

Transcripción: Módulo 3: Datos personales y algoritmos (Video-clase)

[00:00:12] Hola, bienvenidos al módulo 3 de "Periodismo de internet y tecnología: cómo cubrir sus impactos más allá de los gadgets". Este es el módulo de datos personales y algoritmos.

[00:00:28] Venimos del módulo anterior donde vimos la parte material, la parte grande, pesada de la tecnología de internet y sus consecuencias, sus conflictos. Y en este módulo vamos a hablar de los datos, los componentes más pequeños, a veces casi invisibles, que están en el modelo de negocio en internet, pero que también causan gran parte de los conflictos actuales en la red. Su intercambio es permanente tanto en la capa de infraestructura como en la capa de los contenidos. Y sus consecuencias son importantes tanto en los derechos de las personas, en el tratamiento de los datos personales y también son cada vez más importantes porque con la agregación de muchos de los datos, lo que se llama la big data, y el procesamiento a través de fórmulas, lo que se llama algoritmos, la manipulación de los datos puede dar lugar a mayores desigualdades, es decir, puede dar lugar a discriminación, hacer juegos sociales, de géneros, sexuales, económicos y de diversos tipos. Por eso es importante que como periodistas entendamos cuál es la base de los datos personales para poder abordarlos en nuestras historias.

[00:01:58] Vayamos a los puntos.

[00:02:00] Primero vamos a ver qué abarcan los datos personales. Segundo, vamos a ver cuáles son los conceptos importantes a la hora de trabajarlos. Tercero, vamos a ver cuestiones legales para su abordaje. Después nos vamos a concentrar en que más de algoritmos y después vamos a ver algunos ejemplos en la aplicación en temas de historias de tecnología.

[00:02:27] Empecemos por la base. Como veíamos en el módulo anterior, desde los 90 el modelo de negocios de internet se basa en la recopilación de datos privados para analizarlos y vender luego productos o servicios basados en ellos. Volviendo para atrás. Como dijimos en los 70 en internet, el proyecto era un proyecto público mayormente financiado con dinero público y con algunas empresas también privadas. Pero desde los 90 el modelo de negocios de internet pasó a ser un intercambio de datos, ustedes lo ven mucho en las cookies, pero hoy también se ve en los datos de las redes sociales hacia nosotros, parecen gratuitas, pero cuando usamos cada uno de estos productos o servicios se recopilan datos personales nuestros, se construyen perfiles públicos y privados nuestros. Y a su vez, con esos perfiles se nos venden luego otros productos y servicios. Este es el modelo de negocios de internet.

[00:03:36] Las personas no somos tontas, a través de estos procesos, a través de estos intercambios. Ustedes van a ver que en la bibliografía está el capítulo 9, "Dar aceptar" de mi libro "Guerras de internet", pueden ver muy en profundidad cómo funciona este proceso. Las personas hacemos intercambios, utilizamos productos y servicios tecnológicos que son útiles, que son instructivos, que cumplen una función, que andan bien, que funcionan bien, pero en estos intercambios ofrecemos nuestros datos. ¿Cómo lo hacemos? A través de términos y condiciones. Los términos y condiciones son las constituciones de estos países que nosotros navegamos en internet.

[00:04:23] Estos términos y condiciones en general eran, están siendo cada vez menos, pero todavía son muy engorrosos. Hay un estudio que dice que si nosotros tuviéramos que leer los términos y condiciones de todos los productos y servicios de tecnología que utilizamos por año, serían entre 200 y 300 páginas de Word. Pero todos los años, además, tendríamos que volver a leerlo porque cambian los términos y condiciones. Ahora, por ejemplo, en el Reglamento General de la Unión Europea hay una intención de que los términos y condiciones sean simplificados, pero en general no leemos demasiado los términos y condiciones.

[00:05:06] Más allá de eso, cada vez que aceptamos los términos en cuanto a datos hay consecuencias individuales, pero también hay consecuencias colectivas. Las bases de datos están presentes en los intercambios. Nosotros, los individuos, en nuestro derecho, en el derecho occidental, somos sujetos titulares de los datos. ¿Qué significa esto? Nosotros tenemos una autonomía, autodeterminación informativa sobre los datos. ¿Qué significa esto? Que tenemos el derecho a saber qué se hace con estos datos que se recopilan. Cada uno de nosotros, en nuestros países en donde vivamos, tenemos que tener - a veces tenemos, a veces no, a veces

mejores, a veces peores - distintas leyes de protección de datos personales que nos garantizan esta autodeterminación informativa. Es decir, yo puedo ir a pedirle a algún tipo de autoridad estatal que me diga qué se hizo con mis datos, qué se hizo con mis datos recopilados, ¿sí? Esta es la capacidad de que yo pueda reclamar por mis datos personales. ¿Cómo se utilizó mi información personal? Esto es un derecho. Al mismo tiempo tenemos otro derecho, que es parecido, pero no es el mismo, que es el derecho a la privacidad. El derecho a la privacidad es el derecho a ser dejado solo, a pensar para decidir algo. Es un derecho que tiene que ver más con lo político, con poder tomar una decisión sin injerencia de un tercero. La privacidad es la base de la discusión política, del Estado democrático y por eso también debe ser protegida.

[00:07:02] Estos derechos, además, están protegidos por la Carta de Naciones Unidas de la ONU y por nuestros derechos fundamentales en la Carta de la ONU y en otras leyes de las Naciones Unidas.

[00:07:13] Es decir, hay, muchas veces se suele decir "tenemos derechos digitales". En realidad no es que tenemos derechos digitales, los derechos humanos que tenemos offline, digamos, también rigen en nuestra vida online. La Declaración Universal de Derechos Humanos y otras cartas de Naciones Unidas rigen también en nuestra vida online y los mismos derechos que tenemos deben respetarse en todas las esferas. Y esto está claramente determinado y Naciones Unidas lo ha dicho así. Es muy importante esto cuando trabajen cualquier tipo de historia de tecnología, si no hay algún tipo de ley específica, ustedes sepan que los mismos derechos de Naciones Unidas, de cualquier cualquier Carta de Derechos de Naciones Unidas, rigen para el ámbito de internet. Y también es importante saber esto porque muchas veces se busca sobre regular, se buscan crear leyes ad hoc, leyes nuevas para casos de internet. Y no necesariamente se debe dar el caso, muchas veces las leyes existen y lo que hay que hacer es directamente usarlas para internet, porque lo que están es hechas. Incluso muchas leyes comunes en nuestros países rigen también para el ámbito de internet.

[00:08:38] Volvemos al tema de los datos. Los datos están en bases, en bases de datos. Las leyes de protección de datos, de datos en general, deben proteger tanto los datos que se recopilan en bases de datos privadas, de empresas - pero no sólo de empresas de redes sociales o de internet, puede ser una base de datos de un supermercado online, de cualquier o un supermercado offline, cualquier base de datos que recopile información - o una base de datos del Estado. Los Estados también deben recopilar datos de los ciudadanos para tomar decisiones, para brindar servicios, para garantizar derechos, y también estas bases de datos están protegidas. Volvemos. Tenemos el derecho a la autodeterminación, que se nos diga qué hace, qué se hace con esos datos. Primera diferenciación: hay bases de datos públicas y bases de datos privadas.

[00:09:32] Nosotros en América Latina tenemos una tradición que abreva de la tradición europea. Tenemos leyes de datos que son comprensivas, que protegen a todos los sectores. En cambio, en Estados Unidos hay leyes que son sectoriales, es decir, se protegen los datos o hay leyes de datos que se refieren a datos de salud, datos de educación o niñez, datos bancarios, etcétera, etcétera, etcétera.

[00:10:05] Eso es importante que lo tengan en cuenta, porque mucha de la información de tecnología que nosotros leemos o que consumimos o que estamos recorriendo, cuando hacemos una búsqueda para algún artículo, viene de Estados Unidos, está escrita desde periodistas de Estados Unidos o anglosajones. Entonces es muy importante que tengamos en cuenta esta diferenciación. Para más información sobre esto, hicimos una entrevista con Valeria Milanese, abogada especialista en derecho y datos personales en derecho informático y datos personales, que les va a servir muchísimo para profundizar sobre esto. Les recomiendo mirar.

[00:10:44] Hay otra diferenciación muy importante. Respecto de datos personales.

[00:10:53] El dato personal, ¿qué es? Es una información que identifica o permite identificar a una persona. Mi nombre, mi número de identificación, mi domicilio, etcétera. Y después tenemos los datos personales sensibles, que pueden ser datos de salud, pueden ser datos de afiliación política, de afiliación sindical, de creencias religiosas, de orientación sexual, etcétera. ¿Por qué son sensibles? Porque en caso de estar en manos de alguien, pueden permitir la discriminación

de esa persona en base a ellos. Por ejemplo, si alguien si yo declaré que tengo algún antecedente de salud y esa base de datos cae en manos de una persona o de un futuro empleador y me niega un trabajo en base a esa precondition de salud, puede estar discriminándome previamente. O lo mismo, una afiliación política o partidaria, un gobierno que está censurando o que puede tomar medidas en mi contra, puede generarme un problema de una persecución política.

[00:12:06] Y hay otra cuestión muy importante también en los datos sensibles o en los datos personales está hoy en día la biometría o reconocimiento facial. Y hay otra cuestión muy importante también es que, aunque no sean datos sensibles, hoy en día la big data - la agregación de muchos datos de una persona - aunque los datos en principio no sean datos sensibles, la agregación de muchos datos sobre una persona puede generar un dato sensible en definitiva. Puede generar un perfil de datos que termine siendo sensible contra la persona.

[00:12:47] Muy bien, aquí vemos que empieza a haber conflictos. ¿Por qué? Porque las empresas tienen datos de los usuarios sobre este modelo de negocios del que hablábamos, este modelo de negocios del que Shoshana Zuboff, una autora norteamericana, hoy llama capitalismo de la vigilancia. Datos para vender, datos que se recopilan para después vender. El Estado, a su vez, tiene datos de los ciudadanos que se usan para brindar derechos y servicios en principio, pero también en otros momentos se pueden utilizar recopilados para vigilar o para espiar a los ciudadanos.

[00:13:31] Y las empresas, junto con los Estados, pueden cooperar para brindar estos servicios y garantizar derechos, para desarrollar al país, pero también pueden cooperar para la vigilancia, por ejemplo, como lo denunció Edward Snowden en 2013 en un caso Estados Unidos y muchos otros casos. Entonces, estos intercambios de datos pueden estar funcionando bajo las reglas, con legalidad y con derechos o pueden no hacerlos. Y ahí entra el conflicto y ahí entran, por ejemplo, investigaciones que podamos hacer al respecto.

[00:14:09] Hay algunos conceptos importantes que también me gustaría dejar claros, porque van a aparecer muchas veces en las historias y otras veces no van a aparecer, y por eso los quiero aclarar.

[00:14:20] ¿De qué hablamos cuando hablamos de autorregulación?

[00:14:25] La autorregulación se dice que es el mecanismo regulatorio preferido de la internet. ¿Se acuerdan cuando en el módulo anterior hablábamos de la gobernanza? En la gobernanza de internet hay como una regla que es que bueno que la internet se regula sola.

[00:14:43] La autorregulación ha sido como el método en que las grandes empresas, sobre todo las grandes empresas de plataformas como Facebook o como Google, han dado todo para decir "bueno, nosotros nos damos nuestros propios estándares. Nosotros ponemos nuestras propias reglas", pero eso tiene algunos riesgos. Porque si las mismas empresas, cuando tienen conflictos con temas públicos, establecen sus propias reglas, eso puede llevar también a otros problemas. Por ejemplo, si las mismas empresas de plataformas van a poner reglas para organizar el discurso de odio, organizar el debate en público, eso puede traer otros problemas. Entonces hay que tener cuidado o hay que poner entre paréntesis la palabra autorregulación.

[00:15:30] Otro tema importante cuando estamos hablando de derechos y datos es tener en cuenta los casos de la jurisprudencia en casos de internet. Les pongo dos ejemplos muy brevemente. Casos de derecho al olvido en España, porque en otros lugares es distinto, el caso Mario Costeja. Caso Apple versus el Estado Estados Unidos por el desbloqueo de celulares. Cada tanto hay algunos casos judiciales que cambian algunas reglas o que se vuelven importantes porque marcan algún precedente. También un tercer punto de conceptos a considerar es que las jurisdicciones en internet van a ser siempre un tema conflictivo o a discutir, y que siempre van a estar presentes en nuestras historias y que lo van a tener que tener en cuenta. ¿Por qué? Porque hay muchas empresas internacionales - Facebook, Google, Apple - que vamos a tener que abordar temas que cuando vamos a tener que resolver un tema local, van a, estas empresas muchas veces decir o ustedes se van a enfrentar como periodistas a, bueno, hay que recurrir a la sede internacional o a la sede en Irlanda o a la sede en tal país para resolver

el tema. Hay un ejemplo en nuestra bibliografía de una nota de Irina Sternik sobre este tema reciente de un caso en Argentina que les recomiendo leer, pero siempre va a ser estén en Colombia, estén en México, estén en Perú siempre van a tener este tipo de cuestiones jurisdiccionales cuando estén trabajando con temas de tecnología.

[00:17:10] En temas de datos personales hay una cuestión muy importante que es el Reglamento General de Protección de Datos. El RGPD de la Unión Europea que tiene también jurisdicción en compañías de Estados Unidos y en muchos de nuestros países, porque obviamente al pasar entre un país, entre en país y otro, al ceder datos entre un país y otro en el mundo, también nos afecta. ¿Pero por qué tiene incidencia? Porque se ha convertido en el faro para el cual también Latinoamérica puede llegar a mirar, a mirar y adaptar algunas de sus leyes de datos personales y como modelo a seguir.

[00:17:53] Lo menciono y menciono alguno de sus preceptos, del Reglamento General de Datos de Europa, porque es como el modelo de datos personales guía en este momento. ¿Qué es lo que establece? Entre otras cosas, el derecho a saber cómo se utilizan mis datos, para qué fin se utilizan mis datos - esto debe estar definido previamente - por cuánto tiempo se conservan, con quién se comparten. Mi derecho como titular del dato a reclamar o retirarlo de una base de datos o a retirarlos, en caso de que sean distintos datos, el derecho a que se me presenten los términos y condiciones de una manera clara y sencilla, el derecho a opinar sobre las decisiones automatizadas - esto incluye el tema de los algoritmos y sus sesgos -, y la capacidad que tiene en este caso la Unión Europea y sus Estados de multar si los privados o agencias públicas no cumplen con alguno de estos preceptos. Les recomiendo que repasen el Reglamento General de Protección de Datos de Europa como norma general, por la cual guiarse como horizonte general en el cual guiarse. Después hay cuestiones de los datos que se mezclan con el tema de la seguridad y en eso vamos a trabajar, seguridad por otro lado, porque los temas de seguridad no son específicamente de datos.

[00:19:30] Ahora en este módulo quiero pasar al tema de algoritmos. ¿Por qué quiero pasar al tema algoritmos y por qué lo estamos vinculando con el tema de los datos? Porque hoy los datos agregados, digamos, hoy la capacidad que tenemos de recopilar y acumular una gran cantidad de datos y procesarlos a través de fórmulas algorítmicas para tomar decisiones están muy, muy presentes, muy, muy presente en nuestra vida cotidiana. Es lo que muchos denominan inteligencia artificial.

[00:20:05] Yo prefiero, y muchas preferimos llamarlo decisiones automatizadas. ¿Por qué? Porque son una gran cantidad de datos que, pasando por una fórmula, permiten tomar una decisión automáticamente, pero no están automáticamente porque esa fórmula la construye una persona, la construye alguien. Entonces, Cathy O'Neil, de quien tenemos una TED Talk, una charla TED en nuestra bibliografía, dice que hoy los algoritmos son opiniones expresadas en código, es decir, no son ingenuos. Nosotros podemos estar ordenando la realidad de acuerdo a una fórmula matemática y que con eso podemos estar creando nuevas realidades que sean más desiguales y por eso necesitamos cada vez más estudiar los algoritmos.

[00:21:00] También lo menciono porque Philip Alston, el Relator Especial para la pobreza de la ONU, ha dicho recientemente que estas decisiones automatizadas en las sociedades pueden crear mayores desigualdades y mayor pobreza, y que corremos el riesgo de automatizar la pobreza si no nos ponemos a estudiar cómo se toman estas decisiones que muchas veces, si se las dejamos solamente a la tecnología, las automatizamos. Y entonces como periodistas tenemos una gran responsabilidad en ir a abrir esas cajas negras, investigar qué está sucediendo y construir muchísimas más historias, preguntar cómo se están tomando esas decisiones.

[00:21:50] Hay un ejemplo de cómo se están tomando esas decisiones en una nota que les compartí, que es la de Mariana Pernas, en Chequeado, sobre la videovigilancia en Buenos Aires. Se instalan cámaras de videovigilancia, se toman decisiones sobre las vidas de los ciudadanos. Vayamos a preguntar cómo se toman esas decisiones. Vayamos a hacer lo que hizo Cathy O'Neil en su libro "Armas de destrucción matemática". Es muy costoso, pero podemos hacerlo.

[00:22:18] Las empresas, además hoy, están avanzando y diciendo "bueno, nuestros algoritmos son más éticos, estamos incluyendo ética en los algoritmos.

[00:22:28] Yo prefiero pensar en una construcción social de los algoritmos, en que las empresas no sean las únicas que tomen las decisiones, que nosotros podamos investigar sobre cómo están construidas estas decisiones. Que haya universidades, académicos, sociedad civil que pueda participar en otras construcciones. Les recomiendo muchísimo el trabajo de Julia Angwin, hecho en ProPublica, hoy en The Markup, que siempre ha sido una periodista que ha seguido muchísimo estos temas en términos de: el derecho a saber qué decisiones toman los algoritmos.

[00:23:08] ¿Qué temas necesitan todavía ser muchísimo más innovados en términos de decisiones automatizadas? Vigilancia y biometría. En todas las ciudades del mundo, y de Latinoamérica muchísimo más, están y se siguen implementando decisiones de vigilancia y biometría por seguridad o por inseguridad, se suele decir, o por terrorismo. Y esas decisiones suelen ser tecnologías que no están muy explicadas. Los temas de salud y educación, que cada vez se automatizan más, suelen no estar muy explicados. Los temas crediticios o bancarios también suelen ser temas que hay que investigar mucho más. Y los usos de bases de datos masivas por parte de las empresas y de los Estados también deben llevarnos a investigar más sobre ellos. Hay ejemplos en este módulo, como les decía, en principio la nota de Chequeado y otra nota mía también sobre datos y aplicaciones para la pandemia del COVID. Podemos empezar por esos, podemos compartir otras investigaciones, y nos vemos en el siguiente módulo. Gracias.