-0

© INSTITUTO REYES HEROLES
Insurgentes Norte, 59, Colonia Buenavista,
Alcaldía Cuauhtémoc, C.P.: 06539, CDMX

© PARTIDO REVOLUCIONARIO INSTITUCIONAL

Editor: Javier Sánchez Galicia
Diseño gráficos y portada: Fabián González
Diseño editorial y maquetación: Israel Hernández

Este libro fue realizado como parte de las tareas editoriales para las actividades específicas que lleva a cabo el Partido Revolucionario Institucional, de conformidad con lo dispuesto por el Reglamento de Fiscalización del Instituto Nacional Electoral.

Queda estrictamente prohibida su venta o reproducción para fines contrarios a los establecidos en el propio reglamento.
Hecho en México

Índice

Presentación de la colección **9**

Presentación **11**

Introducción **13**

1 La nueva sociedad hiperconectada **17**
1.1 ¿Cómo es la sociedad de hoy? 20
1.2 ¿Cómo deberían ser los políticos? 24

2 De la investigación tradicional al big data y análisis de audiencias **29**
2.1 Los datos 29
2.2 "Acepto los términos y condiciones": la entrega de nuestros datos 45
2.3 ¿Cómo se usan los datos? 48
Segmentar información *52*
Visualizar *53*
Comprender coyunturas *54*
Fake news y deepfake *56*
2.4 Procesar y utilizar el big data: data mining 56

3 Big Five e inteligencia artificial **65**
3.1 Inteligencia social digital aplicada en la política 71
3.2 ¿Cómo definir un proyecto de este tipo? 73
3.3 Monitorización de hashtags 74
3.4 Mejores prácticas 75
3.5 Mejores prácticas del social listening 76
3.6 De la inteligencia social digital al uso de la #IA para el perfilamiento de audiencias 76

Conclusiones **81**

Bibliografía **85**

Cuadernos de
CAPACITACIÓN
ELECTORAL

Presentación de la colección

Es evidente que los partidos políticos están en crisis en todas partes. Sobrevivir políticamente en un mundo de grandes desafíos y transformaciones requiere altas dosis de compromiso e inteligencia. Un partido como el Revolucionario Institucional debe repensarse y renovarse permanentemente para estar a la altura de los tiempos y mantenerse como una alternativa viable para los mexicanos.

Llegó la hora de asumir seriamente que México cambió y que nosotros, como partido, también debemos hacerlo. Hoy más que nunca tenemos que ser un partido moderno, cercano a la gente, un partido que ofrezca soluciones, un partido que conecte con los jóvenes y enarbole las demandas de las mujeres, un partido que fije posiciones progresistas e informadas sobre los grandes problemas nacionales y globales, tales como el Covid-19, el cambio climático, los derechos humanos, el medio ambiente, la inteligencia artificial, entre muchos otros.

El PRI siempre ha tenido claro que la política tiene que profesionalizarse. Eso nos diferencia de las demás opciones políticas. Nadie le ha apostado como nosotros a un proceso de actualización permanente, a la formación de liderazgos. Demos pasos importantes en la preparación de nuestros cuadros, con herramientas útiles para el análisis de los fenómenos

políticos, sociales y electorales. Con ese propósito, nuestro partido, a través del Instituto de Formación Política "Jesús Reyes Heroles", pone a disposición de sus militantes y simpatizantes, y del público en general, la presente Colección editorial "Cuadernos de Capacitación Electoral".

Alejandro Moreno Cárdenas

PRESIDENTE DEL CEN DEL PRI

Presentación

Con el escándalo mayúsculo de Cambridge Analytica, en el que la consultora empleó datos de millones de usuarios de Facebook para influir a los votantes en las elecciones presidenciales estadounidenses, el poder del *big data* se reafirmó una vez más. Los partidos políticos, en aquel país, se nutren de estas empresas especializadas en *marketing* para llevar a cabo campañas que otrora eran propias de compañías publicitarias y que, hoy, se utilizan para entender al electorado a partir de sus interacciones digitales.

El asunto es tan delicado que muchos hablan de la tiranía de los datos (Han, 2020) o del algoritmo como nuevo instrumento de ingeniería o control social al servicio de las élites que dominan la política y la economía global. Una publicación reciente, titulada *La era del capitalismo de vigilancia* (Zuboff, 2020), sostiene que la expansión de las tecnologías de la información y la comunicación en nuestro mundo introduce cambios estructurales en el modelo capitalista vigente. Las personas se han convertido en el producto del mercado, en la medida que las empresas vigilan todos nuestros movimientos y retienen toda la información necesaria para encauzar nuestras preferencias y emociones en función de los intereses de las élites. Para muchos se trataría de un escenario distópico, pero bastantes elementos en nuestra realidad para corroborarlo.

El concepto de *big data* política consiste en captar los enormes volúmenes de información que producen los

votantes —especialmente en medios *online*— con el fin de procesarlos y analizarlos para conocerlos mejor (intención de voto, preocupaciones, gustos, intereses, etcétera). Esto permite predecir su comportamiento y, con base en ello, es posible microsegmentarlos, para luego entregarles propaganda personalizada. De ahí, la importancia de examinar en todas las implicaciones sobre el tema del *big data*, lo cual constituye el objetivo del presente volumen elaborado por *Alberto Rentería*. Más específicamente, en este libro, se estudia el uso del *big data* y la inteligencia artificial en campañas electorales. Para ello, se abordan cuestiones básicas, como la sociedad de la información o la sociedad hiperconectada; la investigación del *big data* y el análisis de audiencias; el *big five* y la inteligencia artificial.

El presente volumen forma parte de la colección Cuadernos de Capacitación Electoral, un esfuerzo editorial que aspira a reunir los principales temas, aportes, experiencias y a los especialistas de la materia electoral indispensables, para proveer una adecuada formación, tanto teórica como práctica, a todos los interesados en la comunicación política, ya sean profesionales o principiantes. Con ello, el Instituto de Formación Política "Jesús Reyes Heroles" cumple sus objetivos centrales: contribuir a la formación política de los mexicanos, en general, y a la capacitación de los profesionales de la política, ya sean consultores o cuadros partidistas, en particular. Sirva la ocasión para agradecer al presidente del Comité Ejecutivo Nacional del Partido Revolucionario Intitucional, Alejandro Moreno Cárdenas, por su apoyo decidido de impulsar este tipo de empresas culturales que se suma a muchas otras realizadas por el PRI.

Mtro. Lázaro Jiménez Aquino
Presidente del Instituto de Formación Política
"Jesús Reyes Heroles"

Introducción

El *big data* ha llegado a la política, por lo cual la lectura de los datos permitirá entender el comportamiento de los usuarios-electores. En los próximos años, las estrategias de los partidos políticos y sus candidatos podrán segmentar, apuntar y disparar a objetivos específicos del mercado electoral, de manera tan precisa como inaudita. Entre los beneficios de utilizar la data se encuentra la predicción del comportamiento electoral de la audiencia, pues permite extraer gustos e intereses para diseñar programas políticos ajustados a sus necesidades. Además, mide las opiniones sobre problemas que importan a la soecidad y provoca una mayor interacción entre la política y los ciudadanos.

Es innegable que los usuarios tienen dudas sobre lo que podría pasar con sus datos personales. Así, el sonado caso de Cambridge Analytica, incendiarios y redes sociales llevaron a Facebook a su mayor crisis de reputación. La Comisión Federal de Comercio de Estados Unidos (FTC) investigó a esta red social por compartir datos de cincuenta millones de usuarios a la compañía británica. Con ello, Cambridge Analytica utilizó la información para predecir las decisiones de los votantes e influirlos a favor del candidato presidencial Donald Trump.

En los últimos años, el uso de datos ha tenido un impacto sin precedentes en los procesos electorales de todo el mundo. Conceptos como *big data*, *thick data*, *bots*, *chatbots*, inteligencia artificial, *analytics*, manejo de redes sociales son esenciales para identificar tendencias, dirigirse a públicos

específicos e influir directamente en la intención de voto. Durante la crisis global del coronavirus, el *big data* ha sido una de las herramientas más utilizadas para paliar los efectos de la pandemia. China la utilizó para identificar pacientes infectados en tiempo real y contener la propagación de COVID-19, aunque el caso del gigante asiático presenta matices polémicos por cuestiones de vigilancia y control social. También Corea del Sur usó técnicas de *big data* y análisis de datos para evitar la propagación. Y lo propio están haciendo ahora distintas naciones de Occidente.

La pandemia provocada por el Coronavirus generó miedo e incertidumbre en el mundo. Ante el confinamiento, que parece no tener fin, solo podemos estar seguros de que nos enfrentamos a un cambio radical de hábitos y patrones de conducta. Pasamos por una etapa de transición en la cual todos estamos conectados, redefiniendo prioridades vitales. Esta redefinición nos obliga como gobiernos, como políticos y como consultores, a conocer a fondo el comportamiento de las nuevas audiencias; escucharlas para que nos escuchen. Esto significará revolucionar los mensajes y los modos de comunicarlos; investigar, analizar, interpretar sus necesidades, sueños, anhelos, miedos y comportamientos será el proceso que marcará la diferencia entre un buen o un mal diseño de campaña política, ya sea para comunicar, posicionar o contener. Una comunicación sustentada en datos será, además, el camino adecuado para lograr un cargo de elección popular.

En este cuaderno de capacitación electoral nos referiremos al nuevo paradigma, el cual parte de una evidente sociedad hiperconectada, con la posibilidad de generar millones y millones de datos, pero, sobre todo, procesarlos en tiempo nunca antes imaginado. Leernos a través de los datos nos permite reconocer distintas audiencias a las cuales no llegaremos con los mensajes generales que aún se utilizan en campañas tradicionales.

1 La nueva sociedad hiperconectada

Fue en 1968, en la campaña de Robert F. Kennedy a la Presidencia de Estados Unidos de América, cuando se utilizó la tecnología más avanzada en materia de *marketing* político. El ejercicio para identificar la intención de voto entre los estadounidenses se llevaba a cabo a través de llamadas telefónicas. Para ello, se utilizaba un sistema análogo de tarjetas que contenían los datos de cada uno de los votantes. La pregunta que realizaban los encuestadores era por quién votaría la persona. La tarjeta presentaba cuatro respuestas que podían ser clasificadas. Si declaraban el voto a favor de Kennedy, se colocaba la clave 1, se agradecía al interlocutor y se clasificaba. Si el votante no estaba seguro, pero se identificaba o tenía cierta inclinación hacia Bobby Kennedy, se tenía un discurso preparado cuyo objetivo era intentar convencer al votante y, nuevamente, se clasificaba la tarjeta con las claves 2 o 3. Por último, si el votante declaraba apoyar a Jimmy Carter (el contrincante), lo clasificaban con la clave 4 y la tarjeta era separada.

Desde 1968 a la fecha, más allá de la asombrosa evolución de la tecnología, la utilización de los *call centers*, las encuestas telefónicas y robotizadas, las encuestas tradicionales

y los *focus groups*, ¿hemos reflexionado sobre el hecho de que no ha habido un cambio significativo en la forma de hacer campañas políticas? Y esto a pesar de que la tecnología ha dado pasos agigantados desde 1968 hasta la fecha. Tanto es así que su evolución ha rebasado al *marketing* político actual.

Ahora, la tecnología nos ofrece muchas maneras de usar e interpretar datos, sin embargo, pareciera que seguimos pensando de forma análoga, cuando actualmente la sociedad está hiperconectada.

Las campañas políticas de las cuales seremos testigos en este momento histórico requieren de una nueva perspectiva: el poder está cambiando de manos y de forma. El poder del siglo XXI lo tenemos en un dispositivo de 14 pulgadas, es decir, será preciso ejecutar campañas políticas de 14 pulgadas. Si lo hacemos, dejaremos que los usuarios sean quienes tengan el poder de incidir, trabajar, decidir y comunicarse políticamente a través de los distintos medios y plataformas a las que tienen acceso.

Esta capacidad, que también es una forma de poder y de empoderamiento, ha transformado las interacciones de una manera obligada y aparentemente sutil. En la práctica, la clase política y la sociedad siguen dividiéndose; sus intereses y preocupaciones parecen no tener puntos de convergencia. La sociedad no quiere a los políticos ni cree en ellos, el

desencanto es notable y lo vemos en los bajos porcentajes de participación de la ciudadanía en las elecciones.

Entre tanto, los políticos no acaban de entender las demandas ciudadanas desde los altos puestos en que se encuentran, donde la realidad es otra, "distinta y distante". La cerrazón de la clase política y el mantenimiento del *statu quo* están incrementando la distancia cada día más. La clase política no reconoce que ya hubo un cambio en el contexto. La tecnología y la vida digital han impactado en nuestras conductas, intereses, expresiones, demandas y miedos. Somos una sociedad hiperconectada y requerimos de políticos que sepan escucharnos a través de esta caja de cristal de la vida digital.

1.1 ¿Cómo es la sociedad de hoy?

La ciudadanía se ha empoderado con la accesibilidad a la tecnología y el uso de las redes sociales. Desde sus *smartphones*, los ciudadanos vigilan, denuncian, critican, opinan y participan sobre los temas que les interesan.

Este empoderamiento pareciera hacer corto circuito con los gobiernos, con los partidos y con los políticos, porque a ellos no les interesa escuchar a la población ni entender que los ciudadanos poseen la capacidad de comunicar, de generar información y, con ello, tendencias, críticas, propuestas y soluciones.

Quizás esta sea una razón por la cual la población se ha vuelto más desconfiada, escéptica, exigente y crítica. Mientras los políticos no se transformen y se incorporen al nuevo paradigma, la brecha seguirá ampliándose, y seguirán sin conectar con la gente.

No obstante, la respuesta está al alcance de nosotros, mucho más cerca de lo que creemos. Actualmente, podemos saber qué piensan, qué desean y qué temen los ciudadanos. Los retos de las campañas que siguen son escuchar y saber atender sus necesidades y demandas. Estar a la altura de sus expectativas y no decepcionarlos. Políticos, gobernantes, analistas y consultores tendrán que evolucionar hacia este nuevo paradigma en el que la sociedad hiperconectada deja de ser una masa estandarizada, para convertirse en individualidades que expresan sus opiniones, reclamos, sentimientos y temores. En este contexto, los individuos se agrupan en torno a intereses, gustos y perspectivas afines, y cada grupo requiere un trato adecuado y personalizado para sentirse atendido, reconocido y valorado.

A partir de ahora, debemos considerar que la parte de la población que posee dispositivos electrónicos y acceso a la red está conectada en promedio 8 horas 23 minutos al día (Global Digital Overview, 2020). El dato anterior se publicó antes de la pandemia del COVID-19. A partir de ella, es probable que estemos conectados casi el doble de tiempo. Sin duda, este nuevo escenario implica un cambio profundo y significativo en el comportamiento humano.

Cuando alguien se me acerca como consultor y me dice: "quiero hacer comunicación política para 'usuarios digitales'", vemos que en la misma solicitud hay un sesgo. Lo cierto es que, como mencioné anteriormente, hoy vivimos en una *sociedad hiperconectada*, podría, incluso, atreverme a afirmar que la mayoría del electorado es digital. Más adelante me referiré al volumen de esta sociedad hiperconectada.

Debemos partir de la idea de que la población está conformada por ciudadanos informados. Pero, ¿qué significa "estar informados" en esta nueva realidad en la que la información es intensiva y extensa, y se transmite en tiempo real en todo el mundo? Actualmente los algoritmos producen en nuestro entorno un sesgo de la realidad. Se conecta a sus redes sociales y se "informan" desde su burbuja, de modo que se junta con personas que piensan igual que tú o que tienen gustos,

preferencias y opiniones similares a las tuyas. Esta situación impide ver, conocer y entender puntos de vista diferentes. En el mejor de los casos, ampliamos un poco nuestra búsqueda, pero eso es todo. Tal sesgo proviene de los algoritmos, y se forma a partir de los clics que hacemos.

Profundizaré en el uso de los algoritmos más adelante, pero quiero plantear por ahora que *la ciudadanía está informada dentro de burbujas*, es decir, en la actualidad no hay una forma orgánica en la que puedas comunicarle a todos el mismo mensaje. Me refiero a una campaña política general y totalizadora que signifique para todos, porque todos nos encontramos —debido a esta sociedad hiperconectada— en burbujas distintas.

Los ciudadanos también son *vigilantes*, porque permanentemente están viendo qué sucede. Les interesan algunos temas, muchos de ellos artificialmente inducidos. Esa vigilancia puede generar tendencias mundiales, como la que se produjo con la publicación del video que documenta el abuso policial en contra del afroamericano George Floyd, en Minneapolis. La publicación de este video se unió a un *hashtag* histórico que denunció y visibilizó el racismo en el cuerpo policial del país más rico del mundo. Ahí, se evidenció cómo la sociedad se ha vuelto vigilante. Sin duda, esto significa un riesgo y un reto para los gobiernos y los políticos de todo el orbe. En este nuevo paradigma, deben sentirse vigilados a todas horas, y, por lo tanto, atender de manera oportuna y pertinente a esta ciudadanía vigilante, crítica y exigente.

Por otro lado, los ciudadanos son *colaborativos*: si les pedimos ayuda para alguna causa, para denunciar alguna anomalía, para participar en alguna actividad o para compartir su conocimiento, podrán ayudar en lo que se requiera, si tienen a la mano la herramienta adecuada.

NUEVO CONTEXTO

1.2 ¿Cómo deberían ser los políticos?

No cabe duda que esta nueva ciudadanía nos obliga a modificar la manera de hacer política y de convencer a los votantes. En primer lugar, partamos de que la política debiera ser *permanente*. Es decir, se tendría que trabajar mientras la población está conectada: 24 horas, los 365 días del año. Los políticos deberán estar permanentemente monitoreando lo que sucede en la vida digital/real de la sociedad: no hay horarios.

Otra característica de la política es que es *omnisciente*, es decir, tiene que estar viéndolo todo, percibiéndolo todo, para poder generar información. Hoy, lo podemos constatar en la forma en que los gobiernos del mundo han manejado la pandemia. En medio de una de las crisis más dolorosas y profundas de la historia, los gobiernos y los políticos no han comprendido el problema más importante que hay. Tal y como lo ha comentado Mario Riorda en días recientes, no estamos hablando de hacer una comunicación de crisis, sino de hacer una comunicación *de riesgo*.

De igual manera, necesitamos *políticos plataforma*. Esto significa que los políticos tienen que ser como una aplicación. Y habrá que preguntarnos, ¿cómo queremos que funcione una aplicación en nuestros dispositivos electrónicos? Siguiendo esta misma analogía, nuestra respuesta es sencilla: los políticos tienen que ser como WhatsApp, funcionar como Telegram; tienen que ser rápidos, ágiles y con capacidad de respuesta inmediata.

En la actualidad, la dolencia de la política como acción pública, y de los políticos como agentes de representación y escucha de lo ciudadano, es la falta de velocidad en las reacciones para responder a las demandas de la ciudadanía. A todas luces, se sigue viviendo en el paradigma anterior.

Entonces, ¿cómo tendría que ser la comunicación política en la sociedad hiperconectada? Pues bien, esta comunicación debe dirigirse de manera directa al ciudadano, llegándole a través de distintos medios, utilizando diferentes mensajes, en distintos lugares, de forma adecuada y personalizada.

Con seguridad, todos hemos escuchado, desde la campaña de Barack Obama en 2008 y pasando después por la de Donald J. Trump en 2016, que se habla de *microsegmentación*. De esta forma, en una sociedad hiperconectada y en un contexto social, donde el entendimiento de lo relevante se ha asentado en la premisa "cada cabeza es un mundo", la pregunta que surge es: ¿Cuántos políticos, organizaciones y representantes han hecho campañas con microsegmentación, y las han implementado? Muchas veces solo nos quedamos con la intención o el discurso. y terminamos armando la campaña con el mensaje.

Sí, la nueva comunicación política debe ser *desintermediada*. La ciudadanía consulta información con menor frecuencia, pues en la actualidad acceden, usan e interpelan a las fuentes de origen con mayor regularidad y cercanía. La ciudadanía digital dialoga con los gobiernos e interpela a los políticos de manera directa, a través de sus redes sociales, sin intermediación.

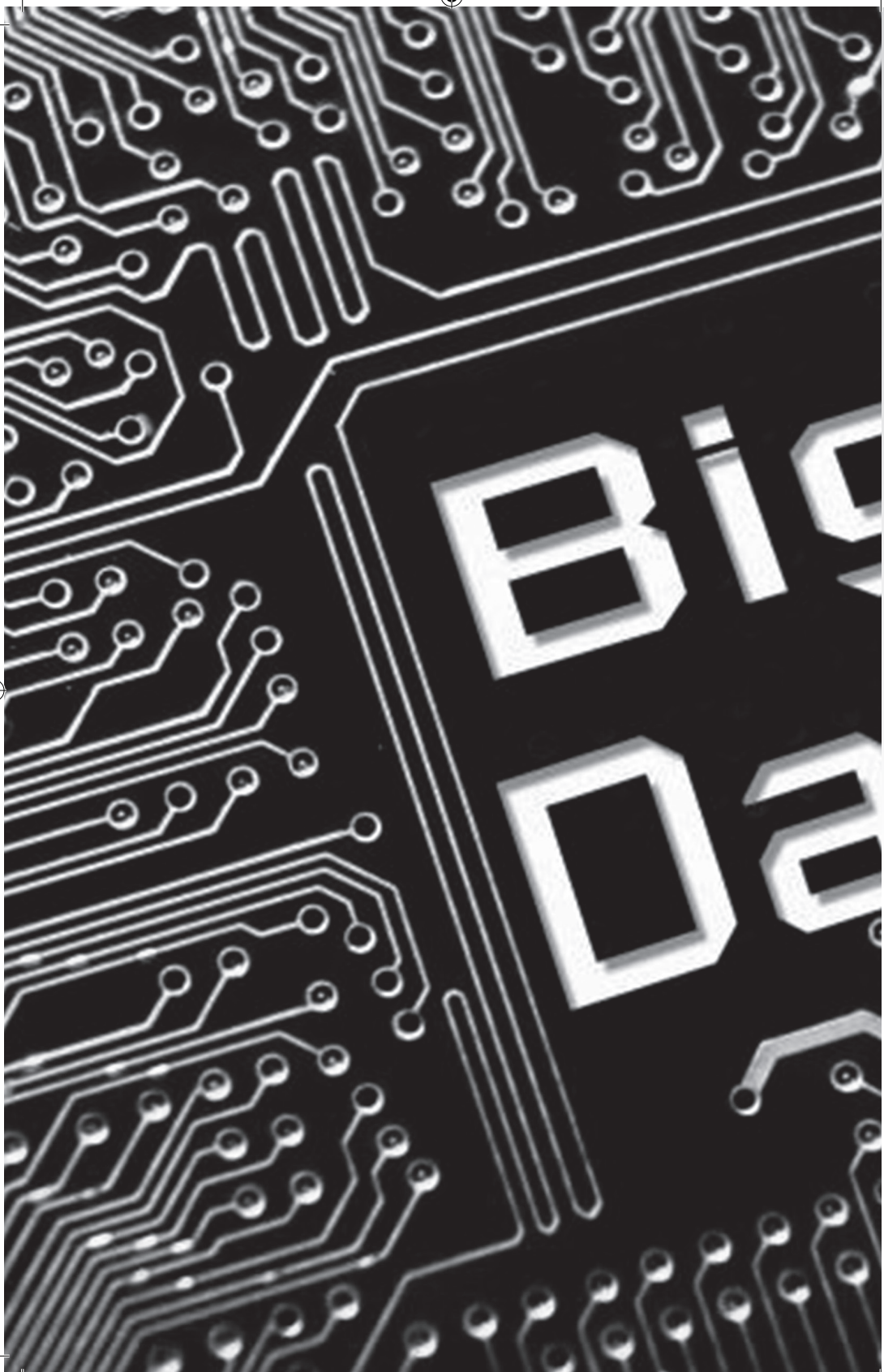
Este es el contexto que ha cambiado desde la campaña de Robert F. Kennedy en 1968 hasta el día de hoy. Lo más desafortunado es que los políticos siguen siendo análogos, en tanto que la comunicación y la sociedad se van transformando a pasos agigantados. Reflexionemos desde nuestro día a día: ¿La política o los políticos de hoy han logrado descifrar o entender este importantísimo cambio? ¿Nos entienden a los ciudadanos como entes singulares con nuestras propias preocupaciones e intereses? ¿Nos ofrecen soluciones como lo harían ante una sociedad demandante e hiperconectada como la que somos hoy?

En este marco, los políticos tienen que aprender y adaptarse a las nuevas tecnologías y a las nuevas plataformas. La comunicación de hoy es ágil y sumamente veloz. De ahí que deban adaptarse a las necesidades de forma permanente. Los ciudadanos tienen expectativas de participar, de contribuir, de ayudar más al bienestar de su sociedad.

Sin duda, en la era poscovid-19, los políticos deberán funcionar como lo hacen las buenas aplicaciones: actualizarse de manera permanente. En otras palabras, los políticos deberán convertirse en una versión *beta infinita*, con el fin de adaptarse rápidamente a los cambios. La velocidad con la que los políticos se adapten será directamente proporcional a su velocidad de crecimiento, y no solo los políticos, nosotros mismos, como consultores, tenemos que estar en permanente adaptación. Debemos abrirnos a la innovación, pues ya es una realidad que la tecnología puede ayudar, a contribuir a escuchar mejor a la sociedad, con la finalidad de atender sus demandas.

La tecnología, como premisa, es un medio, no un fin; este lo determinamos nosotros en función de nuestros objetivos políticos. Sin embargo, hoy buscamos conocer como punto de partida qué esperan los usuarios de nosotros. En respuesta, los ciudadanos quieren políticos que se "actualicen", de acuerdo con sus demandas, que puedan estar en sintonía con sus problemas, preocupaciones y reclamos, que respondan de manera eficiente, que no se traben. Si lo logramos, en definitiva, seremos parte de la evolución, no solo de la comunicación política, sino de la manera de hacer política.

NUEVO CONTEXTO


Big
Da

2 De la investigación tradicional al *big data* y análisis de audiencias

2.1 Los datos

Para la investigación de mercados electorales, contamos con las herramientas tradicionalmente probadas, tales como las encuestas, los *focus group*, los sondeos, entre muchos otros. Las encuestas, por ejemplo, nos permiten visualizar una fotografía de un momento de la realidad. La forma de tomar esta fotografía pareciera no haber cambiado en un largo tiempo, sobre todo en lo que se refiere a su metodología. En cada ejercicio estadístico, preguntamos a los ciudadanos que deseamos convencer, sobre sus problemas más vigentes.

Con la sistematización, análisis e interpretación de estas fotografías vamos tejiendo los puntos de la ruta que nos llevará a nuestro destino. Ante la pandemia ocasionada por el SARS-CoV-2, los *focus groups* se han trasladado al ambiente digital. En tanto, los muestreos han evolucionado a métodos novedosos que también aprovechan la tecnología y se

desarrollan en línea. Sin embargo, en todos estos ejercicios siempre esperamos que la persona que está contestando la encuesta nos diga "la verdad, solo la verdad y nada más que la verdad". Buscamos que nos declare sus intereses y preferencias, que nos descubra lo que trae en la mente y que esa información nos arroje *insights* relevantes para utilizar en la conformación de nuestra estrategia.

Pero ¿cuál es el problema de estas herramientas? Sin desestimar su importancia, debemos reconocer que el problema es que la información que nos están entregando los entrevistados es binaria, es decir, información racional. Y, siguiendo algunas conclusiones del estudio de Stephens Davidowitz, el *efecto* de las respuestas racionales es que no son verdad. Veamos con detenimiento esta conclusión.

En su libro *Todo el mundo miente*, Davidowitz (2014) afirma que las respuestas racionales tienen un sesgo que se denomina el *sesgo de la deseabilidad*. Para llegar a esta conclusión se llevó a cabo el siguiente ejercicio: en Estados Unidos de América, se les preguntó a hombres y mujeres acerca de su conducta sexual. Específicamente, se les preguntó cuántas veces a la semana tenían relaciones sexuales y si usaban o no preservativo. Las mujeres respondieron que tenían relaciones una vez a la semana y que en 20% de las ocasiones utilizaban preservativo. Esto significa que ellas requerirían 1.1 mil millones de preservativos. Cuando se preguntó a los hombres, no faltos de presunción, contestaron que tenían muchas más relaciones y que usaban preservativo la mayoría de las veces. Según sus respuestas, se requerirían 1.6 mil millones de preservativos al año. Los resultados no coinciden y alguien está mintiendo. La realidad es que la industria solo produce seiscientos millones de preservativos al año.

Este es un claro ejemplo que nos muestra la facilidad con la cual los seres humanos mienten cuando están frente a un encuestador o facilitador, ya sea presencial o telefónico.

Las personas guardan sus pensamientos más íntimos para sí mismas. No obstante, cuando se encuentran frente a una computadora y nadie las está observando, entregan datos o información personal, porque los están entregando a un dispositivo móvil o computadora. No tienen el sesgo de lo que el entrevistador pueda pensar de ellas.

Así, lo que se denomina *sesgo de la deseabilidad social*, es decir, "yo quiero que me vean como un miembro relevante de la sociedad, con ciertos comportamientos, con ciertas características", propicia que los ciudadanos ofrezcan información que no proporcionarían en otras circunstancias, y muestren su intimidad a los diversos dispositivos conectados a la red.

Al final, la obtención de datos —que no son otra cosa más que fotografías instantáneas del momento— ocasiona, como se mencionó anteriormente, que nos convirtamos en poseedores de lo más íntimo del ciudadano. Si alguien preguntara: ¿quiénes ven pornografía? Por supuesto que todos o una gran mayoría diría que nadie.

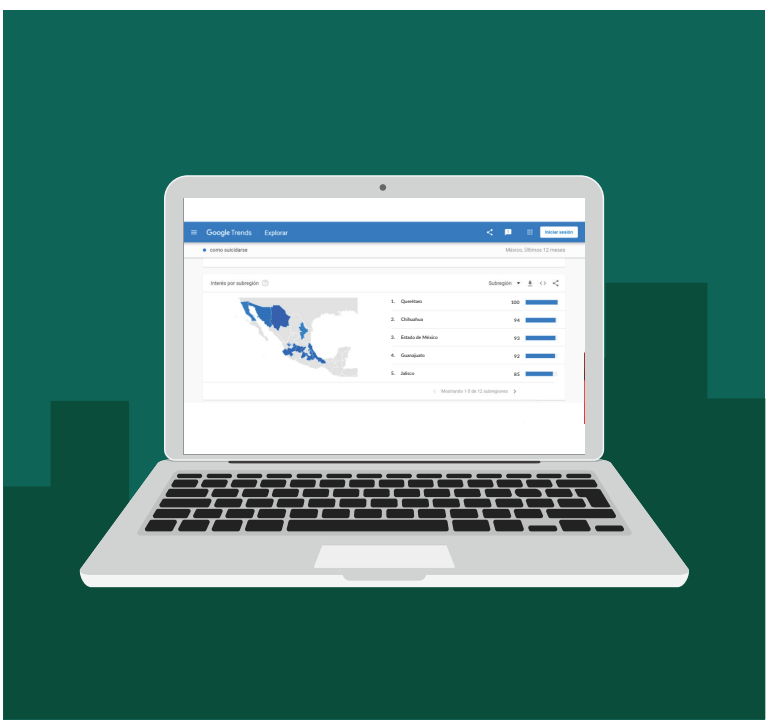
Muchas personas no responderían a esa pregunta, pero es completamente diferente cuando las personas teclean y realizan búsquedas en Google. Al hacerlo están entregando una información íntima a alguien que puede investigarles con mayor profundidad que un encuestador.

Esa es la diferencia entre lograr compartir la información, tener datos con algunos sesgos en las variables de los métodos tradicionales o usar este nuevo paradigma que tiene que ver con el uso de los algoritmos, la inteligencia de audiencias o *big data*.

Todo el mundo miente. Todo el mundo embellece el currículum. Pero, ¿qué sucede cuando las personas le dicen a Google o a cualquiera otra plataforma digital cosas que no le confiesan a nadie más? Entonces la intimidad está siendo revelada en los medios sociodigitales.

Detengámonos a pensar cuántas fotos de platillos que nos gustan hemos compartido en los últimos seis meses, cuántas fotografías de lugares que visitamos, cuántas consultas hemos realizado acerca de cosas que nos preocupan o que nos inquietan. Por ejemplo, actualmente, se han disparado las búsquedas en México sobre qué hacer si alguien enferma de coronavirus.

De esta manera, puedo ir viendo de forma específica el comportamiento de la población. Echando mano de simples herramientas que cualquiera de nosotros puede utilizar, como Google Trends, podemos empezar a buscar información que nos vaya dando indicios de cómo piensa nuestra audiencia y lo que quiere. O, ¿por qué no pensar —léase como un ejemplo pragmático— en cómo podemos mejorar políticas públicas?



Si yo preguntara a la población cómo ayudar a prevenir el suicidio en México, seguramente cada uno tendría respuestas distintas: "hay que revisar los datos", "hay que hacer campañas de prevención", "hay que buscar en dónde hay más índices de suicidio", etcétera. Sin embargo, ¿qué sucede si cambiamos un poco el chip y lo que hacemos es ir a Google Trends y buscar la frase "cómo suicidarse"? Entonces, Google Trends nos dirá en qué ciudades se está buscando con mayor recurrencia información acerca de cómo suicidarse.

¿Qué nos indican estos recursos? Los datos están ahí. Se trata de datos públicos, de información que solamente tenemos que ir a buscar en Google, y que se pueden encontrar en cada uno de los países. Es casi seguro de que si existiera el dato de dónde se ha dado el mayor número de suicidios, de enero a la fecha, prácticamente podría haber una coincidencia muy relevante en esta información. Los informes no solamente nos dirán en dónde y en qué ciudad ocurre un fenómeno, sino también cuál es la búsqueda específica que se ha realizado.

Reiteremos: la información es pública y accesible para todos. Google Trends brinda la información más relevante acerca de las tendencias de búsquedas de la audiencia con la que se desea establecer una interacción, una comunicación. A veces, anclados en nuestros usos, pasamos informaciones, datos y carga de procesamiento en las áreas de comunicación o en las de la consultoría política. Nos vamos a lo más complejo, pero no nos damos cuenta de que mucha de esta información está disponible a un clic, en la computadora o en un *smartphone*.

Imagínense, si ustedes buscaran información sobre temas como violencia de género o cómo evitarla, lo que encontraremos es un mapa de cada uno de los países, específicamente en donde hay mayor violencia de género. Así, no tienes que buscar los grandes algoritmos, porque al final de cuentas

nadie te dirá si se quiere suicidar. "¿Ha pensado usted en suicidarse?, ¡por supuesto que no!".

El sesgo de la deseabilidad social dice que no confesamos nuestra intimidad, pues ello es políticamente incorrecto. Sin embargo, la estamos declarando todo el tiempo, por poner un ejemplo, en Google. Todas nuestras búsquedas y clics cuentan, y las entregamos sin chistar al investigador más infalible. Pensemos acerca de lo que le hemos "confesado" a Google desde que lo utilizamos.

Cuando nos preguntamos, ¿por qué estas empresas tecnológicas han adquirido tanto valor? La respuesta es simple: porque conocen nuestros datos, porque saben a qué hora nos conectamos normalmente, a qué hora nos despertamos, a qué hora hacemos ejercicio, o a qué hora hacemos nuestra primera búsqueda del día.

Les invito a que revisen su historial de búsquedas de un año a la fecha, y al final recibirán una importante lección sobre la transformación de la información y el poder que confiere a quien la posee. Los algoritmos están empezando a saber mucho más de nosotros que nosotros mismos. Hoy en día existen muchas plataformas que contienen todas nuestras declaraciones en medios sociodigitales, no solo es Google.

Pensemos acerca de cuántos de nosotros damos nuestro correo electrónico de Gmail, o accedemos con nuestra cuenta de Facebook o Twitter para utilizar, cuando menos, otra plataforma. Es decir, usamos una cuenta para "loguearnos" en Netflix, Spotify, Amazon Prime, LinkedIn, TikTok, Uber, Rappi, etcétera. Cada vez que accedemos a ellas, lo que hacemos es sumar capas de datos, capas de información a cada una de ellas.

Y entonces, sí es verdad que el poder está cambiando de manos. De manera que si el poder está en los ciudadanos que forman parte de una sociedad hiperconectada, y el poder se encuentra también en los datos, ¿por qué seguimos pensando en hacer las campañas políticas de siempre?

Cabe preguntarnos si queremos hacer campañas políticas como siempre con el habitual manejo de crisis o si queremos evolucionar a este nuevo paradigma, usar los datos (legalmente obtenidos) y entender cómo se generan, cómo se ordenan y cómo podemos utilizarlos para comprender mejor a los electores.

Por otro lado, el uso de datos con fines de *marketing* no es nuevo. Google Trends tampoco es nuevo. En el año 2008, en la campaña de Barack Obama a la Presidencia de los Estados Unidos de América, el director de campaña, Jim Messina, decía que los datos son la mejor vía para alcanzar la verdad. En la campaña, la mayoría de las personas, de los analistas y de los profesionales se conformaron con la revelación de cómo se utilizaban los datos para generar mensajes. Pero lo realmente relevante fue que la data se compartía en todas las áreas de campaña, con la finalidad de tomar todas y cada una de las decisiones.

Cuando hablamos de data no hablamos solamente de la información que se genera o que se clasifica como producto de una encuesta, sino de toda la información que se va generando en el proceso de campaña. Para ejemplificar, podemos recordar el concepto de lo que se llamó "La Cueva de Obama". Expertos matemáticos e informáticos trabajaron secretamente en Chicago, durante casi dos años. Fueron ellos quienes cruzaron toda la data e interpretaron y tradujeron en *insights* toda la información.

Sin lugar a duda, otro de los grandes retos es lograr que la metodología de investigación tradicional pueda conjuntarse con la inteligencia social digital. Esto significa conseguir una sinergia entre investigación, *big data* e inteligencia artificial. Los datos que se capturan en territorio, por ejemplo, tendrían que sistematizarse. Es lo que le da pies y manos a la información obtenida a través de la data. En otras palabras, se

necesita un gran volumen de datos de distintas fuentes para ayudar a tomar decisiones.

La Cueva de Obama se dedicaba a rastrear comportamientos de amigos, a buscar tendencias, a segmentar criterios sobre votantes. Así, demostraron que el ganador sería el candidato que más compartiera afinidades y preocupaciones de manera general con la mayoría de la población.

¿Qué pasa en muchas de las campañas políticas? Normalmente empieza el periodo de la preprecampaña y vamos recopilando datos. Sin embargo, nunca acabamos de sistematizarlos en un solo bloque, de clasificarlos o analizarlos con una sola metodología; de catalogarlos en un solo lugar.

En muchas de las campañas políticas, la información es pobremente sistematizada y los datos, por lo tanto, terminan dispersos y se vuelven inútiles: "¿dónde está el archivo de Excel?", "¿dónde está la base de datos?" No, tienen sistemas de gestión política, sistemas de gestión de proyectos, sistemas electorales, pero cada uno de ellos es un silo, como en el sistema digital.

Los silos no se comunican entre sí y todo resulta en una pérdida de información y de valor. Al no haber comunicación entre los silos, se pierde el valor potencial de la articulación de toda la información, tanto territorial como digital. Entonces, surge la pregunta obligada: ¿cómo hacer para que todos entiendan la importancia de los datos?

Otra parte muy importante es que las campañas ganadoras, como la de Barack Obama, se ganan con anticipación. En México, por ejemplo, en algunos estados las campañas durarán treinta días, lo cual antes no sucedía. Convencer a los ciudadanos en un mes es tarea imposible, por ello, será importante trabajar de forma anticipada, con el objetivo de conocer quiénes son los votantes potenciales.

Martin Hilbert, alemán experto en redes, sostiene que los datos son el nuevo petróleo. Quien tiene el control de

los datos —afirma— controlará Estados y naciones, o por lo menos dominará su sector. Aquí es donde nos damos cuenta de la verdadera importancia de las empresas tecnológicas, que son las más redituables y las de mayor valor en bolsa actualmente.

¿Por qué valen tanto estas empresas? ¿Es por el *core* de negocio o por los datos que poseen y que utilizan? Es común pensar que Amazon vale más que Google, por sus ingresos, pero, en mi opinión, Amazon vale más por la información que tiene de nosotros —además de todas las utilidades que obtiene—.

También empezamos a ver un nuevo tipo de guerra que se relaciona con la geopolítica de los datos, es decir, con cómo vamos a pelear el esquema de datos. La reciente batalla entre Samsung y Huawei se dio por el control del sistema operativo. ¿Por qué peleaban por el sistema operativo? Por la importancia de tener Google en su sistema. Los datos, los datos, los datos.

Supongamos lo siguiente: si yo tengo un teléfono chino y un sistema operativo americano, ¿quién tiene mis datos? ¿Qué gobierno puede acceder a mis datos? Si tengo un teléfono de procedencia china y voy a tener un sistema operativo chino, el gobierno que es dueño de mis datos es China. En este tenor, lo que vamos a empezar a ver son *guerras* entre los gobiernos por obtener la mayor cantidad de datos.

Lo hemos atestiguado en las recientes protestas en Hong Kong: cómo se identificaba a las personas a través del uso del *big data*, cómo los inconformes debían usar mascarillas para ocultarse y protegerse, cómo tenían que hacerlo para que el gobierno no pudiera identificarles. Este es el nivel de invasión que tendremos en el futuro cercano. Quedémonos con esta

idea: *los datos son el nuevo petróleo*. ¿Qué hacemos con ellos? ¿Qué hacemos con la información que nos llega día a día?

En décadas anteriores, era muy fácil manejar la información que nos llegaba, pues teníamos y producíamos pocos datos. Sin embargo, hoy en día, la *infoxicación* hace que tengamos y generemos millones de datos. ¿Cuánta data generamos al día? ¿Cuántos correos enviamos diariamente? ¿Cuántos mensajes? ¿Cuántas veces posteamos contenidos? ¿Cuántas veces subimos fotografías?

No cabe duda de que hoy vivimos en la era del *dataismo*. Yuval Noah Harari, uno de los pensadores más relevantes y brillantes de este siglo, y autor de libros como *Homo Deus*, *Homo Sapiens* y *21 lecciones para el siglo 21*, afirma que con el paso del tiempo confiamos más en los macrodatos y en los algoritmos computacionales, los cuales están sustituyendo a la sabiduría humana. ¿Estamos dejando las decisiones a los datos? Es decir, ¿una inteligencia artificial, un algoritmo, decide por cada uno de nosotros? ¿Es posible?

El ejemplo más claro de cómo los algoritmos nos están desplazando es la aplicación de tráfico Waze. Si queremos ir del punto A al punto B, le cedemos nuestra capacidad de intuición y decisión a un algoritmo, que decide por nosotros. Esto es solo un ejemplo de cómo vamos permitiendo que nuestra información y nuestras decisiones estén controladas por algoritmos.

¿Qué pasa cuando cedemos a un algoritmo algo tan básico como la ruta que nos llevará del punto A al punto B? Eso es el dataismo y eso es lo que estamos viviendo. Cada vez más, los usuarios están permitiendo que los algoritmos incidan en sus decisiones.

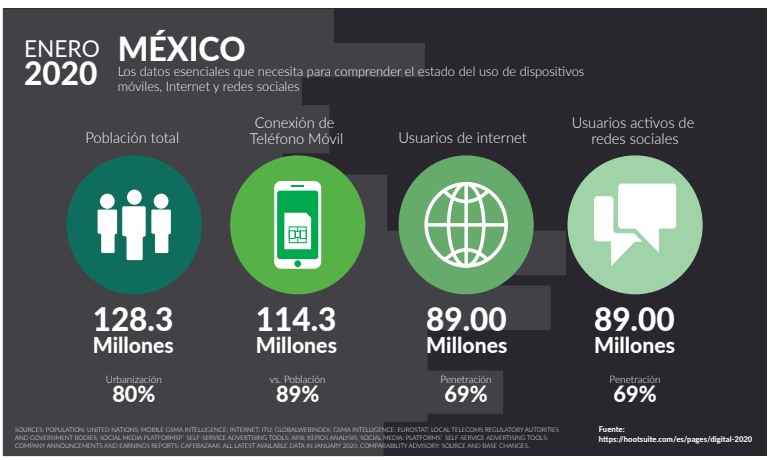
Cuando nos conectamos a Waze para que elija una ruta de viaje por nosotros, también solemos conectarnos a Spotify y enlazamos nuestro viaje. Así, no solo accedemos a dar el dato de las trayectorias que usamos para llegar del punto A al

punto B, sino también qué nos gusta escuchar del punto A al punto B y si siempre lo escuchamos. En un día de navegación, Waze sabrá los comportamientos de movilidad de los usuarios y sus hábitos musicales. Esta data es sumamente valiosa, porque, después de poco tiempo, resulta la rutina de comportamiento de un usuario: cómo se traslada, cómo se mueve en la ciudad, entre otras cosas que servirán para generar un perfilamiento.

Al dejar que los algoritmos decidan por nosotros, utilizamos nuestra memoria con mucha menor frecuencia, dando paso al efecto Google. Estamos cediendo nuestra capacidad de memorizar eventos importantes al recordatorio de un calendario o una red social. Si el calendario no se activa o Facebook no notifica, las consecuencias pueden ser muy desagradables. ¿Cuántos nos aprendíamos los números telefónicos de amigos y familiares antes de que existieran los teléfonos móviles? ¿Cuántos teléfonos pueden recordar?

Al permitir que los algoritmos decidan o incidan en nuestras elecciones personales, dejamos en manos de terceros nuestra memoria y nuestra capacidad de intuición. En tanto que nuestros rastros digitales están a disposición de todos aquellos que quieran utilizarlos para *x* o *y*.

El conjunto de los datos que dejamos en la red tiene un alto valor. En el año 2012, la compañía estadounidense IBM afirmaba que cada persona producía, en promedio, 500 MB de datos cada año. Se estima que para el año 2025 cada ser humano generará 62 GB, o más, de datos de información. Hoy, nos encontramos cerca de este tamaño de data, porque el volumen de información que creamos es cercano a esas dimensiones en un solo día.



En enero de 2020, en México, había 128.3 millones de personas, de las cuales, 114.3 millones estaban conectadas mediante un teléfono móvil. De estos, al menos 90% eran *smartphones*, es decir, dispositivos que tienen acceso a WhatsApp, Facebook, Instagram y, en ocasiones, a Twitter. En zonas semiurbanas y rurales puede no haber internet, sin embargo, tienen conectividad a través de dispositivos móviles.

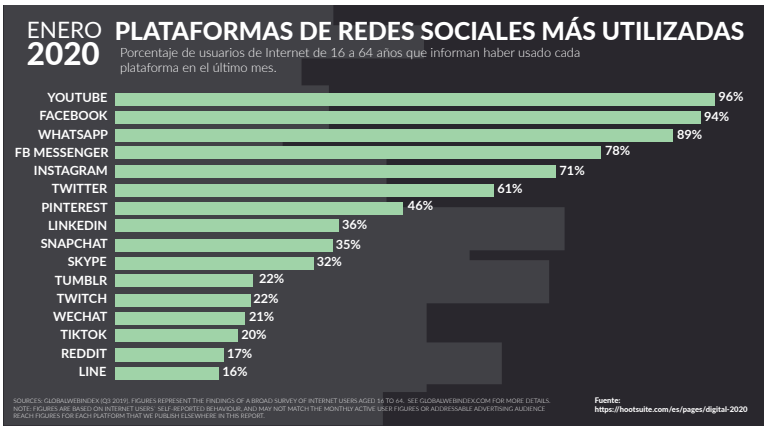
Esto significa que incluso si no hay internet, se entregan datos en los móviles, mientras pasamos mucho tiempo conectados.

Por otro lado, es preciso considerar el tiempo que pasamos en redes sociales. En promedio, estamos conectados a la red mundial 8 horas 23 minutos, de las cuales son tres en redes sociales. No obstante, la pandemia ha aumentado el tiempo de conectividad hasta llegar a la mayor parte del día.

El uso de las redes sociales se incrementó, así como el tiempo que las personas pasamos viendo la televisión, y el tiempo que dedicamos a escuchar música. Si todo el tiempo estamos hiperconectados, entonces los políticos tienen un reto: saber qué pasa al menos en esas 8 horas con 23 minutos, y en esas tres horas con 25 minutos. Deben averiguar qué está sucediendo, que están pensando y diciendo los usuarios y, sobre todo, deben identificar dónde lo están diciendo.

Los mexicanos consumimos mucho entretenimiento a través de la plataforma de videos YouTube, sin embargo, Facebook, Whatsapp, Messenger e Instagram pertenecen a una sola compañía: Facebook. Habrá que preguntarnos quién domina realmente los datos. No es YouTube, sino Facebook, es

a esta red social a la que los usuarios le entregan sus datos desde distintas plataformas. Después, tenemos a Twitter, Pinterest y demás redes sociales.



La plataforma de videos cortos TikTok, que este año ha tenido una penetración de 20%, el próximo estará cerca de los porcentajes de LinkedIn. También estamos presenciando que todas las generaciones entran al juego de TikTok, y no solamente en México, sino en toda América Latina. Quienes antes decían "es una red para chavitos", ahora también la disfrutan.

Todos los segmentos poblacionales están jugando en la hiperconectividad, y las brechas generacionales se vuelven borrosas. Al mismo tiempo, todos estamos dejando información íntima y valiosa. En sesenta segundos, en el mundo, ingresamos a nuestras cuentas un millón de veces, se intercambian 41 millones de mensajes, vemos casi setecientas mil horas de Netflix o de televisión, hay 1.4 millones de *swipes* en Tinder, se hacen casi cuatrocientos mil *scrolls* en Instagram. Son sesenta segundos de datos permanentes, es decir, de información que se genera todos los días, de hábitos de consumo, costumbres, antipatías, filias, fobias, gustos, deseos y tipos de comportamiento de cada uno de los ciudadanos.

Si en sesenta segundos podemos identificar toda esta información, surge el reto de sistematizarla, clasificarla, detallarla, elaborar *insights* a partir de ella. Bajo un correcto protocolo de procesamiento, podríamos ser más asertivos en la interacción y comunicación con los ciudadanos de lo que actualmente logran los medios tradicionales.

Se trata de empezar a llevar a cabo un cambio de paradigma, y uno de los fines de este cambio será reflexionar acerca de los usos, beneficios y perjuicios de esta tecnología. En el transcurso de estos 60 segundos vamos dejando nuestros rastros digitales. Estos rastros van construyendo un patrón

en función de lo que hacemos, buscamos, jugamos, escuchamos, compramos, etcétera.

No obstante, los rastros digitales provenientes del uso de redes sociales o de teléfonos inteligentes ya no nos están bastando, pues ahora también estamos conectando nuestro cuerpo, nuestra casa o nuestro automóvil a distintos dispositivos, con el fin de hacernos la vida "más fácil". Ahora también monitorean nuestra salud.

2.2 "Acepto los términos y condiciones": la entrega de nuestros datos

Los periodistas Kashmir Hill y Surya Mattu (2018) se dieron a la tarea de investigar lo que los dispositivos saben sobre nosotros. Analizaron la cantidad de data que se genera en una casa, con el objetivo de identificar a los usuarios cada vez que se conectaban a un teléfono, a la televisión, a la radio o a su teléfono móvil. La casa tenía un sensor que medía los datos que salían para hacerlos llegar a sus centrales.

Algo tan sencillo y tan íntimo como un cepillo de dientes —electrónico, por supuesto— generaba datos cuando se prendía y apagaba. ¿De qué sirve saber esta información? Para las empresas, empleados y aseguradoras de Estados Unidos de América, uno de los datos más valiosos es el de los seguros dentales. Cuando un ciudadano acuda a una para que le den un seguro dental, buscará el nombre —si usa un cepillo electrónico—, verá sus hábitos de higiene y cepillado de dientes y, en función del análisis de la data, le cotizará un buen precio, si tiene hábitos saludables, o subirá el costo de la póliza si tiene hábitos poco saludables. Algo tan simple como el uso de un cepillo de dientes genera data valiosa. Por ello, esa data se vende y se compra.

No solo eso, ambos periodistas también investigaron de forma específica qué pasaba con los televisores inteligentes. Estos están conectados todo el tiempo, siempre están enviando datos, aunque pareciera que están apagados. Y, entonces, ¿qué transmiten? Cuando adquirimos un televisor digital y comenzamos la programación del equipo, solemos aceptar los términos y condiciones (sin leer los términos de referencia). En el contrato que firmamos sin leer se especifica que autorizamos a mandar no solo información detallada de lo que estamos viendo cuando encendemos el televisor, sino de todas las aplicaciones a las que nos conectemos.

Con nuestro correo electrónico nos conectamos a varias plataformas. Ahora imaginemos que nos conectamos a un televisor inteligente. ¡A cuántas aplicaciones nos conectamos! Netflix, Amazon Prime, Spotify, YouTube, entre muchas otras.

Incluso, cuando transmitimos la data desde el teléfono al televisor, ambos serán utilizados para perfilar consumo y

consumidores. Nuevamente, pareciera que los datos o rastros digitales no están ahí, pero muchas empresas de todo el mundo empiezan a utilizar la información para perfilar a los consumidores.

De la misma manera, los dispositivos Echo de Amazon y Google Home permiten interconectar cualquier aparato o dispositivo con bluetooth o wifi. Echo y Google Home tendrán toda la información de cada uno de los aspectos que conforman nuestra vida diaria.

Los usuarios consentimos la entrega de esa data por voluntad propia. Cada vez que aceptamos los términos de servicio y los términos de privacidad de las compañías, lo que estamos haciendo es dar permiso para que accedan a cierta información. Por ejemplo, si se instala la aplicación de Instagram, lo primero que solicita es que se le permita acceso al teléfono, a la ubicación, a las fotografías y a las llamadas. Cuando se le da acceso al teléfono en general, se le da acceso a la galería de imágenes y no solo a las que se publican, sino a todo lo que hay en la galería de imágenes del teléfono. Instagram tiene acceso a todos los videos y fotografías de nuestros *smartphones*.

Para que nos demos una idea de la complejidad de este tema, pensemos en que quizás una de las razones por las cuales no solemos leer esos contratos, términos y condiciones, es que para hacerlo necesitaríamos, en el caso de Instagram, nueve minutos y 42 segundos, en total 2451 palabras. Para leer los términos de servicio de Microsoft, se requiere una hora con once minutos, pues tiene 17 121 palabras, lo equivalente a la obra *Macbeth* de Shakespeare. Recordemos que Microsoft tiene acceso a todo lo que guardamos en nuestra computadora.

Otro ejemplo es la red social Facebook, cuyo aviso toma diecisiete minutos y doce segundos para leerlo, y tiene 4132 palabras. Lo que pasa con Facebook es que, a través de su plataforma, ingresamos a otras. Al "loguearnos" con Facebook, Google o Twitter, lo que atestiguamos es que la plataforma se vuelve un consumidor más grande, pues no solo tiene acceso a lo que voluntariamente introducimos en la plataforma, sino que tendrá acceso a capas y capas de información, es decir, data, data y más data.

Si el servicio que ofrecen estas plataformas no tiene costo, entonces, ¿quién es el producto? Nosotros somos el producto, nuestros datos son el producto. No es gratuito que estas plataformas nos permitan entrar a gastarnos miles de datos, millones de terabytes. No es gratuito; nuestros datos son el negocio. Ahora, empecemos a aprovecharlo también nosotros.

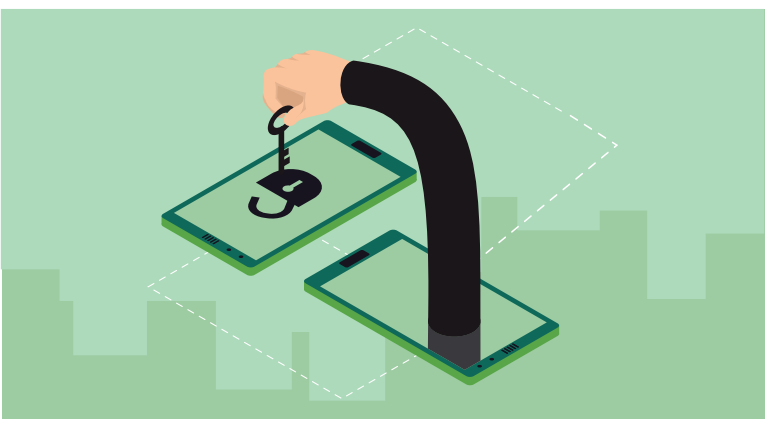
2.3 ¿Cómo se usan los datos?

Nuestra data sirve para que Netflix, por ejemplo, pueda presentarnos miles de posibilidades o accesos distintos, con el fin de que cuando entremos a la plataforma no veamos las mismas portadas. Netflix llega a medir información detallada a través de identificar cuándo pausamos, cuándo

dejamos una serie a la mitad, y va seleccionando y descartando información a través de la sistematización de esos comportamientos.

Imaginemos que nuestra actriz favorita es Umma Thurman. Si vimos *Kill Bill*, el algoritmo ofrecerá las series o películas que tengan a Umma Thurman en la portada. Eso nos llevará a ver *Pulp Fiction*, con ella en la portada principal. En cambio, si nuestro perfil se inclina por John Travolta, el algoritmo nos ofrecerá portadas que lo incluyan.

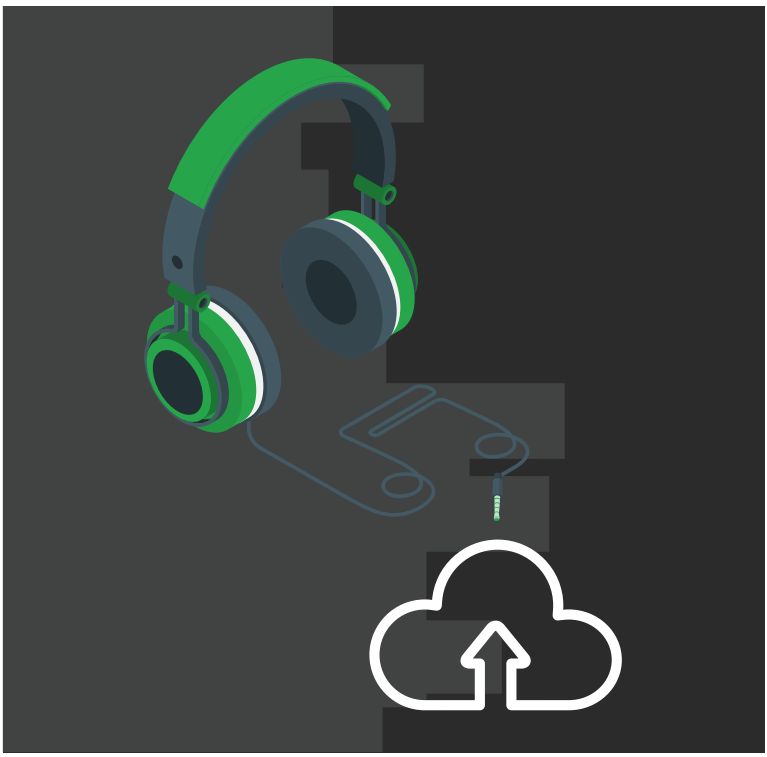
El algoritmo empalma nuestros gustos con el producto que nos están ofreciendo, y ello se basa en el comportamiento que mostramos al elegir determinada serie o película. De tal manera, cada vez que terminamos de ver algún contenido, la plataforma nos recomienda ver otra cosa basada en nuestro gusto anterior. Las elecciones del algoritmo se basan en categorías como personajes o tiempos de consumo. Si vemos una serie de diez capítulos y luego vemos tres series con ese mismo número de capítulos, lo que el algoritmo inferirá es que nos gusta ver series cortas, por lo cual nos ofertará series de duración semejante, con el fin de que los espectadores consumamos más, que estemos contentos y no cancelemos la suscripción. De esta manera funcionan los algoritmos.



Otro ejemplo es el uso de los algoritmos en Spotify. La aplicación de música tiene dos grandes criterios y fue la primera en usar algoritmos para cambiar o sugerir *playlists* en función de estados de ánimo. Si bien iTunes fue una de las primeras plataformas que revolucionó la música digitalizada o *streaming*, Spotify aplica con creatividad los algoritmos para conocer los gustos y preferencias de los usuarios y ofertar contenido en función de ello. Para tal efecto, produce listas de reproducción categorizadas de acuerdo con una amplia variedad de géneros, estilos, orígenes, usos e incluso tipos de escuchas. Por ejemplo, sus listas de reproducción "zodiacales"; en función de la fecha de cumpleaños del usuario, que este introdujo al registrarse, el algoritmo realiza una búsqueda entre todos los usuarios para construir una lista de reproducción en función de cada signo zodiacal. Con cada uno, lo que hizo la herramienta fue identificar la canción que se reproduce con mayor frecuencia en un rango de mes y la integra a una *playlist* que tiene pocas variaciones, lo cual significa que son similares en América Latina o en Estados Unidos de América.

Spotify introdujo un cambio muy relevante acerca de cómo entregar la música. Descubrió que las personas empezaron a armar listas de reproducción para una tarde de barbacoa o de carne asada o de comida o de café. Se dieron cuenta de que los usuarios empezaron a compartir música en función de sus momentos anímicos, por lo cual, hoy, pueden entrar y ver que la música más escuchada se elige en función de las *playlists* de estados de ánimo. Nuevamente, le cedemos al algoritmo la capacidad de seleccionar música por nosotros. En años anteriores, a muchos de nosotros nos tocó grabar cintas, editar y grabar nuestras propias listas, Seleccionábamos el orden de la música. Ahora elegimos "cena con amigos", "cena con pareja", y normalmente nos quedamos satisfechos, porque si no

fuera así, nos daríamos de baja de la aplicación. Así es como se usan los algoritmos en el área empresarial y comercial.



Ahora bien, ¿cómo se pueden usar todos estos datos en el ámbito político? Podemos usarlos para:

a) segmentar información;
b) visualizar;
c) comprender coyunturas, y
d) usar *big data* para manejo de *fake news* y *deepfakes*.

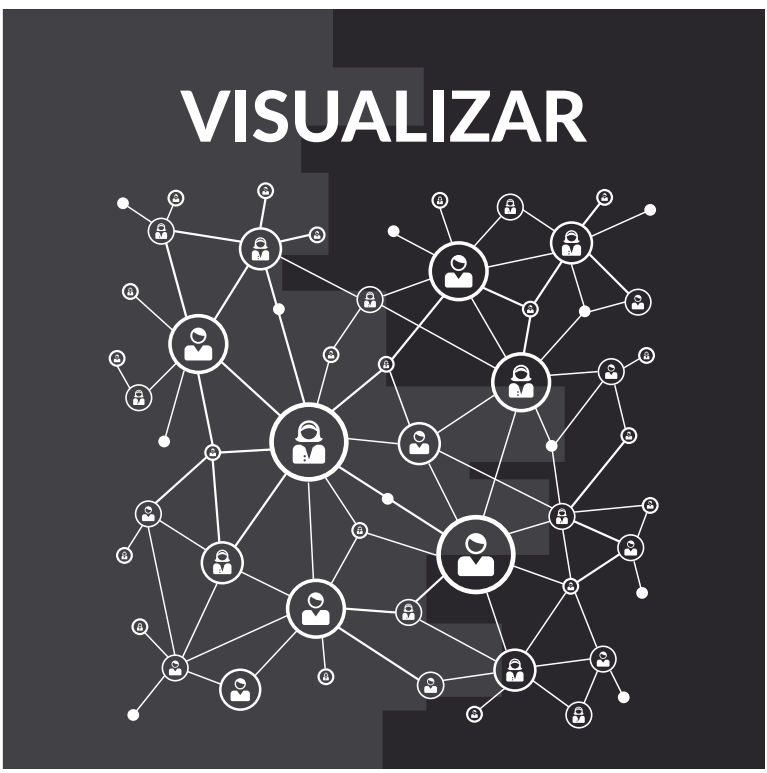
Segmentar información

Cada declaración que hacemos es de intención. Las respuestas a preguntas de un cuestionario, de una entrevista, de una encuesta de campo o de un *focus group* nos dejan un sesgo. La diferencia del *big data*, en relación con los demás métodos, es que cuando damos un clic, en teoría, nadie nos está mirando, de manera que cada vez que compartimos algo, incluso un gusto culposo musical o una serie o una fotografía, o hacemos *check in*, o posteamos en nuestro muro de Facebook "estoy viajando a", "estoy haciendo esto", "estoy haciendo aquello", estamos declarando nuestra intención. Y la estamos dejando para que los analistas puedan interpretar qué dicen nuestras declaraciones, cómo lo dicen, dónde lo dicen, cuándo lo dicen, y en qué contexto lo dicen.

Toda esa es información de valor, pues nos permite segmentar con base en declaraciones que normalmente son emocionales. Aquí no entra el tema de "me van a ver", sino que incluso se llega a ver una declaración del tipo "quiero que me vean", "quiero que sepan".

Visualizar

Los datos nos permiten visualizar información, actores y afinidades de medios de comunicación entre políticos y medios institucionales, gobierno y ciudadanos. Una visualización permite definir cómo cimbrar y en dónde cimbrar. Si quieres tener resonancia, tienes que pegar en el centro y hacer que todos los demás lo sientan. La única manera de visualizar esto es mapeando la manera en cómo interactúan los medios sociodigitales y los distintos elementos que se quiere analizar. Entonces, se puede inferir quiénes son los actores más relevantes para que se pueda saber dónde atacar y, con ello, generar resonancia en todas las áreas.



Dave Troy (2014) procesa datos para ver los lugares, no como vecindarios, sino como relaciones entre personas. Con su proyecto PeopleMaps, agrupa a personas a quienes siguen en Twitter y con quienes hablan por esta red. Luego, trabaja con colaboradores locales para analizar de qué hablan con más frecuencia esas personas. ¿El resultado? Fascinantes mapas no geográficos de los intereses y comunidades de una región.

Recientemente se realizó un ejercicio de cómo se ven las ciudades al mapear las diferentes tribus y su localización georreferenciada en San Francisco, California. Se pueden mapear grupos a partir de datos específicos o de comportamientos, temáticas, gustos o preferencias, con la meta de proyectarlo de manera visual.

Comprender coyunturas

El análisis de la data permite entender contextos y momentos muy concretos. Por ejemplo, en el caso del coronavirus, las herramientas de escucha social como Brandwatch permiten medir lo que está pasando exactamente en todos los países de Europa y América Latina. En este *insight*, el miedo y la tristeza son las emociones que con más frecuencia expresan los usuarios, sumándose a ellas la incertidumbre acerca del tiempo que durará la pandemia y los confinamientos.

Así como podemos ver esta data, sería posible identificar cómo los políticos estarían contribuyendo a que el miedo y la tristeza de la población disminuyeran si sus acciones y discursos otorgaran certidumbre. Pero eso es lo que menos hacen, por lo menos en el caso de México, de modo que reina una polarización en torno a lo que se debe de hacer para enfrentar la pandemia. El resto de América Latina más o menos se encuentra en la misma situación. Ni hablar de Brasil o de Estados Unidos de América. Al parecer, estamos más preocupados por el tema económico y no atendemos a los ciudadanos quienes expresan miedo y tristeza. Ello evidencia

cómo la comunicación política se está encontrando con un muro o, por otro lado, está ayudando a amplificar la brecha entre los políticos y los ciudadanos. El comunicador político no está diciendo algo relevante para el ciudadano. Sólo está diciendo lo relevante para sí mismo.

Los políticos y la política tienen su fundamento en el acto de escuchar, pero hoy en día los datos están disponibles para que los vean y les ayuden a planear e implementar mejores políticas públicas. Sin embargo, políticos de todas las corrientes y partidos están comportándose de manera análoga, mientras que la ciudadanía ya está hiperconectada. Los políticos están desconectados de esta realidad digital, se han quedado con la idea de que Twitter lee el humor social o que lee a la sociedad política.

Aunque la opinión pública no necesariamente está ahí, la opinión política y del círculo rojo sí, y el problema es que los políticos se quedan en las burbujas de comunicación de Twitter. Y entonces, como no salen de esa burbuja, creen que todo se está conversando ahí. Twitter no necesariamente es representante de la sociedad digital. Eso se encuentra en otras plataformas: Facebook, Instagram, Snapchat, que son más nobles y se prestan para la conversación. No obstante, la red social Twitter se ha caracterizado por mostrar una polarización mayúscula. Lo cierto es que todas las burbujas pueden reventar.

La opinión de la población acerca de una coyuntura específica también puede ser analizada a través del *big data*. Por ejemplo, la plataforma Brandwatch tiene una herramienta en la cual se puede lanzar una pregunta a través de *adds* publicitarias, a audiencias segmentadas, para que, en tiempo real, les aparezca la pregunta y los usuarios responden; ya no en horas, sino en segundos o en minutos. Así es como se puede recopilar información de la sociedad con respecto a temas específicos. Esta plataforma sirve para llevar a cabo estudios

nacionales, pues se trata de sondeos digitales que utilizan inteligencia artificial.

Fake news y *deepfake*

Se trata del uso de los algoritmos y del *big data* para copiar, imitar, simular y reproducir personajes, muchas de las veces, con el objetivo de generar campañas negativas. Se toman fotos de medios digitales, generando un fotorama que se une a un sampleo de la voz del personaje. Así, la inteligencia artificial produce expresiones y es tan potente que transforma a un personaje en otro y hasta puede hablar por él. En breve, los datos públicos se utilizan para manchar la imagen de otros.

Gutiérrez Rubí, experto en tecnopolítica, afirma que "un buen uso de los datos puede contribuir a la mejora de la equidad, del bienestar, a la igualdad, al interés general [...] o a salvar el planeta, al poder medir el impacto de nuestra huella". Así, las ventajas de usar estas herramientas de datos son: *a)* Obtener información en un par de horas, sin trabajo de campo, sin sesiones de *focus groups*, sin tiempo de espera; *b)* Extraer datos relevantes y a nuestra disposición, pues nos permite identificarnos con el otro, ponernos en los zapatos del otro.

2.4 Procesar y utilizar el *big data: data mining*

El *big data* es la recopilación de grandes volúmenes de datos, desde diversas fuentes hasta su posterior procesamiento, con el fin de recogerlos, almacenarlos, buscarlos, compartirlos, analizarlos, visualizarlos, procesarlos y entenderlos.

Uno de los problemas que suelen presentarse en campañas políticas es la ineficiente estructuración, clasificación y sistematización de los datos. Por lo regular se organizan en hojas de cálculo de Excel, en otras plataformas, en otros sistemas; los datos están dispersos y desestructurados. Ello

es así porque, cuando se les quiere integrar, no se pueden vincular entre sí, pues tienen formatos distintos. El *big data*, en cambio, es altamente eficiente en su procesamiento. El *big data* solamente se refiere a la recopilación en tanto que el proceso que realizan las plataformas es reunirlos en un solo servidor.

El *data mining* es la capacidad de comprender la información. Cuando las imágenes o datos recopilados entran en proceso, se elaboran los *insights* que nos permitirán entender los sueños, los anhelos y los deseos más profundos de un cuerpo social. Por ejemplo, si quisiéramos saber el sentimiento o humor social de una ciudad, tendríamos que cuestionar a todos los habitantes. Sin embargo, lo que hacen las herramientas de escucha social, con procesos de *big data*, es leer la información y convertirla, a través del *data mining*, en *insights* con valor y con sentido.

La aplicación del *big data* en política es muy relevante porque la política basada en datos es conversación e interacción permanente. Los analistas deben ser capaces de tomar esos datos y darles valor. El uso del *big data* nos ayuda a tener potencia predictiva.

Si en una campaña política estamos comunicando y le pedimos a la gente que nos deje su correo electrónico, pretendemos que esa información vaya a una base de datos, para ser clasificado de acuerdo con cierto criterio. Esa base de datos nos permitirá enviar información personalizada.

Si en una campaña política —ya sea como consultores o como parte de un gobierno— logramos atrapar y recopilar los datos, procesarlos y darles, un valor, ello es más que suficiente para articular e interpretar información, darle sentido y trascendencia de acuerdo con la estrategia. No solo eso, cuando se aplica el proceso del *data mining*, lo anterior nos permitirá hacer predicciones sobre el comportamiento, gustos o antipatías de la audiencia.

En una campaña, el *big data* sirve para hacer *microtargeting* electoral. Como se mencionó algunas líneas arriba, no se puede hacer campañas para todos, ni campañas generalizadas. Es necesario planear y ejecutar campañas para cada uno de los segmentos de la audiencia y micropúblicos. Con mensajes personalizados y adecuados a cada una de estas mentalidades. Cuando se hace un análisis específico, los resultados que arroja nos dicen quién es la persona que tenemos enfrente, y esa persona no es solamente una edad o los valores sociodemográficos, sino también los valores psicográficos que lo caracterizan. Y a todo ello se llega a partir de las herramientas de recopilación y procesamiento de datos.

Esto nos permite hacer comunicación política muy dirigida y, básicamente, tener la información de los electores para compartirlos con el equipo de campaña. Cada área puede hacer su trabajo, sumando esta información adicional que antes no se podía obtener de manera tan sistematizada.

El *big data* se alimenta de tres fuentes que son útiles para la política:

1. Datos sociológicos, económicos, culturales, electorales, entre otros.
2. Con quién se relaciona y qué le interesa a la persona (redes sociales junto con el comportamiento del entorno digital).
3. Rastro digital y consumo, es decir, la información pública: qué compra la persona, en dónde está suscrita. Ello potencia el *microtargeting*.

Mientras más fuentes de información podamos incorporar, más relevante será la comunicación. Cada vez que el usuario tiene una interacción, completa un perfil, interactúa con un contenido, marca eventos vitales (como compromisos, cambios de trabajo, conclusión de estudios), declara interés

en algunos eventos, hace *check in* en plataformas, escribe comentarios o sube un *post*, obtenemos valor de esas interacciones. No solo se trata del análisis de los datos, sino de observar cómo toda esta información genera patrones de comportamiento, y el algoritmo nos ayuda a encontrar más personas semejantes.

Cada interacción crea una realidad que la persona vive y todos y cada uno de nosotros lo hacemos: vamos dejando nuestro rastro digital. Una manera muy sencilla de hacer un análisis de audiencia es ir a las métricas de Facebook y ver quiénes son estos usuarios. Realizar este análisis solamente arroja un beneficio: saber acerca de nuestros seguidores y lo que genera la huella digital.

Hoy, la mayoría de las personas que buscan hacer *marketing* digital se enfocan en esta huella. Se concentran en generar posteos, anuncios y dinámicas que atraigan seguidores y generen el famoso *inbound politing*. Eso vale, sin duda. Sin embargo, su limitante es que no podemos saber sobre las personas que hablan de nosotros, pero no nos siguen; aquellos que están afuera de nuestra burbuja. A pesar de que revisemos las métricas de Facebook, de Twitter y de otras redes, son silos, no es el universo completo.

La inteligencia social digital y el análisis de audiencias a través de *big data* no consiste en analizar estadísticas simple y sencillamente a partir de las plataformas. Esos datos solo van a hablar de las personas que ya nos siguen, es decir, solo lo que tenemos en nuestra huella digital. Esa que nosotros mismos producimos y a quienes seguramente ya conocemos, con quienes ya tenemos contacto y a quienes no necesitamos convencer. Lo que no se puede saber —porque no hay forma de que las estadísticas tradicionales lo muestren— es lo que se denomina *sombra digital*. Esa sombra digital es todo aquello que se habla de nosotros, de nuestra marca, de nuestro

tema. Y que no aparece en nuestra red social, podríamos decir, lo que se dice a nuestras espaldas.

Las interacciones para el uso de audiencias implican que no necesariamente vayamos a las estadísticas. Las estadísticas son datos aislados que no se comunican entre sí, porque están elaborados en silos y solamente hablan de un perfil. Las estadísticas nos ayudan a obtener un perfil y el de nuestra audiencia, de nuestros fans que ya están con nosotros, por lo cual difícilmente se encontrarán detractores en esa audiencia.

De lo que hablamos es de obtener la información de las conversaciones que hay en la vida digital del mundo. Esta información nos permite identificar de qué están diciendo en una zona geográfica determinada; quién habla de esos temas, qué y cómo; con qué humor lo dicen; en qué contexto se dice. Inclusive, se puede monitorizar cuándo y hasta cuánto se habla de ello. Todo esto gracias al *big data*.

Esta visión es el nuevo paradigma y está al alcance de todos. Nos presenta el reto técnico y estratégico de articular la sombra y la huella digital; entender lo que tenemos en frente, a la vista de todos, y mirar lo que está atrás y otros no ven. De ver cómo nos ven.

Las herramientas identifican cada texto y un robot comprende lo que escribes en un periodo de hasta tres años atrás y puede identificar cómo te has comportado, detectando patrones y estableciendo correlaciones. Este tipo de identificación textual también se hace con fotografías y, hoy en día, se hace hasta con video. Las herramientas de inteligencia artificial analizan el tipo de video que se sube, analizan los metadatos y se indica qué tipo de consumidor eres.

Las herramientas tienen que ser calibradas con esta información objetivo. Ello significa indicarles qué queremos que busquen: la meta, sigue siendo humano. Las herramientas buscan, detectan y captan la información, en tanto que el *data mining* hace la interpretación, a través de lo que se llama

crawlers, que son los motores semánticos que procesan y entregan la información en función de la calibración realizada. Así, se nos muestra cuántas conversaciones hay, cuáles son los temas o tópicos de los que habla la audiencia específicamente y qué valores tiene.

De esta manera, la herramienta convierte valores desestructurados de una audiencia general, de huella y sombra digital, en grupos. En los cuales ya se tienen *insights*. Esto significa que se tiene la capacidad de determinar qué sentimientos tienen, se pueden hacer predicciones o identificar a los *influencers*, entre otros aspectos. Integrar y analizar con este perfil sociodemográfico y psicográfico del electorado nos permite crear con mayor precisión los mensajes que generarán empatía con mayores posibilidades, para conectar con nuestras audiencias.

Esta era nos invita a reflexionar acerca del cambio de paradigma de las campañas políticas. Estas no pueden mandar un mensaje a todos, porque, de ser así, le producirá un corto circuito a cualquiera de las tribus que puedan estar con nosotros y no les podemos estar hablando de algo que no es relevante para ellas. Todos hemos participado en campañas en las que dicen: "yo hice microsegmentación" y, al final, la microsegmentación acaba siendo un problema, porque la campaña es global.

Obtener valor de cada interacción
Insights
Sentimiento
Simulación
Predicción
Influyente
Comunidades

3 *Big Five* e inteligencia artificial

¿Qué podemos hacer para pasar de solamente tener información sociodemográfica, a tener la información que representa "la joya de la corona", es decir, la información psicográfica? El reto está en buscar, por medio de palabras clave, todas las correlaciones que puedan llevarnos a la información que buscamos. Por ejemplo, en el caso de un candidato en México, buscaremos si tiene motes, apodos, si ha participado en otros cargos de otras administraciones. Hay que escribir palabras de relevancia.

La mejor analogía es imaginarse como investigadores privados, con el fin de reflexionar y pensar cómo se buscaría al candidato y qué palabras claves son las que nos serían más útiles. Entonces, la herramienta buscará la información. Debemos tener cuidado y subrayar que la herramienta no identificará el perfil, pues la búsqueda no va dirigida a recopilar la huella digital, sino que va a todas las fuentes, como periódicos, medios, redes. Todo lo que está relacionado con la hiperconectividad en la cual nos encontramos. La herramienta atrapa toda esa información que luego produce estos *insights* de los que hemos venido conversando.

En 2012, los investigadores Michal Kosinski, David Stillwell y Graepel (2013) aplicaron una prueba denominada "My personality". De acuerdo con las respuestas, hacía una clasificación de cinco rasgos de comportamiento, se categorizarán en cinco personalidades, conocidas como las Big Five:

1. La apertura: Nuevas experiencias + impulsos + capacidad de planear
2. La extroversión: Actitudes sociales + ¿eres alegre?
3. La conciencia: Perfeccionismo + actividad + estilo de vida
4. La agradabilidad: Cooperativismo + ¿eres considerado?
5. La neurosis: Preocupaciones + miedos + autoconfianza + obsesiones + irritabilidad

Las *Big Five* identifican qué tan abierta, introvertida, consciente, perfeccionista, agradable, considerada o neurótica es la persona sujeta de la prueba. Cuando este ejercicio se realizó, Kosinski y Stillwell pensaron que habría poca gente que les daría sus datos. Sin embargo, muchos usuarios llegaron y, al hacer la prueba, cedieron la información necesaria. Los investigadores de la Universidad de Cambridge se preguntaron qué pasaría si cruzaran esa información con los "me gusta" de Facebook. Dado que ya sabían qué comportamiento tenía cada persona, descubrieron a través de correlaciones, que 68 "me gusta" en Facebook podían predecir el color de piel de una persona con una precisión de 95%. Al contar con esta información de personalidad, más los "me gusta" se podía identificar su orientación sexual en 88%. Y no solo eso. También se podía identificar su orientación política con un grado de certeza de 85%. Además, se podía deducir su grado de inteligencia, si consumía o no drogas o alcohol, su creencia religiosa, e incluso si sus padres estaban separados.

Cuando empezaron a ver cómo era una persona y a correlacionarla con todos los *likes* que daba, asignaron puntajes y compararon los datos con toda la información disponible: edad, género, y todos los cruces. Así, fueron teniendo hallazgos como que a todas las personas abiertas les gustaba más Salvador Dalí, la meditación y las charlas TED, y empezaron a agruparlos en tribus. Los hombres que daban más *likes* a Mac tendían a ser homosexuales, en tanto que uno de los factores que indicaba la heterosexualidad era dar *like* a Wu Tang Clan, una banda de hip hop originaria de Nueva York.

En su estudio, Kosinski y Stillwell (2013) descubrieron que, al alinear un perfil con todos los rasgos de su huella digital, se pueden identificar patrones y empezar a predecir qué le gustaba a cada persona. Por ejemplo, los seguidores de Lady Gaga eran más extrovertidos que los que seguían a filósofos. A partir de esas hipótesis empezaron a comparar datos y a hacer correlaciones que empezaron a definir en tribus y audiencias.

Los dos investigadores infirieron que con diez *likes*, la plataforma puede conocerte mejor que un compañero de trabajo o un colega. Con 150 *likes*, la herramienta puede conocerte mucho mejor que la persona con la que cohabitas o incluso que tu familia. Con trecientos *likes*, te puede conocer mejor que tu pareja.

El promedio de *likes* mensuales es fue de 227. Pensemos cuántos *likes* damos en promedio y a qué damos clic. Podríamos llevarnos una sorpresa.

La plataforma nos puede conocer mejor que nosotros mismos por el sesgo que mencioné anteriormente. Tenemos miedo a decir quiénes somos, pero estamos compartiendo mucha información y mucha data, y nunca en una sola red. Es importante que sepamos cómo funciona esto, porque muchas veces nos dicen "es que los que hacen esto nos ven la cara". Las plataformas nos ven la cara porque saben qué

hacemos, cómo lo hacemos, a qué hora lo hacemos. La data está aquí.

Imaginemos que no solo tenemos la cantidad de *likes*, sino que con información adicional podemos perfilar a una persona con todo aquello a lo que le da "me gusta", todo lo que comparte y todas las huellas digitales que está generando.

Como ya se dijo, las plataformas tienen los datos. Es importante saber que siete de cada diez personas tienen su perfil abierto. Entonces, ya tenemos la configuración no solo de una persona, sino de distintas personas a las que estamos buscando. Kosinski y Stillwell desarrollaron un algoritmo que elabora hipótesis a partir de una fotografía y de la revisión de muchos datos: cómo una persona abierta muestra el cuello, cómo es su nariz, si cuida mucho su cabello, etcétera. Estas hipótesis "disparan", y lo que van descubriendo es que cada foto, cada video, determina un grado de comportamiento específico. Todo se apoya en los rastros digitales.

Lo que hace la herramienta es identificar criterios ya establecidos, científicamente válidos, de cómo son las personas. Y a partir de eso hace un análisis de *big data* a través de un proceso de *data mining* y de inteligencia artificial. Le colocan parámetros a las herramientas que dicen cómo es una persona extrovertida y lo traducen en códigos binarios. Si alguien quisiera saber cómo es otra persona a nivel de comportamiento, lo que tendría que hacer es completar la encuesta de 179 ítems de *Big Five*. Pero, en este caso, lo que se hace es buscar la audiencia que se desea, colocar los parámetros, y la herramienta ya sabe cómo llegar a cada uno de ellos.

Si no me creen, entren a *applymagicsauce.com*, donde cualquiera puede descargar sus datos de Facebook, Twitter o Instagram, y el sitio les irá guiando en el proceso. Al hacerlo, la plataforma arroja los rasgos de comportamiento, con 95% de certeza. Estamos hablando de una investigación completa.

Un análisis de personalidad con base en inteligencia artificial, que ve rasgos, posturas fotográficas, analiza videos y texto, innegablemente logra identificar de manera muy asertiva el comportamiento. Ello no significa que si el resultado es que eres extrovertido siempre te comportes así, simplemente es tu tendencia.

La herramienta dará el valor demográfico que se usa en todas las plataformas a través del *data mining* tradicional, pero, hoy en día, la "joya de la corona" es tener el valor psicográfico. Todos los profesionales que nos dedicamos a la tecnopolítica nos preguntamos cómo podríamos obtenerlo de forma más rápida.

Los datos están resguardados por la Universidad de Cambridge, pero el objetivo del ejercicio es demostrar cómo, a través de los rastros digitales, estamos dejando nuestros sueños, anhelos y deseos más profundos. No es lo que decimos, es lo que hacemos, y muchas de las veces no estamos

conscientes de lo que hacemos. Eso es el cómo de la magia de la información y de la inteligencia artificial.

Cuando ya tenemos el resultado de este análisis, la interpretación nos dice que los de mentalidad abierta utilizan cierto tipo de lenguaje, expresiones y palabras. En tanto que los de mentalidad cerrada usan otro tipo de lenguaje, expresiones y palabras.

Las campañas deben contar con científicos de datos al interior de sus cuartos de estrategia. Se debe contar con equipos especializados en el análisis de la data, equipos que conozcan de manejo de información; sociólogos y psicólogos, porque son quienes verdaderamente pueden sacarle jugo a esta información y entonces sí, elaborar estrategias de acuerdo con la nueva humanidad digitalizada.

IBM Watson es uno de los motores para el procesamiento del lenguaje natural y el aprendizaje automático. Para utilizarlo, es necesario trabajar de la mano de profesionistas en la ciencias de datos, que nos ayuden a integrar la información con la perspectiva deseada. El procedimiento básico consiste en introducir la información, y la plataforma va dando los *insights* ya armados.

Existen algunas otras herramientas que vamos a ver en las campañas políticas que se avecinan. Estas podrán determinar de forma muy específica los cinco rasgos del comportamiento. Cuando el usuario introduce esta parte del análisis de inteligencia artificial se obtienen los valores en los cuales se sustentan, las necesidades que tiene y luego las *Big Five*.

En este punto, se realiza la interpretación. Es ahí donde se necesitan especialistas en neuropsicología o sociología, pues ellos lograrán identificar y dar *insights* a los equipos que construyen la parte publicitaria y la parte de estrategia, con el fin de llegar a convencer a distintas tribus, en distintos momentos. Siguiendo con el ejemplo de un candidato en México, con el análisis, obtendremos de cinco a diez audiencias distintas.

Es decir, cincuenta personalidades, cincuenta tribus que se van agrupando, pero ya tenemos diez formas de agrupación que son más fáciles de estudiar para generar productos de comunicación. El reto es generar seis, siete o diez campañas que irán cambiando por zonas geográficas.

¿Esto es nuevo? No. ¿Ya se probó alguna vez? Sí, en Argentina. Mauricio Macri ganó así la Presidencia de la República Argentina. En Estados Unidos de América está el caso de Donald Trump, en el que Cambridge Analytica hizo el análisis de una forma poco ética, pues no pidió permiso para el uso de los datos.

Hoy se puede hacer sin restricciones, porque sigue habiendo acceso a los datos de las plataformas, y podemos construir perfilamientos muy específicos. Lo más ético es respetar los perfiles de los usuarios.

3.1 Inteligencia social digital aplicada en la política

Un principio básico de la política es escuchar a la ciudadanía. Todos los políticos de todos los partidos y corrientes ideológicas dicen que lo hacen. Sin embargo, nos encontramos en una época donde la tecnología —como lo hemos comentado a lo largo de este cuadernillo— nos permite acceder al humor social de la audiencia digital en su totalidad, las peticiones, intereses, preocupaciones e inclusive tendencias de conductas están al alcance de su mano y las de sus equipos.

La inteligencia social digital es una técnica que permite identificar los temas (percepción, preocupaciones, sueños, anhelos y miedos) de la sociedad digital y obtener *insights* para entender el humor social respecto del gobierno, de los políticos, o de cualquier tema de interés, ya sea estructural o coyuntural. Se trata, pues, de entender a la sociedad hiperconectada de una manera profunda, en todo lugar y en todo momento.

Para que esto suceda es importante determinar el universo de personas, temas, instituciones o delimitar la zona geográfica por analizar. Este proceso se realiza a través del desarrollo de *algoritmos booleanos*, mediante los cuales limitamos nuestro caso de estudio y creamos todas las excepciones que se requieran. Esto puede realizarse con cualquier herramienta de escucha social disponible en el mercado.

Una vez que la herramienta nos arroja la información, comienza el verdadero reto, que es analizar la información más allá del método tradicional de solo ver lo cuantitativo. Además, hay que entender de manera detallada y profunda lo que le preocupa a la sociedad para poder atenderlo, ya sea con el desarrollo de la gestión pública, la realización de políticas públicas, o bien, generando campañas que informen acerca de las acciones realizadas. Con ello, se atiende al principio de escuchar, pero al mismo tiempo, se atienden las demandas de la sociedad. En una campaña política, se pueden ofertar propuestas que conecten con las necesidades de los electores.

El reto para identificar los temas es atrapar la mayor cantidad de rastros digitales del objetivo que estemos canalizando. Para tal fin, es importante determinar el universo que buscamos y, a partir de ello, leer y analizar una muestra representativa de las conversaciones, siempre quitando los *bots*, la inducción, los medios de comunicación, las cuentas identificadas con políticos. Con el objetivo de evitar sesgos y obtener información relevante de los ciudadanos reales que están conectados. A partir de ese proceso, es relevante determinar cuáles son los temas, necesidades y demandas de estos ciudadanos, con la finalidad de definir qué acciones se deben llevar a cabo para atenderlas.

3.2 ¿Cómo definir un proyecto de este tipo?

Este cuadernillo tiene una finalidad didáctica, y con el propósito de compartir cuál sería, desde mi perspectiva, el ejercicio inicial para poder aprovechar estas herramientas y técnicas a las que nos estamos refiriendo, planteo cuatro puntos.

Primero. Definir qué se quiere analizar y en qué audiencia hay interés. ¿Qué se quiere investigar? Entre más específico sea el objeto de estudio, mejores, más valiosos y útiles serán los datos. Podemos, si así lo requiere el equipo de estrategia, hacer el estudio para analizar cómo ven a la competencia. Si se quisiera, por ejemplo, buscar el perfil de Andrés Manuel López Obrador, se tendría que programar la herramienta con todos los motes o palabras con los cuales se puedan referir a él, desde las campañas anteriores. No basta con las referencias actuales; presidente, candidato, todas las palabras que puedan correlacionarse con el tema principal y que arrojen una mayor cantidad de información.

Segundo. Una vez que se tiene la información, debemos rescatar metodológicamente la que arroja valor a la investigación. En este proceso, convertimos el *big data* en *small data*.

Tercero. Escoger los canales en los que se quiere hacer. ¿A quién se quiere escuchar? Si ya se tiene en mente la audiencia a la cual se quiere llegar, entonces la herramienta arroja en dónde están los datos que pueden ayudar a llegar a esa audiencia.

Cuarto. Elegir la herramienta que se quiere utilizar. Algunas herramientas que ya tenemos a las mano y nos pueden ayudar a sumarnos a la era del *big data* pueden ser:

- *Google Alerts*. Permite escoger con exactitud temas que generen alertas. De esta manera, avisan cuando se genera una conversación en torno al gobierno que

se desee analizar, los servicios que ofrece, o bien, un personaje.

- *Google Trends*. Herramienta gratuita de Google Labs que muestra los términos de búsqueda más populares del pasado reciente.
- *Social Mention*. Es gratuita, hasta una cantidad de datos. Analiza los datos con mayor profundidad, basándose en cuatro categorías: fuerza, sentimiento, pasión y alcance.
- *Talkwalker*. Disponible para más de diez redes sociales. Permite filtrar y obtener datos de hasta hace dos años.
- *Mention*. Genera reportes que permiten obtener información detallada de cada mención recibida. Es de pago, con periodo de prueba gratis.
- *Brandwatch*. Identifica tendencias, crisis potenciales y encuentra tendencias que afectan a determinada marca, producto o mercado.

3.3 Monitorización de *hashtags*

Instagram y Twitter permiten monitorizar *hashtags* y contenido relacionado con un tema determinado. Saber utilizar *hashtags* adecuados es fundamental, porque se podrá conocer quiénes hablan y qué se dice de la marca, competencia o sector correspondiente. Cuando se monitorizan *hashtags*, se considera:

- Número de tuits o publicaciones generadas en torno al tema monitorizado
- Alcance de las publicaciones
- Conversación social generada en torno al *hashtag*
- Quiénes son los usuarios más influyentes dentro de la conversación social.

- La procedencia de publicaciones y punto de partida de los tuits

3.4 Mejores prácticas

- Encuentre menciones sociales sobre su proyecto y vea si se le habla y cuáles son las asociaciones comunes.
- Mientras configura su proyecto, intente reunir todas las variaciones del nombre de su proyecto y póngalo en su consulta.
- Compare los números con el análisis de sentimientos. Esto puede mostrar el estallido de una crisis potencial (si el número de menciones es inusualmente alto y la mayoría de ellas son negativas).
- Piense en el monitoreo de los medios sociodigitales como un vehículo en su camino hacia las mejores relaciones con los ciudadanos y la gestión de la reputación. Responda a las menciones de su proyecto en las redes sociales (especialmente si son negativas).
- Saque conclusiones generales de la analítica social. Hay un momento para enfocarse en cada mención y un momento para tomar una vista panorámica de su audiencia. Haga coincidir su proyecto con las necesidades de los ciudadanos y personalice la comunicación.
- Confronte las tendencias con las preferencias de su audiencia. Describa las fuentes más populares, no para los internautas de todo el mundo, sino específicamente para aquellos interesados en su campo.
- Compare su proyecto con sus competidores en análisis precisos, como género, geolocalización, histogramas, horarios y fuentes para encontrar el apalancamiento o detectar el nicho.

- Verifique sus ideas y decisiones con su audiencia. Escuche la voz del ciudadano para demostrar que realmente le importa.

3.5 Mejores prácticas del *social listening*

- Tenga en cuenta el alcance y relevancia del *influencer*, verifique la participación de su audiencia; un experto carismático acorde con su estrategia de marca.
- Escuche las preguntas o problemas de los ciudadanos y atiéndalos para ser confiable. Use las ideas para crear su estrategia de contenido.
- Aproveche la comunicación uno a uno con los ciudadanos que descubra. No subestime a ninguno. Haga el mejor esfuerzo para contactarlos.

3.6 De la inteligencia social digital al uso de la #IA para el perfilamiento de audiencias

Por lo general, los esfuerzos por mejorar la comunicación digital suelen quedarse en escucha social y —en el mejor de los casos— en *inteligencia*. Sin embargo, en la actualidad, ante una sociedad hiperconectada, no basta con tener claras las ideas para conectar con la audiencia. Es necesario descubrir quién está del otro lado de la pantalla, a nivel sociológico y psicográfico, para poder dialogar de forma empática con "cada uno de ellos".

Para lograr este propósito, es necesario usar el modelo de análisis de audiencias a través de inteligencia artificial. A partir de la identificación, clasificación y carga de usuarios en la herramienta de IBM Watson, permite al robot revisar los perfiles de estas audiencias que han sido previamente seleccionadas.

Una vez que el robot revisa el perfil (textos, fotografías, imágenes y videos), se identifican variables sociodemográficas y de comportamiento que permiten identificar conglomerados con rasgos similares, como vimos en el proceso de Kosinski. Por medio de esta metodología, se pueden perfilar audiencias específicas basadas en 179 variables de las *Big Five*.

Una vez que se cuenta con este nivel de perfilamiento, es posible generar la comunicación que permita construir el mensaje adecuado a la audiencia ideal, en el momento correcto, con el sentido idóneo, para la personalidad exacta y con el punto de interés preciso para convencer.

Solo así se puede generar empatía y relevancia con la audiencia, entendiendo cuáles son sus anhelos. Y al saber cómo son sus personalidades, se pueden desarrollar microcampañas que impacten a cada individuo que conforma el conglomerado.

Por ejemplo, imaginemos que es un diputado y que va a proponer una iniciativa de ley para que las armas sean adquiridas como en los Estados Unidos de América. El análisis de inteligencia digital y audiencias nos arroja tres conglomerados:

- Primer perfil: aventurero. De mentalidad abierta, un poco neurótico y ama la variedad, impulsivo, nervioso.
- Segundo perfil: protector, agradable, extrovertido, piensa en los demás.
- Tercer perfil: concienzudo, abiero, es una líder natural, planificador de pensamiento racional.

Estas tres audiencias son muy distintas y quererles ofrecer un mismo mensaje sería erróneo, dado que cada una tiene hábitos digitales y personalidades distintas. Sin embargo, a través de esta metodología podemos ofrecer un contenido que le haga sentido a cada uno de ellos.

Por ejemplo, al perfil *aventurero* se le enviaría algo así:

Al perfil *protector*:

Al perfil *concienzudo*:

Cada mensaje está pensado con el fin de que tenga sentido para cada una de las audiencias con base en la personalidad. Esto permite que una campaña convenza a tres conglomerados distintos de apoyar el objetivo propuesto.

Como sostiene López-Zafra, especialista en análisis de datos y *big data*: "En un mundo donde todos los días, a todas horas, en todos los lugares, los votantes dejan rastro digital, triunfará quien sea capaz de identificar el comportamiento de cada segmento de votantes".

Conclusiones

Después de la pandemia mundial por Covid-19, dimos un salto cuántico como sociedad. Actualmente, atravesamos una transformación acelerada que ha hecho que estemos inmersos en la atmósfera digital. Las actividades en persona se detuvieron y durante meses han tenido que realizarse en una plataforma virtual. Bien por la humanidad, la cual, gracias a la tecnología, ha podido malabarear para sostener su organización actual. Ahora somos una sociedad hiperconectada, con los riesgos y ventajas que ello implica. En este cuadernillo comparto las ventajas que en los próximos procesos electorales podremos tener a la mano.

El uso del *big data* y la inteligencia artificial nos da elementos para poder alcanzar uno de los objetivos más importantes de la actividad política: saber escuchar a la ciudadanía. Estas herramientas dejaron de ser ideas para una película de ciencia ficción o de estar solo al alcance del grupo en el poder. Hoy por hoy, estas herramientas están al alcance de todos, y es la creatividad, el profesionalismo y la convicción disciplinada los que lograrán la diferencia.

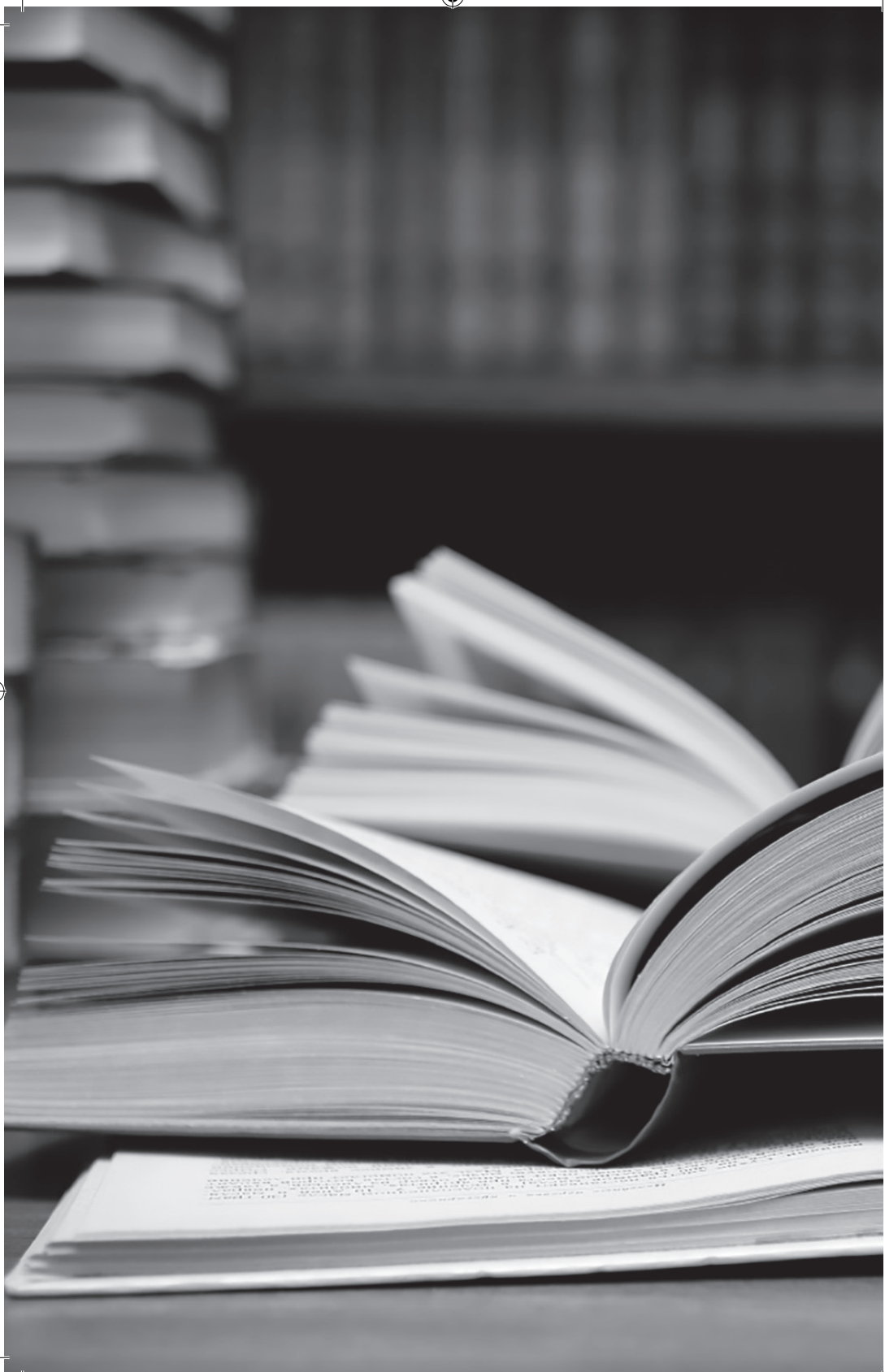
En el ámbito técnico de la consultoría en comunicación política tenemos una tarea que hemos discutido desde hace algún tiempo. Las campañas generales han dejado de funcionar, necesitamos conquistar no solo a "un electorado", sino

a varias audiencias. Cada una con intereses, preferencias y preocupaciones distintas. ¿Cómo presentar a un candidato o candidata que haga clic con cada una de estas audiencias? Ese es el reto. En Tecnopolítica, estamos en ello.

Una vez que logramos desarrollar las estrategias para llegar a cada audiencia, el reto operativo está en implementarlas, monitorizarlas y calibrarlas. Las audiencias son entes vivos que se transforman hora a hora. Los políticos requieren de equipos y *staff* a la orden, no de ellos. Los actores políticos que desarrollen la capacidad de reaccionar y resolver la demanda y preocupación ciudadana con empatía y precisión, avanzarán a los siguientes niveles. Los que se queden en el paradigma tradicional de solo leer o escuchar a su burbuja, y a su círculo rojo o de poder, vivirán la tragedia de la obsolescencia programada.

Alberto Rentería

Es *growth hacker social*, especialista en tecnopolítica, *coolhunting* y nuevos medios comunicación. Ha colaborado en proyectos de transformación digital para empresas, organizaciones sociales y partidos políticos. Con 15 años de experiencia, ha coordinado proyectos sobre política digital en México, Costa Rica, República Dominicana, Guatemala, Panamá y Colombia. Es pionero en el uso de inteligencia artificial para perfilar audiencias sociales y entender el humor social. Es orador frecuente en seminarios, clases, cursos, y talleres en universidades y organismos públicos.



Bibliografía

Davidowitz, S. S. (2014), *Todo el mundo miente, lo que el internet y el big data puede decir de nosotros*, Paidós.

Gaalup, L., *Big Data y Política*, Ediciones B.

Harari, Y. N., *Homo Deus*, Debate.

Hill, K. y Mattu, S. (2018), "Lo que los dispositivos saben (y comparten) sobre ti, *TED Talks*. Recuperado de: https://bit.ly/32quD9f

Kosinski, M., Stillwell, D. y Graepel, T. (2013), "Los rasgos y atributos privados son predecibles a partir de registros digitales del comportamiento humano", PNAS, https://www.pnas.org/content/110/15/5802

Mayer-Shönberg, V. y Cukier, K., *Big Data, la revolución de los datos masivos*, Turner.

Paz Pellat, M. A., *Política 2.0. La reinvención ciudadana de la política*. Infotec-Conacyt.

Silver, N., *La señal y el ruido*, Penguin.

Troy, D. (2014), "Mapas sociales que revelan las intersecciones y separaciones de una ciudad", *TED Talks*. Recuperado de: https://bit.ly/3c47Xis

Weill, P. y Ross, J. W., *IT Governance*, Harvard Business School Press.

Cuadernos de
CAPACITACIÓN
ELECTORAL

Descarga la colección completa

1 Planificación Estratégica de la Comunicación Política en las Campañas Electorales.
Ismael Crespo / Ileana Carletta (Editores)

2 Nuevas formas de propaganda, movilización y emociones políticas en tiempos de Pandemia.
Ismael Crespo / Ileana Carletta (Editores)

3 Cuarto de Guerra y estrategia de campaña. El paradigma militar en acción
Javier Sánchez Galicia

4 Investigación del mercado electoral
Federico Irazabal / Elias Aguilar García

5 Uso del Big Data y la Inteligencia Artificial en Campañas Electorales
Alberto Rentería

6 Estrategias de Ataque y contra ataque en campaña.
Alfredo Dávalos López

7 Marketing Político y campañas electorales.
Alfredo Dávalos López

8 Campaña de tierra fidelización y movilización electoral.
Sheyla Dallmeier / Rafael Andrés Arbeláez

9 El Mensaje de Campaña temas, discurso y argumentación para el debate político.
Javier Sánchez Galicia / René Gonzálo Palacios

10 El ABC de la Operación Política
Lázaro Jiménez Aquino

EL PARTIDO DE MÉXICO

IRI INSTITUTO REYES HEROLES

Uso del Big Data y la Inteligencia Artificial en Campañas Electorales se terminó de imprimir en noviembre de 2020 en los talleres de Piso 15 Editorial.
El tiraje consta de 3,000 ejemplares.

Llegó la hora de asumir seriamente que México cambió y que nosotros, como partido, también debemos hacerlo. Hoy más que nunca tenemos que ser un partido moderno, cercano a la gente, un partido que ofrezca soluciones, un partido que conecte con los jóvenes y enarbole las demandas de las mujeres, un partido que fije posiciones progresistas e informadas sobre los grandes problemas nacionales y globales. El PRI siempre ha tenido claro que la política tiene que profesionalizarse. Eso nos diferencia de las demás opciones políticas. Nadie le ha apostado como nosotros a un proceso de actualización permanente, a la formación de liderazgos. Demos pasos importantes en la preparación de nuestros cuadros, con herramientas útiles para el análisis de los fenómenos políticos, sociales y electorales. Con ese propósito, nuestro partido, a través del Instituto de Formación Política "Jesús Reyes Heroles", pone a disposición de sus militantes y simpatizantes, y del público en general, la presente Colección editorial "Cuadernos de Capacitación Electoral".

Alejandro Moreno Cárdenas
Presidente del CEN del PRI

ISBN: 978-607-98995-3-0

9 786079 899530

PISO 15 Editorial