INTELIGENCIA ARTIFICIAL, MACRODATOS Y METADATOS EN LAS INVESTIGACIONES POLICIALES Y EN EL PROCESO PENAL

Trabajo publicado en *Las nuevas tecnologías y la inteligencia artificial, al servicio del proceso (MARTÍN RÍOS, P., VILLEGAS DELGADO, C., Directores), A Coruña: Colex, 1.ª edición 2023. ISBN: 978-84-1359-779-9, pp. 231-259.*

INTELIGENCIA ARTIFICIAL, MACRODATOS Y METADATOS EN LAS INVESTIGACIONES POLICIALES Y EN EL PROCESO PENAL

Juan Carlos ORTIZ PRADILLO[1]

Universidad Complutense de Madrid

Instituto de Derecho Europeo e Integración Regional (IDEIR)

RESUMEN:

Como ya sucediera con la distinción entre contenidos de una comunicación y datos externos a la misma, la posibilidad de descomponer la información digital en datos y metadatos abre un nuevo horizonte en la prevención y persecución eficaz del delito, por cuanto las actuales capacidades de recolección de grandes conjuntos de información digital (*Big Data*) y su clasificación y análisis (*Data Mining*) a través de herramientas de la Inteligencia Artificial, para extraer y analizar los metadatos, permiten a las autoridades localizar y extraer información cada vez más precisa a la hora de condenar o absolver al acusado. Esa información casi "subatómica", sin embargo, puede tener una enorme afectación sobre los derechos fundamentales de los ciudadanos, de modo que es preciso examinar qué encaje legal pueden tener las actuales técnicas de recolección y análisis de los metadatos, llamados a convertirse en la nueva *regina probatorum* de todo juicio.

ABSTRACT:

As it already happened, after *Malone v. United Kingdom*, with the distinction between the content of a communication and external data of such communication, the possibility of decomposing digital information into data and metadata opens a new horizon in the prevention and prosecution of crime, since the current capabilities of collection of large sets of digital information (Big Data) and its classification and analysis (Data Mining) through Artificial Intelligence tools, allow LEAs to locate and extract increasingly accurate information when convicting or acquitting the accused. This almost "subatomic" information, however, can have an enormous impact on the fundamental rights of citizens, so it is necessary to examine if the current techniques of collection and analysis of metadata, are in accordance with the law and can be considered as necessary in a democratic society.

PALABRAS CLAVE:

[1] juancarlosortiz@ucm.es ORCID: 0000-0001-6092-6137. WoS ResearcherID AAG-7424-2019. El presente trabajo se enmarca en el Proyecto de Investigación "Novel Strategies to Fight Child Sexual Exploitation and Human Trafficking Crimes and Protect their Victims (HEROES)", financiado por la Comisión Europea, dentro del Programa Marco de Investigación e Innovación de la Unión Europea *Horizonte 2020*, Convocatoria H2020-SU-SEC-2020, Proyecto de Investigación 101021801. Investigador Principal: Luis Javier García Villalba.

Inteligencia Artificial, Tecnología, Big Data, Metadatos, Algoritmos, SOCMINT, Investigación Policial, Proceso Penal.

KEYWORDS:

Artificial Intelligence, Big Data, Metadata, Technology, Algorithm, Police Surveillance, Social Network Analysis, Criminal Justice.

I. INTRODUCCIÓN

24 de junio de 2015. Antonio (nombre ficticio) decide salir a cenar con su pareja a un restaurante de la playa de Badalona para celebrar su santo en la noche de San Juan. Esa misma noche, tres hombres entraron a la fuerza en un piso del barrio de El Carmel, en Barcelona; golpearon y torturaron a su ocupante hasta que ésta les indicó dónde se encontraba la caja fuerte y se llevaron 55.000 euros en efectivo. Antonio fue detenido porque la denunciante le señaló en una identificación fotográfica. A pesar de contar con el testimonio de descargo de su pareja, el juzgado ordenó su entrada en prisión provisional, donde permaneció varios meses hasta que se finalmente se acordó su libertad bajo fianza de 10.000 euros.

Al cabo de unos años, la sentencia absolutoria dio por probado que Antonio no estuvo en el lugar de los hechos. En el juicio, además de valorar los "tradicionales" testimonios de los testigos que declararon haberle visto aquella noche en el Restaurante, el tribunal fundamentó su absolución en el "tecnológico" informe pericial sobre los metadatos extraídos de las fotos y videos que su acompañante hizo aquel día y colgó en Facebook e Instagram; los datos de creación y *posteo* de esas imágenes en dichas redes

sociales se consideraron veraces y muy difícilmente manipulables, salvo que se hubiera tenido acceso a los servidores de dichas redes, lo cual resulta harto difícil para cualquier ciudadano.

22 de agosto de 2016. Una joven camina por las calles de una pequeña localidad gallega durante sus fiestas patronales, mientras *wasapea* con sus amigos, hasta que desaparece sin dejar rastro. Su cadáver sería hallado 497 días después en un pozo de una nave a más de 20 kilómetros de distancia. Durante el juicio, el acusado sostuvo que abordó a la joven porque ésta le había sorprendido robando gasoil en un callejón. Sin embargo, el informe pericial efectuado por el Grupo de Apoyo Técnico Operativo (GATO) de la Unidad Central Operativa (UCO) de la Guardia Civil, a partir del análisis forense del teléfono de la víctima ―datos GPS, datación de los mensajes enviados y recibidos, celdas y datos de triangulación con las antenas de telefonía y de conexión a las redes wifi de determinados establecimientos― permitieron reconstruir con gran precisión el recorrido efectuado por la joven por cada una de las calles de la localidad, así como el momento y lugar exactos en que se produjo el ataque, que distaba mucho de ser el callejón mencionado por el acusado.

Ambas historias reales demuestran, no sólo que la tecnología ha ido cobrando un papel protagonista cada vez mayor en el proceso penal, a la par que el propio desarrollo tecnológico de la sociedad, sino que la específica información extraíble de los archivos digitales que se generan, almacenan o transmiten con motivo del uso de dispositivos electrónicos está llamada a convertirse en la *regina probatorum* de un juicio; dicha información será decisiva a la hora de declarar probado dónde y cuándo se produjo un hecho concreto o quién fue su autor, y por tanto, a la hora de condenar o absolver al acusado.

II. LOS MACRODATOS (*BIG DATA*) Y SU ANÁLISIS (*DATA MINING*) EN LA INVESTIGACIÓN DEL DELITO

«El conocimiento es poder» (*ipsa scientia potestas est*); quien posea mayor conocimiento estará en condiciones de acceder y ostentar mayor poder. En segundo lugar, la tecnología, y en particular, el desarrollo de la Inteligencia Artificial, constituyen los actuales motores que están impulsando el conocimiento y la transformación digital de nuestra Sociedad, de modo que puede establecerse una relación directa entre el desarrollo y empleo de la tecnología y la adquisición de conocimiento. Y en tercer lugar, no hay que olvidar que el "combustible" o materia prima utilizada por los algoritmos y programas de las distintas herramientas de Inteligencia Artificial son los datos, y resulta que una de las más notables características de nuestra actual Sociedad es la generación de inmensas cantidades y tipologías de datos en formato digital con motivo de casi cualquier actividad humana, cuya recolección, almacenamiento, tratamiento y análisis a través de técnicas avanzadas de clasificación y segmentación ―*Big Data* y *Data Mining*― permiten tener una visión sin precedentes del comportamiento humano[2].

[2] La Resolución de 14 de marzo de 2017, sobre las implicaciones de los "macrodatos" en los derechos fundamentales: privacidad, protección de datos, no discriminación, seguridad y aplicación de la ley (2016/2225(INI) utiliza el concepto de «macrodato», al que se refiere como *la recopilación, análisis y acumulación constante de grandes cantidades de datos, incluidos datos personales, procedentes de diferentes fuentes y objeto de un tratamiento automatizado mediante algoritmos informáticos y avanzadas*

Si interrelacionamos las tres afirmaciones precedentes, resulta lógico inferir que las autoridades encargadas de la investigación criminal se esfuercen por conocer y emplear las capacidades tecnológicas de la IA en cualquier pesquisa, llamadas a revolucionar el modo en que se indagan, localizan y obtienen evidencias e indicios sobre la perpetración de los delitos y sobre la identidad de los delincuentes.

De hecho, ya asistimos a un auténtico cambio de paradigma en el modo de actuación de las autoridades policiales y judiciales a la hora de llevar a cabo sus funciones de prevención, detección, investigación y enjuiciamiento de los delitos, pues la búsqueda de "vestigios" se ha redirigido hacia el acopio y análisis de "datos"; esto es, toda aquella información en formato electrónico relacionada con el entorno virtual de los sujetos intervinientes en cualquier comisión delictiva. De una parte, el entorno digital de la vícitma, a través de cuyo análisis se habilitarán específicas líneas de investigación que agilicen la investigación de los hechos y conduzcan a la delimitación e identificación de sus atacantes. Y de otra parte, el entorno digital de los lugares en que se produjeron los hechos presuntamente delictivos, así como de los sujetos sospechosos de haber participado en los mismos, de modo que la recopilación y examen de todo ese caudal de información digital podrá ser utilizado como fuentes de prueba acreditativas de la comisión y autoría de los hechos objeto de investigación[3].

En resumen, el uso de la tecnología ha revolucionado las técnicas de investigación delictiva y ha obligado a reformular las pautas y metodologías con la que los investigadores delimitan y examinan el escenario del crimen, así como también ha redimensionado el contenido y alcance de la protección ofrecida por los derechos fundamentales.

1. LA INTELIFENCIA ARTIFICIAL Y LA «PREDICCIÓN DEL COMPORTAMIENTO» APLICADO CON FINES COMERCIALES: *GOOGLE TRENDS* Y EL ESCÁNDALO *CAMBRIDGE ANALYTICA*

Google Trends es una herramienta de acceso libre y gratuita que arroja información gráfica sobre la frecuencia de temas y términos de búsqueda, agregados y segmentados por regiones y períodos de tiempo. Gracias a la recopilación ―anonimizada― de los datos introducidos por los usuarios cuando utilizan el motor de búsquedas de Google, las empresas pueden conocer así la popularidad o número de búsquedas de unas determinadas palabras o frases durante un concreta franja de tiempo y en una determinada zona geográfica, y adecuar así sus campañas publicitarias y la producción y el stock de sus productos a la previsible alza o reducción en la demanda de los mismos.

La red social Twitter ha servido para detectar cuáles son los temas del momento ―*trending topic*―, con el simple análisis de cuáles son las palabras o frases más repetidas en un determinado momento. Desde 2010, Twitter habilitó los trending topics locales, además de los mundiales, y permitió que se pudiera seleccionar entre más de una treintena

técnicas de tratamiento de datos, utilizando tanto datos almacenados como datos transmitidos en flujo continuo, con el fin de generar correlaciones, tendencias y patrones (analítica de macrodatos).

[3] ORTIZ PRADILLO, J. C. (2021) "Big Data, Vigilancias Policiales y Geolocalización: nuevas dimensiones de los Derechos Fundamentales en el Proceso Penal", *Diario La Ley*, núm. 9955, de 18 de noviembre de 2021.

de países y ciudades para conocer cuáles eran los temas más hablados en ese momento y en ese lugar en particular.

El desarrollo de múltiples herramientas basadas en esa idea del análisis de datos procedentes de la Internet de cara a la extrapolación de tendencias ha dado lugar a un modelo de negocio denominado el «capitalismo de la vigilancia[4]»; recabar y analizar los datos extraídos del comportamiento humano en las redes sociales o en sus comunicaciones e interacciones a través de internet para predecir sus comportamientos futuros (el *profiling* o la creación de «perfiles de personalidad» de sus usuarios con fines comerciales), que posteriormente son utilizados para aumentar su volumen de negocio, adecuar la fijación de los precios a futuras demandas, o bien, monetizados a través de su venta a terceros.

En 2004, y durante la temporada de huracanes que azotan las costas de Florida y el mar caribe, la cadena de supermercados WalMart ―con más de 10.500 tiendas y más de 2 millones de empleados en 24 países― llevó a cabo un análisis exhaustivo de los datos recopilados con motivo de sus ventas en los EE.UU. (los productos que vende en cada tienda, Estado o región; la periodicidad con la que se reponen, etc.) y comprobaron que el producto más vendido antes de un huracán no son las linternas, sino la cerveza, o que las tartaletas de fresa aumentan en más de siete veces su tasa de ventas normal. Con dicha información, los ejecutivos de Wal-Mart decidieron enviar camiones repletos de los productos más demandados hacia las tiendas situadas en las zonas que se verían afectadas semanas más tarde por los huracanes y tormentas tropicales[5].

A finales de 2008, la revista *Nature* publicó un estudio que mostraba cómo Google había desarrollado un programa computacional ―*Google Flu Trends*― que identificaba más de 45 términos de búsqueda relacionados con la gripe y los cotejaba con los datos recopilados por el Centro de Control de Enfermedades (CDC) de los EE.UU. Con dicha información, Google era capaz de "predecir" con semanas de antelación en qué regiones tendría una mayor incidencia la gripe, facilitando así la adopción de medidas preventivas y asistenciales por parte de las autoridades médicas[6].

Y durante las elecciones presidenciales de 2016 en los EE.UU., la consultora *Cambridge Analytica*, una empresa con sede en Londres que usaba el análisis de datos para desarrollar campañas publicitarias, obtuvo información de más de 87 millones de perfiles de Facebook a través de una aplicación llamada "Esta es tu vida digital" ―*This Is Your Digital Life*―. Con dichos datos, elaboró perfiles psicográficos (e intenciones de voto) de sus usuarios basados en su información personal, su red de amistades, mensajes y posts, actualizaciones de estado, o "me gusta", que posteriormente vendió a otras empresas y partidos políticos, quienes los utilizaron para enviarles anuncios personalizados, pero también noticias falsas[7].

---

[4] ZUBOFF, S. (2019). *The Age Of Surveillance Capitalism*. New York: PublicAffairs.

[5] HAYS, C. L. (2004) "What Wal-Mart Knows About Customers' Habits". Noticia publicada el 14 de noviembre de 2004 en el periódico New York Times. Accesible en: http://www.nytimes.com/2004/11/14/business/yourmoney/14wal.html.

[6] GINSBERG, J., MOHEBBI, M., PATEL, R. et al. (2009) "Detecting influenza epidemics using search engine query data". Nature 457, 1012–1014. https://doi.org/10.1038/nature07634.

[7] MANHEIM, K. M., KAPLAN, L. (2019) "Lyric, Artificial Intelligence: Risks to Privacy and Democracy", 21 Yale Journal of Law and Technology 106. Accesible en: https://ssrn.com/abstract=3273016.

2. LA INTELIGENCIA ARTIFICIAL Y LA «PREDICCIÓN DEL COMPORTAMIENTO» APLICADO A LA INVESTIGACIÓN CRIMINAL: VIGILANCIA MASIVA Y POLICÍA PREDICTIVA

De igual modo, los gobiernos no han dudado en aprovecharse de este modelo económico basado en el análisis del comportamiento humano para mejorar los resultados de sus políticas públicas y contribuir al desarrollo de sus naciones. El *Big Data* puede facilitar una mejor y más eficiente recaudación tributaria —tanto a la hora de determinar qué tributos pueden ser objeto de alza o reducción, como a la hora de aflorar actividades donde se focaliza la economía sumergida—. Y la aplicación de la IA puede hacer que se reduzca la siniestralidad en los centros de trabajo —a través del examen y análisis de la actividad laboral, se pueden identificar tareas peligrosas y proceder a su robotización—, así como la siniestralidad vial —el análisis del tráfico en tiempo real, a partir de la interconexión entre vehículos, vías y centros de control, permitirá una movilidad más segura; identificará las vías, tramos y condiciones de mayor peligrosidad; y agilizará la toma de decisiones y respuestas ante accidentes de tráfico—.

¿Pueden el Big Data y la IA mejorar la política criminal de los Estados y, en particular, aumentar la eficacia de la respuesta estatal a la delincuencia? La digitalización de la Justicia ya es una realidad presente en múltiples ámbitos de actuación, como puedan ser la automatización de la tramitación de los procedimientos, la creación de plataformas de resolución de litigios en línea, la informatización e interoperatividad de los expedientes judiciales, así como el uso de blockchain en su formación y securización, el desarrollo de actuaciones orales a través de recursos telemáticos y su textualización mediante algoritmos, o la robotización en el cálculo de probabilidades de riesgo para la adopción de resoluciones judiciales, por citar algunos.

En concreto, y en el el ámbito de la investigación del delito, el uso de la tecnología en la recopilación de grandes conjuntos de datos y su análisis inteligente a través de fórmulas matemáticas persigue dos grandes finalidades: la investigación y persecución eficaz del delito presuntamente cometido, mediante la utilización de técnicas de «vigilancia masiva» o «tecnovigilancia[8]», así como la reducción de la tasa de criminalidad y la prevención de conductas criminales mediante la formulación de patrones y pronósticos de «policía predictiva».

2.1. *Data Mining*, vigilancia masiva en tiempo real y reconstrucción de movimientos (geolocalización histórica) de los investigados

Respecto a la primera —la vigilancia masiva—, el empleo de las capacidades tecnológicas al servicio de la investigación del delito y centradas en la búsqueda y recogida de vestigios del delito y a la averiguación de sus autores permite a las autoridades

[8] Además de la expresión "mass surveillance", otros términos anglosajones comúnmente utilizados son los de "electronic surveillance", "Internet surveillance" y "online surveillance". Nosotros preferimos su traducción como "tecnovigilancia". Vid., por todos, LLAMAS FERNÁNDEZ, M., GORDILLO LUQUE, J. M. (2007) "Medios técnicos de vigilancia", en *Los nuevos medios de investigación en el proceso penal. Especial referencia a la tecnovigilancia*, Madrid: CGPJ, Cuadernos de Derecho Judicial, 2007-II, p. 236.

resolver los delitos de forma más rápida, eficaz y segura[9]. Con la expresión «Vigilancia masiva» se alude al uso policial de técnicas de IA y *Big Data analytics* sobre toda esa información digital masivamente recopilada y almacenada por terceros (operadoras, prestadores de servicios de Internet, empresas privadas, administraciones públicas…) de cara a facilitar la investigación del delito mediante la reconstrucción de las circunstancias en que tuvo lugar una determinada conducta, quiénes se encontraban en el lugar de los hechos, cuáles fueron sus interacciones, y dónde quedaron almacenadas las fuentes de pruebas acreditativas de dichas circunstancias. Más allá de tratar de localizar e interrogar a los posibles testigos presenciales, se trata ahora de proceder a la recopilación y análisis de toda esa información digital que permita la identificación y captura de sus autores, y la prueba de su participación en los hechos delictivos. Frente a la fragilidad de la memoria humana, la masiva conservación de archivos digitales constituyen los nuevos "caladeros" a donde acudir.

Y para ello, las autoridades ya no necesitan invertir ―aunque lo hagan― en dicha recolección y conservación; les basta con conseguir la colaboración de las empresas privadas, que son quienes manejan toda esa ingente cantidad de información creada, almacenada o transmitida en formato digital. Los Estados tecnológicamente desarrollados, también los democráticos, se han convertido en lo que BALKIN califica como *Estados Vigilantes*[10], caracterizados por servirse de las capacidades de las empresas privadas, no sólo para brindar servicios sociales valiosos, sino también para identificar problemas, evitar amenazas potenciales y combatir la delincuencia.

Como ejemplo concreto, ya hemos advertido sobre la posibilidad de poder «geolocalizar en tiempo real[11]» a un individuo sospechoso de haber cometido un delito a través de diversas herramientas tecnológicas, tales como el uso de balizas adosadas a su vehículo, ropa u objetos que porte consigo; vigilancias policiales sistemáticas, apoyadas por drones y otros instrumentos dotados de cámaras de visión térmica y nocturna o sistemas de reconocimiento facial o de determinadas características biométricas del investigado; o el acceso en tiempo real a los datos de localización que emita su propio terminal móvil. Pero resulta que la recolección y análisis inteligente de datos personales almacenados por entidades privadas, con otros fines comerciales, también permite reconstruir la actividad llevada a cabo por una persona y conocer con gran precisión los lugares visitados, las rutas utilizadas, las personas con las que contactó, sus hábitos y rutinas («geolocalización histórica[12]»): la policía puede servirse para ello de los datos referidos a la matrícula del vehículo empleado, que indiquen por qué vías y en qué momentos circuló dicho vehículo, dónde estacionó, dónde respostó combustible, o cuándo y dónde fue objeto de una multa; de los datos facilitados por los establecimientos

[9] ORTIZ PRADILLO, J. C.: *Problemas procesales de la ciberdelincuencia*, Colex, Madrid, 2013, p. 28.

[10] BALKIN, J. M. (2008) "The Constitution in the National Surveillance State", *Minnesota Law Review 93* (1), pp. 1-25.

[11] ORTIZ PRADILLO, J. C. (2021) "Big Data, Vigilancias Policiales y Geolocalización: nuevas dimensiones de los Derechos Fundamentales en el Proceso Penal", *Diario La Ley*, núm. 9955, de 18 de noviembre de 2021.

[12] Como específico examen de las diferencias y métodos de geolocalización histórica o reconstrucción de movimientos, vid. ORTIZ-PRADILLO, J. C. (2022) "Vigilancias policiales y utilización de dispositivos de seguimiento, localización y captación de la imagen", Reflexiones en torno al Anteproyecto de Ley De Enjuiciamiento Criminal de 2020 (Coordinadora Olga Fuentes Serrano). Valencia: Tirant lo blanch, pp. 813-881.

hoteleros y de hospedaje para saber dónde se alojó y pernoctó; de los datos PNR[13] remitidos por parte de las compañías aéreas y otras entidades para conocer qué vuelos tomó y con qué itinerarios; de los datos asociados a compras online y operaciones de pago cuando las mismas se hayan realizado a través de medios electrónicos de pago para llegar a determinar lugares habituales de ocio; de los datos almacenados por aquellas empresas dedicadas a ofrecer servicios comerciales basados en la geolocalización de sus usuarios cuando llevan a cabo prácticas deportivas (*Strava, Endomondo, Runtastic, Wikiloc*, etc.) para determinar sus rutinas e incluso su domicilio; y por supuesto, también podrán acudir a las operadoras de servicios de acceso a la telefonía móvil e internet para que cedan los datos de tráfico y de localización del terminal empleado por el investigado conservados en virtud de las diversas disposiciones legales en materia de cesión de tales datos conservados.

La geolocalización histórica o "reconstrucción de movimientos" de una persona sospechosa de haber participado en unos hechos delictivos concretos resulta, cada vez, más habitual en la operativa policial y en la práctica judicial.

El 26 de octubre de 2015, sobre las 07.55 horas, varios encapuchados accedieron armados a una sucursal bancaria de la localidad de Vallbona d'Anoia (Barcelona, España), se apoderaron de más de 11.000 euros y huyeron en un vehículo BMW azul. No pudieron ser identificados por los testigos ni por las cámaras de seguridad del banco, ni tampoco se encontraron huellas dactilares. Sin embargo, las autoridades lograron identificar y condenar a los autores del atraco.

Para ello, una de las primeras actuaciones fue acudir a las operadoras para obtener los datos de geolocalización de los teléfonos móviles que se encontraban operativos en la dirección de la sucursal bancaria y a la hora en que se produjo el atraco. Debidamente analizados dichos datos, se comprobó que varios de sus titulares presentaban antecedentes penales por robo con violencia y se acordó la interceptación judicial de sus comunicaciones, junto con la cesión de los datos de tráfico de llamadas y datos de localización de las mismas en un periodo de tiempo tanto previo como posterior a la comisión de los hechos. Se comprobó que varios de los sospechosos habían estado días antes del atraco en las inmediaciones de la sucursal. Y las filmaciones de las cámaras de la entidad habían grabado a una persona con la misma vestimenta que una de las personas que entraron en el banco el día de los hechos. Esas filmaciones también facilitaron que varios testigos reconocieran a los acusados, y finalmente, en el registro domiciliario de estos se halló vestimenta con las mismas características de la utilizada en el atraco[14].

2.2. La Inteligencia Artificial y la elaboración de pronósticos delictivos

[13] Las siglas PNR significan «Passenger Name Record» —Registro de Nombres de Pasajeros— y se refieren a informaciones que determinadas empresas y compañías aéreas deben remitir a los Estados, en virtud de lo establecido en la Directiva (UE) 2016/681 del Parlamento Europeo y del Consejo de 27 de abril de 2016 relativa a la utilización de datos del registro de nombres de los pasajeros (PNR) para la prevención, detección, investigación y enjuiciamiento de los delitos de terrorismo y de la delincuencia grave. En España, dicha Directiva ha sido incorporada a través de la Ley Orgánica 1/2020, de 16 de septiembre, sobre la utilización de los datos del Registro de Nombres de Pasajeros para la prevención, detección, investigación y enjuiciamiento de delitos de terrorismo y delitos graves.

[14] STS, Sala Penal, secc. 1ª, núm. 484/2018, de 18 de octubre. ECLI: ECLI:ES:TS:2018:3575.

Constituye el sueño de cualquier gobernante ser capaz de erradicar el crimen, o al menos reducirlo a su mínima expresión, anticipándose a su realización a través de la adopción de medidas que eviten o disminuyan los factores que facilitan la comisión de delitos. Pero los actuales programas de «policía predictiva» poco tienen que ver con el relato *El informe de la minoría* (The Minority Report, escrito por Philip K. Dick en 1956 y llevado al cine por Steven Spielberg en 2002) en donde unos seres ―los "precognitivos"― prevén todos los delitos antes de que ocurran y una división de la policía llamada *PreCrime* detiene a los sospechosos antes de que puedan cometer delitos reales.

La policía predictiva o *PredPol* se basa en la recopilación y análisis de grandes conjuntos de datos, sobre los que se aplican algoritmos que tienen en consideración múltiples factores espaciales, personales y ambientales (v. gr., los días en que se abonan los salarios, la ubicación de tiendas de bebidas alcohólicas, datos de ejecuciones hipotecarias y lanzamientos judiciales, condiciones climáticas, rutas de escape, celebraciones, eventos multitudinarios…) para la elaboración de patrones y pronósticos delictivos basados, bien en la determinación de posibles lugares más propensos a ser objetivos de hechos delictivos, bien en la elaboración de listas de sospechosos o personas más propensas a cometer actividades criminales. Esto es, las autoridades policiales también emplean técnicas de predicción del comportamiento basada en la recopilación masiva de información digital y su análisis inteligente, gracias a herramientas de Inteligencia Artificial y *Machine Learning*, para la elaboración de patrones de delincuencia y mapas de riesgos con los que reducir la delincuencia mediante la elaboración de "pronósticos" delictivos.

El mayor ejemplo de estas nuevas técnicas policiales lo encontramos en la ciudad de Nueva York. En 2012, su Departamento de Policía, en colaboración con Microsoft, anunció el lanzamiento del programa "Sistema de Conciencia de Dominio" ―*Domain Awareness System,* DAS―; una medida inicialmente ideada para la lucha contra el terrorismo (sobre todo, tras el atentado terrorista en 2010 en Times Square), pero que actualmente es utilizado para la investigación de cualquier tipología delictiva. El sistema se nutre de los datos recopilados por las más de 9.000 cámaras CCTV y de reconocimiento automático de matrículas dispuestas en las calles de Nueva York, torres con sistemas de detección de tiroteos y de emisiones de agentes químicos, las llamadas efectuadas al 911, datos extraídos de los informes de detenciones policiales, o de citaciones y órdenes judiciales.

A nuestro juicio, lo llamativo es que las autoridades hayan potenciado y focalizado sus esfuerzos en el desarrollo de esa "policía predictiva" en el terreno netamente policial y apenas se haya impulsado su utilización con perspectiva global en todo el ámbito de las Administraciones Públicas, pues el *Big Data* y el empleo de la IA por parte de la Administración debería ser la mayor prioridad política para impulsar el desarrollo de nuestra Sociedad y la erradicación de la corrupción y el fraude.

Hace más de una década que la Comisión Europea, a través de la Oficina de Protección de Datos, impulsó el desarrollo de la herramienta *Arachne* para identificar los proyectos susceptibles de presentar riesgos de fraude, conflicto de intereses o irregularidades en la gestión de los fondos europeos FSE y FEDER. Dicha herramienta informática se basa en la prospección y en el enriquecimiento de datos provenientes de la gestión de tales fondos (beneficiarios, códigos de identificación fiscal, número de trabajadores, volumen de negocios, subcontratistas, miembros del consorcio, accionistas, perfiles de personas del medio político ―PMP―, así como de sus familiares y allegados, etc.) para facilitar la clasificación del riesgo y la emisión de alertas (*red flags*). Y de la

misma manera, la Oficina Europea para la Lucha contra el Fraude (OLAF) lleva tiempo ensayando sistemas automatizados basados en el *Data Mining* para examinar lenguaje dudoso y sospechoso que ayude a detectar conductas fraudulentas, colusión, doble facturación, etc., como por ejemplo, la expresión "borre este mail después de leerlo[15]".

Por ello, deben ser bienvenidas e institucionalmente potenciadas aquellas iniciativas como la puesta en marcha en 2015 en la Comunidad Valenciana el proyecto del Sistema de Alertas Tempranas (SALER), también calificado como *Sistema de Alertas Tempranas Anticorrupción* (SATAN), con el que aprovechar la digitalización de las actuaciones administrativas y la generación de ese gran volumen de información en el seno de las Administraciones para elaborar patrones de comportamiento que posibiliten la identificación de malas prácticas que puedan devenir en casos de fraude o corrupción.

III. LOS METADATOS Y SU UTILIZACIÓN EN EL PROCESO PENAL

La fase de investigación tiene por principal objetivo "averiguar y hacer constar la perpetración de los delitos con todas las circunstancias que puedan influir en su calificación y la culpabilidad de los delincuentes, asegurando sus personas y las responsabilidades pecuniarias de los mismos" (art. 299 LECrim). Para ello, las autoridades policiales y judiciales son las encargadas de localizar y obtener la mayor cantidad posible de información ―y salvaguardar las fuentes de prueba que albergan tal información― sobre los hechos objeto de investigación y sus posibles autores. Por su parte, la fase de enjuiciamiento tiene por finalidad practicar, sobre tales fuentes de prueba ―los recipientes de esa información adquirida durante la investigación―, aquellos medios de prueba útiles y pertinentes para lograr la convicción del órgano judicial sobre los hechos afirmados por la acusación y desvirtuar así la presunción de inocencia.

El uso masivo e ininterrumpido de dispositivos electrónicos y de sistemas de la información y la comunicación, no sólo arroja un mayor volumen de información sobre la cual aplicar el *Big Data* y la IA para extraer aquélla relevante para la causa; también arroja una mejor y más precisa tipología de información sobre la que discernir, escoger y visibilizar aquel específico dato demostrativo de la autoría o de la conducta atribuidas al sujeto investigado. De igual modo que la medicina forense permitió en su momento a las autoridades precisar el veneno utilizado o identificar el ADN de las muestras halladas en el escenario del crimen o sobre el cuerpo de la víctima, la informática forense capacita a las autoridades a concretar las coordenadas espaciales desde donde se efectuó una concreta comunicación, el dispositivo utilizado para ello, o las distintas modificaciones operadas sobre un concreto archivo informático.

Dicho con otras palabras; como quiera que nos hemos convertido en la versión moderna de Hansel y Gretel[16] y constantemente vamos dejando *migas de pan* en múltiples formatos digitales, a partir de las diversas interacciones que llevamos a cabo

[15] Véanse las actas del Seminario "Proceedings of the workshop on Use of big data and AI in fighting corruption and misuse of public funds - good practice, ways forward and how to integrate new technology into contemporary control framework", desarrollado el 23 de febrero de 2021 por la Dirección General para las Políticas internas de la UE. Accesibles en: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/691722/IPOL_STU(2021)691722_EN.pdf.

[16] ORTIZ PRADILLO, J. C. (2017) "Desafíos legales de las diligencias de investigación tecnológica", *El Proceso Penal. Cuestiones fundamentales*, (Coord. Olga Fuentes Soriano), Valencia: Tirant lo Blanch, pp. 303-316.

(comunicaciones a través de nuestros teléfonos móviles y otros dispositivos, navegación web, uso de redes sociales, compras online, tránsito por lugares videovigilados, etc.), las autoridades policiales tratarán de localizar, recopilar y singularizar aquella concreta modalidad de información digital que interese a los efectos de la investigación: el dato de localización GPS en el momento de los hechos, la marca y modelo del dispositivo utilizado en la elaboración del video, o el *hash* identificativo del archivo ilícito enviado.

Así, en el examen forense de los diversos dispositivos electrónicos que pudieran pertenecer a la víctima, a terceros o al presunto autor de los hechos, en busca de información trascendental para la averiguación de los hechos, la información decisiva a la hora de condenar o no al sujeto investigado no sólo estará compuesta por los "datos" localizados en la memoria de tales dispositivos o sus elementos periféricos (imágenes, videos, registro de llamadas, chats, mails, etc), sino también por los datos de esos datos; sus «metadatos».

El metadato se define en el Anexo del Real Decreto 4/2010, de 8 de enero, por el que se regula el Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica, como "Dato que define y describe otros datos". Y con mayor precisión, el artículo 42 del Real Decreto 1671/2009, de 6 de noviembre, que desarrollaba parcialmente la Ley 11/2007, de 22 de junio, de acceso electrónico de los ciudadanos a los servicios públicos (ya derogado) disponía que *Se entiende como metadato, a los efectos de este real decreto, cualquier tipo de información en forma electrónica asociada a los documentos electrónicos, de carácter instrumental e independiente de su contenido, destinada al conocimiento inmediato y automatizable de alguna de sus características, con la finalidad de garantizar la disponibilidad, el acceso, la conservación y la interoperabilidad del propio documento*.

Los metadatos pueden ser entendidos, por tanto, como aquellos datos que suministran información sobre el contenido, tipología o formato, origen, soporte, actividad e historial del archivo. También han sido gráficamente descritos como «datos de actividad»: un grupo de etiquetas y marcadores que describen "todos los registros de todas las cosas que haces en tus dispositivos y todas las cosas que tus dispositivos hacen por su cuenta[17]". Cuando enviamos un correo electrónico, los datos de tráfico arrojan información sobre la fecha y la hora de la comunicación y la dirección IP; un análisis más pormenorizado de los metadatos y las cabeceras TCP permiten identificar el servidor utilizado, y en último término, el equipo que conectó con dicho punto de encuentro. El historial de ubicaciones de un teléfono móvil permite conocer el domicilio concreto de un sujeto (ej., dónde pasa las horas de madrugada sin actividad concreta) o su lugar de trabajo. Los metadatos de los archivos enviados por Whatsapp permiten conocer qué dispositivos modificaron tales archivos. Y el historial, duración y potencia de las conexiones a puntos de acceso Wi-Fi en centros comerciales permite saber cuánto tiempo permanecemos en cada uno de los establecimientos de dichos centros, y con ello, si nos gusta más la tienda de deportes, la tienda de animales, o la zona de restauración.

Por todo ello, los procesos judiciales "se han llenado de conversaciones de chat, audios, fotografías y todo tipo de grabaciones, (…) hoy al alcance de cualquiera a través de algo todavía más suave que un click[18]" y una de las pericias cada vez más habituales

[17] SNOWDEN, E. (2019) *Vigilancia permanente*. Esther Cruz Santaella (Trad.), Barcelona: Editorial Planeta, p. 195.

[18] NIEVA FENOLL, J. (2022) "El tránsito de la fe a la tecnología en el proceso penal", *Diario La Ley*, núm. 9986, Sección Tribuna, de 11 de enero de 2022.

en la investigación criminal es el informe forense de los metadatos extraídos de esos audios, fotografías y chats.

1. METADATOS EXTRAÍDOS DE LOS INSTRUMENTOS DE CONVICCIÓN

Como decimos, el análisis forense de imágenes y videos digitales localizados en los dispositivos informáticos aprehendidos se ha convertido en una prueba clave de la comisión de múltiples conductas delictivas.

Durante el proceso de captura y fabricación de la imagen, los dispositivos introducen una enorme cantidad de información ―metadatos― en dicha imagen: características técnicas de la imagen, fecha y hora de generación, localización GPS, presencia o ausencia de flash, distancia de los objetos, tiempo de exposición, apertura del obturador, o la marca y modelo dispositivo utilizado, por poner algunos ejemplos[19]. El empleo de específicas herramientas de extracción y tratamiento de los metadatos de imágenes y videos (por ej., *ExifTools, Gspot* o *MediaInfo*), pueden arrojar muchísima información técnica (tamaño del archivo, versión, fecha de creación y modificación, códecs de audio, número de fotogramas por segundo, sistema operativo utilizado para su captura, idioma, o georreferenciación del lugar donde fue creado) que, tras su debida incorporación al proceso como medio de prueba, permitirá atribuir a un sujeto, sin género de dudas, una determinada conducta.

En otras ocasiones, el análisis se dirige a localizar y analizar toda aquella información identificativa del dispositivo empleado, y en su caso, de su usuario, con el fin de probar su participación en los hechos objeto de investigación. Existen diversas herramientas, calificadas como «fingerprinting», que sirven para extraer información acerca del dispositivo informático utilizado en la navegación web, con el objetivo de identificarlo y singularizarlo. Las empresas utilizan habitualmente dichas técnicas con el fin de identificar y rastrear a los usuarios que visitan sus páginas web o se conecten a sus servidores, para así poder construir un perfil de los mismos a los que ofrecer sus productos y servicios de una manera más personalizada, pues dado que es común que las personas no compartan sus equipos (menos aún, sus teléfonos móviles), individualizar el dispositivo supone individualizar a la persona que lo utiliza. Cualquiera que haya buscado en Google información sobre un determinado producto o le haya dado "me gusta" en Facebook a una específica noticia, comprenderá ahora por qué, al teclear cualquier dirección URL de una página web, es muy posible que el banner o barra de anuncios de dicha web le arroje información sobre ese producto o sobre productos y servicios relacionados con la noticia anteriormente marcada.

A través de dichas técnicas de *fingerprinting*, y aunque el usuario elimine las cookies o utilice VPN u otros servicios de anonimización, se puede llegar a recoger información muy concreta del dispositivo, como por ejemplo el sistema operativo utilizado, el modelo y versión del navegador, resolución de la pantalla, arquitectura de procesador, listas de fuentes de texto, plugins o dispositivos instalados, direcciones IP, el

[19] ESTEBAN COBO, M. (2016) Herramienta para la extracción automática de metadatos en vídeos de dispositivos móviles. [Trabajo Fin de Grado]. https://eprints.ucm.es/id/eprint/38697/.

idioma configurado en el sistema, o el huso horario[20]. La combinación apropiada de toda esta información permite confeccionar una suerte de *huella digital única* del dispositivo que lo singulariza y, por lo tanto, diferencia de forma univoca a cada usuario en internet[21].

Con similar objetivo, la pericial forense tratará de localizar, asegurar y extraer toda aquella información que permita determinar qué dispositivo fue utilizado en una determinada comunicación o interacción a través de la red (datos de sensores del dispositivo, puntos de acceso Wi-Fi utilizados, dispositivos bluetooth activados, etc.) y cuál fue su usuario (por ej., datos del usuario insertados en la aplicación utilizada, datos del titular del correo electrónico introducidos en dicha aplicación, etc.).

1.1. Los metadatos como elementos probatorios de la autoría delictiva

El análisis forense de los metadatos puede resultar trascendental en el enjuiciamiento de delitos relacionados con la libertad e indemnidad sexual o con la explotación y corrupción de menores. El informe pericial de los agentes policiales sobre los metadatos extraídos de los archivos localizados en los dispositivos, instrumentos y materiales incautados en el domicilio del investigado o de las comunicaciones efectuadas o recibidas por el equipo informático igualmente aprehendido en dicho domicilio permite determinar, entre otra información, las cámaras fotográficas o videográficas empleadas para la elaboración del material ilícito, el lugar en donde tales archivos fueron creados, los dispositivos electrónicos utilizados para su almacenamiento y envío a través de la red, los *nicks*, perfiles de redes sociales y de correos electrónicos utilizados y las páginas web y foros empleados para su difusión[22].

Pero el análisis forense de los metadatos sirve para la averiguación de cualquier delito. El *computer forensics* cobra especial valor en la investigación de la delincuencia económica y contribuye a esclarecer delitos contra la Administración Pública y defraudaciones económicas consistentes en la falsificación de documentos públicos y privados, la malversación de caudales públicos, tráfico de influencias, prevaricación administrativa y cohechos activos y pasivos. Como ejemplo, en una de las diversas piezas de la denominada «Trama Gürtel», los tribunales fundamentaron su convicción en los metadatos de los archivos intercambiados a través de correos electrónicos entre los integrantes del grupo criminal y diversos empleados de la Agencia Valenciana de Turismo. A partir de dicha información, se pudo probar que la trama se valía de información privilegiada sobre los criterios de adjudicación para alterar y modificar los informes de condiciones técnicas y manipular los procedimientos para la contratación pública, así como para la creación y alteración de las facturas emitidas, asignación de estas a actividades no realizadas o referidas a los sobornos efectuados, y también sirvieron para atribuir su autoría a los acusados[23]. Así, por ejemplo, uno de los informes de la

---

[20] AGUILERA DÍAZ, V., SEISDEDOS, C. (2020) *Open Source INTelligence (OSINT): Investigar personas e Identidades en Internet*. Madrid: Oxword, p. 49.

[21] Véase el documento *Fingerprinting o Huella digital del dispositivo*, publicado por la AEPD en 2019. Disponible en: https://www.aepd.es/sites/default/files/2019-09/estudio-fingerprinting-huella-digital.pdf.

[22] Una explicación detallada de los metadatos encontrados en una operación contra la explotación sexual infantil y la corrupción de menores se aprecia en la STS núm. 395/2021, de 6 de mayo. ECLI:ES:TS:2021:1737.

[23] STS núm. 214/2018, de 8 de mayo de 2018. ECLI:ES:TS:2018:1551. Los metadatos de los correos interceptados, así como de las facturas manipuladas también fueron objeto de examen en otra de las piezas

Unidad de Delincuencia Económica y Fiscal (UDEF) de la Policía Nacional acreditó que las facturas emitidas por la empresa Orange Market, si bien tenían fechas diferentes de varios meses entre ellas, habían sido elaboradas consecutivamente y en solo tres minutos: entre las 18:46 y las 18:49 del 17 de julio de 2007, y ello fue probado gracias a los metadatos extraídos a los archivos Excel de dichas facturas[24].

1.2. Los metadatos como elementos probatorios de la inocencia del acusado

Pero lo más trascendental del análisis forense de los datos y metadatos localizables en las fuentes de prueba reside en que, como evidencias científicas, están llamadas a desterrar prejuicios y ser debidamente valorados por encima de otros medios de prueba más *tradicionales*, pero menos científicos, como pueda ser el ancestral valor probatorio de la confesión del reo, precedida del habitual y ritual juramento, o de los testimonios prestados por terceros[25].

La Audiencia Provincial de Sevilla condenó a Moisés (nombre ficticio) a 6 años de prisión como autor de un delito de lesiones agravadas de los artículos 147.1 y 149.1 del Código Penal, al estimar probado que, entre las 1:30 y las 5:00 horas de la madrugada del día 15 de diciembre de 2013, estando en las inmediaciones de la puerta de una discoteca, propinó puñetazos y patadas en la cara y cabeza de su víctima. Como pruebas incriminatorias, tuvo en consideración el reconocimiento del acusado efectuado por el lesionado, quien ante la policía reconoció fotográficamente al acusado como su agresor, lo reconoció después durante el sumario en rueda de identificación practicada en sede judicial, y lo reconoció por última vez en el propio acto del plenario, así como la declaración de los amigos que le acompañaban en el momento en que fue objeto de la agresión, quienes sustentaron los mismos reconocimientos sin ninguna duda, y ciertas declaraciones espontáneas del acusado cuando fue detenido.

Aunque el testimonio de la víctima, corroborado por otras declaraciones testificales, es prueba valorable y suficiente para desvirtuar la presunción de inocencia, es preciso recordar que la existencia de otros elementos probatorios contradictorios (declaraciones de descargo efectuadas por el acusado, igualmente corroboradas por otros testigos aportados por la defensa) exigen un especial celo —junto con una necesaria motivación— por parte del tribunal juzgador a la hora de valorar todo el bagaje probatorio, y especialmente, todo el material probatorio periférico que ambas partes aportan precisamente para facilitar que el juzgador pueda testar la verosimilitud de los diferentes testimonios. Y es, entonces, cuando la *prueba electrónica* puede jugar un papel decisivo a la hora de condenar o absolver al acusado.

En los hechos enjuiciados por la Audiencia de Sevilla, ese otro material probatorio periférico y de naturaleza digital fue el siguiente:

En primer lugar, una de las cámaras de videovigilancia ubicadas en la calle en la que se encuentra la discoteca donde se produjo la agresión recogió unas imágenes en las

---

de dicha trama, también corroboradas por el Tribunal Supremo en su STS núm. 86/2022, de 31 de enero. ECLI:ES:TS:2022:396.

[24] Véase la noticia publicada en http://www.elladodelmal.com/2012/06/la-financiacion-ilegal-de-francisco.html.

[25] En el mismo sentido, vid. NIEVA FENOLL, J. (2022) "El tránsito de la fe…", op. Cit.

que el atacante no identificado regresó, un momento posterior a la agresión, para increpar al lesionado. Dicha imagen fue datada a las 5:37 horas, (y, como veremos, la concreción del momento exacto de la agresión resultará determinante).

En segundo lugar, y en apoyo de la versión de descargo del acusado, quien sostuvo que participó en una cena navideña con unos amigos en otro restaurante, se aportaron fotografías y conversaciones de WhatsApp datadas entre la 1:10 y la 1:54 de la madrugada). Junto con la declaración de los amigos del acusado que afirmaron haberse dirigido a otro pub de la zona aproximadamente hasta las 3:00, se aportó una fotografía tomada por el teléfono de la novia del acusado que, según la certificación notarial, el contenido de los metadatos del archivo fotográfico refleja que la fotografía fue tomada a las 3:58 horas de esa madrugada. Otras fotografías capturadas en el exterior del pub estarían tomadas a las 4:47 horas.

Y en tercer lugar, compareció en el plenario el taxista que llevó al acusado y a otro acompañante a su localidad de residencia. Dicho testigo relató que iniciaron el trayecto sobre las 5:50 o 6:00 de la mañana y que el mismo duró 30 o 35 minutos aproximadamente, lo cual vino corroborado por el extracto bancario de la tarjeta de crédito con el que el acusado abonó dicho servicio, registrado a las 6:29 de esa madrugada.

Con toda esta información, el Tribunal Supremo estimó el recurso de casación por vulneración del derecho a la presunción de inocencia formulado por Moisés, tras valorar que "la credibilidad que la sentencia de instancia ha conferido al reconocimiento que los testigos de cargo hicieron del acusado como el autor de los hechos, se enfrenta a un conjunto de testimonios y *vestigios objetivos que lo desmienten*[26]".

2. METADATOS OBTENIDOS DE FUENTES ABIERTAS

2.1. Internet y las Redes Sociales como "lugar de comisión del delito" y como lugar de recolección de los vestigios y las fuentes de prueba

La reformulación de la noción del "escenario del crimen" propiciada por el avance tecnológico tiene, entre otras consecuencias, su traslación a la noción del "lugar del delito", que no sólo debe referirse a un espacio físico y geográfico. *El delito en su forma más convencional convive ahora con nuevas formas de ciberdelincuencia en las que su ejecución se desarrolla enteramente en redes telemáticas que, por definición, no son inmovilizables en un espacio físico perfectamente definible. El ciberespacio ofrece un marco digital diferenciado de la realidad puramente física como espacio del delito. La experiencia más reciente enseña que las redes sociales no son sólo el instrumento para la comisión de algunos delitos de muy distinta naturaleza. Pueden ser también el escenario en el que el delito se comete, ya sea durante todo su desarrollo, ya en la ejecución de sólo algunos de los elementos del tipo*[27].

Por tanto, aunque no resulte posible acceder al dispositivo empleado para la comisión delictiva, ello no impide que las autoridades busquen, localicen y examinen

[26] STS núm. 120/2019, de 6 de marzo. ECLI:ES:TS:2019:737.

[27] STS núm. 547/2022, de 2 de junio. ECLI:ES:TS:2022:2356.

datos y metadatos obtenidos de fuentes abiertas o de la propia Internet. De hecho, el ciberpatrullaje o navegación web de los agentes de policía ha sido considerada una actividad legítima que no requiere de una autorización judicial previa, de modo que la recopilación de aquella información que la policía adquiere con motivo de dicha navegación no distaría de aquella información referente a los instrumentos o vestigios del delito que la policía localiza y aprehende con motivo de su actividad investigadora.

Un ejemplo concreto lo encontramos en la metodología empleada por el Grupo de Investigación y Análisis de la Agrupación de Tráfico de la Guardia Civil (GIAT). Este departamento, con agentes especializados en cada uno de los sectores de Tráfico de cada provincia, tiene por misión la investigación y persecución de los delitos relacionados con la seguridad vial, como puedan ser excesos de velocidad, conducción temeraria, o la falsificación de la documentación referida a vehículos y conductores.

Cuando se sube y difunde a través de las redes sociales un video ilustrativo de la comisión de un posible delito contra la seguridad vial, el GIAT analiza la información documentada en el video de cara a la posible identificación de la vía o tramo de carretera donde se han producido los hechos, y a partir de ahí, tratar de identificar a sus autores (identificación de los vehículos y de los rostros que aparecen en el vídeo) mediante su cotejo con los datos en poder de la DGT (ej., los datos de los titulares de los vehículos o las imágenes de las cámaras de tráfico que, en otros puntos kilométricos, sí identifiquen las matrículas de los vehículos que aparecen en el video). Pero también se sirven de la información que aporta el propio hecho de subir el vídeo a la red social; los metadatos extraíbles del vídeo y del perfil del usuario que lo colgó o difundió.

Tras la detención de los integrantes de *La Manada*, por la violación cometida el 7 de julio de 2016 en Pamplona, la Policía Foral de Navarra analizó los videos e imágenes contenidos en el teléfono móvil de uno de los acusados y se encontraron evidencias de otro delito; se encontraron grabaciones efectuadas entre el 30 de abril y el 1 de mayo de 2016 que, aparentemente, reflejaban un abuso sexual cometido sobre otra víctima en el asiento posterior de un vehículo. El rastreo y análisis de la información publicada en las redes sociales permitió a las autoridades ubicar el lugar de los hechos (Pozoblanco, Córdoba) y determinar la identidad de la víctima.

En concreto, la policía llegó hasta el perfil de Facebook de una discoteca de una localidad que había celebrado sus fiestas locales en dichas fechas, y en donde se habían subido centenares de fotografías de aquellos días. El rastreo de los *likes* dirigió a la policía hasta un perfil de una joven cuya fotografía era similar a la que aparecía en el vídeo. En otras fotografías publicadas por dicha joven en su perfil de esa red social se constató que coincidían el vestido, el reloj o ciertos rasgos que aparecían en el vídeo de la agresión. Gracias a la colaboración de la policía local de Pozoblanco, se localizó a la víctima.

Los integrantes de esa jauría humana, condenados por la agresión cometida en Pamplona, fueron finalmente condenados también como autores de un delito de abusos sexuales y otro contra la intimidad (por la grabación del vídeo) por los hechos cometidos en Pozoblanco[28]. Y ello, gracias en gran parte a la investigación sobre fuentes abiertas y redes sociales.

2.2. Técnicas e instrumentos de recolección de los metadatos de la información transmitida a través de Internet y las Redes Sociales

[28] Sentencia del Juzgado de lo Penal número 1 de Córdoba de 14 de abril de 2020, confirmada por la SAP Córdoba, secc. 2ª, núm. 306/2020, de 11 de noviembre. ECLI:ES:APCO:2020:1292.

Como acabamos de comprobar, la recopilación y análisis de esa información (y los metadatos asociados a la misma) extraída de internet se lleva a cabo a través del empleo de las herramientas de inteligencia sobre fuentes abiertas o de acceso público (OSINT, por sus siglas en inglés *Open Source Intelligence*), lo cual amplía exponencialmente las capacidades policiales de localizar información relevante para la prevención y persecución de actividades delictivas, así como identificar y georreferenciar a un determinado sujeto en función de su actividad en Internet.

Más allá de atribuir a un usuario específico el empleo de una concreta dirección IP, también resulta conveniente tomar conocimiento de otros metadatos producidos con motivo de dicha comunicación web, como pueda ser la geolocalización o los datos del terminal empleado por un concreto perfil o del nick utilizado en las redes sociales.

Por señalar un ejemplo concreto, y con el objetivo de identificar y localizar desde dónde *tuitea* un sujeto que puede estar cometiendo a través de Twitter un posible delito de incitación al odio, discriminación o violencia contra una persona conforme al art. 510 CP, existen herramientas como *Cree.py* para geolocalizar a los usuarios de servicios web, como Twitter y Flickr, a partir de los metadatos de los *tuits* o fotos publicadas (información GPS del dispositivo, tuits enviados desde una determinada ubicación, triangulación basada en la IP desde la que se hizo el tuit, etc). De igual modo, y en aras a fundamentar una condena penal por delitos de captación y adoctrinamiento terrorista y de enaltecimiento del terrorismo, resulta en ocasiones esencial probar la autoría de un determinado perfil desde el cual se escriben posts y se da "me gusta" a vídeos e imágenes de extremismo yihadista[29].

Junto con dichos metadatos, también resultará relevante el análisis inteligente de los contenidos transmitidos en abierto, como puedan ser las palabras utilizadas, las imágenes o videos intercambiados, o las personas mencionadas en dichas interacciones cibernéticas. Para ello, existe una categoría especial de herramientas y técnicas ideadas para la extracción de información de las principales redes sociales denominadas «*Social Media Intelligence* (SOCMINT)» o Inteligencia basada en las Redes Sociales.

Tales herramientas aprovechan la información generada por el *Big Data* para localizar aquella información generada o transmitida a través de la Red que pueda resultar útil de cara a las legítimas funciones de las autoridades estatales.

Al igual que las empresas son las primeras en utilizar dichas técnicas con propósitos comerciales como, por ejemplo, la detección de tendencias o el análisis y mejora de su reputación online, a partir del examen y la categorización del perfil de sus seguidores (edad, sexo, ubicación geográfica,…), los gobiernos, los servicios secretos y los investigadores policiales también monitorizan las Redes Sociales para, bien prevenir la comisión de posibles ilícitos, o bien obtener una valiosa información sobre hechos e individuos que sean objeto de una concreta investigación judicial, pues un examen profesional del perfil de un usuario, sus interacciones y los metadatos generados con aquéllas (contactos, términos empleados, ficheros intercambiados, geolocalización de las interacciones o imágenes difundidas, información del equipo empleado, identificación del correo electrónico asociado a la cuenta, etc.) pueden facilitar enormemente la labor policial de proceder a su identificación, su localización o la determinación de su pertenencia a un determinado grupo criminal.

La «Operación Araña» llevada a cabo por la Guardia Civil durante los años 2014, 2015 y 2016 contra supuestos delitos de enaltecimiento del terrorismo y humillación a las víctimas cometidos a través de las redes sociales Facebook y Twitter, y que se saldó con

[29] Un extenso examen de la prueba consistente en los "me gusta" de un específico perfil de Facebook puede leerse en la SAN (Sala de lo Penal, secc. 4ª) núm. 16/2021, de 23 de septiembre. ECLI:ES:AN:2021:4196.

más de 70 detenciones en toda la geografía española[30], consistió en la monitorización por parte de las autoridades policiales de dichas redes sociales mediante técnicas de fuentes abiertas para la localización de contenidos y publicaciones y para la identificación de los responsables de dichas publicaciones.

Para ello, la determinación de la dirección IP utilizada en una concreta interacción a través de Internet suele ser el punto de partida de muchas investigaciones. Y el Tribunal Supremo ha reiterado que la obtención policial de una dirección IP con motivo del ciberpatrullaje en la red es "información de conocimiento público para cualquier usuario de Internet y que el propio usuario de la red es quien lo ha introducido en la misma[31]".

La geolocalización de una dirección IP resulta relativamente sencilla. La ICANN (Corporación de Internet para la Asignación de Nombres y Números) es la entidad que gestiona las direcciones IP a nivel mundial y las reparte a los Registradores Regionales de Internet (RIR). El Registrador Regional que gestiona las direcciones en Europa es el RIPE (Network Coordination Centre - RIPE NCC), quien a su vez procede a una posterior subdelegación de recursos a sus clientes, los Registros de Internet Local (Local Internet Registries, LIRs[32]), quienes distribuyen el direccionamiento IP en cada una de esas regiones a determinadas operadoras (ej., Telefónica), y éstas finalmente son las que las ceden a los proveedores de acceso. Por lo tanto, y en función del rango de dirección IP, es fácil determinar inicialmente si una determinada dirección ha sido asignada a la operadora de un determinado país y presumir que el usuario de dicha IP se halla dicho país, y dentro del mismo, qué nodo utiliza para acceder a Internet.

A partir de ahí, cuando las fuerzas y cuerpos de seguridad de cualquier país tienen conocimiento de que una específica dirección IP ha sido empleada para cometer una específica conducta delictiva (v. gr., enviar o recibir material de explotación sexual infantil), comunican dicha información a las autoridades policiales del país al cual se tiene asignada dicha dirección IP, quienes se ocuparán ―en el caso español― de solicitar el oportuno mandamiento judicial para recabar de las operadoras los datos del titular de la conexión (los datos que permitan la identificación y localización del terminal o del dispositivo de conectividad y la identificación del sospechoso, en palabras del artículo 588 ter k LECrim).

2.3. Colaboración público-privada en el rastreo de metadatos de información ilícita transmitida a través de Internet y las Redes Sociales

Un ejemplo concreto de colaboración entre autoridades estatales, el sector privado y organizaciones no gubernamentales, a la hora de detectar posibles conductas delictivas en Internet, es el referido a la lucha contra los abusos sexuales y la explotación sexual de los menores, incluida la pornografía infantil.

La necesidad de adopción de medidas eficaces contra dichas actividades se recoge en múltiples textos legislativos, tales como la Convención de las Naciones Unidas sobre

[30] Véase la nota de prensa difundida el 13 de abril de 2016 por el Ministerio del Interior. Noticia disponible en la dirección URL: http://www.interior.gob.es/prensa/noticias/-/asset_publisher/GHU8Ap6ztgsg/content/id/5847566.

[31] SSTS núm. 1058/2006 de 2 noviembre, 1299/2011 de 17 de noviembre, 342/2013 de 17 de abril, 1025/2013 de 26 diciembre y 16/2014, de 30 de enero.

[32] El listado de esos Registros de Internet Local que ofrecen servicios en España puede ser consultado en la página web: https://www.ripe.net/membership/indices/ES.html.

los Derechos del Niño de 1989, el Protocolo facultativo de la Convención de las Naciones Unidas sobre los Derechos del Niño de 2000 relativo a la venta de menores, la prostitución infantil y la utilización de los menores en la pornografía, el Convenio del Consejo de Europa de 2007 sobre la protección de los menores contra los abusos sexuales y la explotación sexual, o la Directiva 2011/93/UE del Parlamento Europeo y del Consejo, de 13 de diciembre de 2011, relativa a la lucha contra los abusos sexuales y la explotación sexual de los menores y la pornografía infantil y por la que se sustituye la Decisión marco 2004/68/JAI del Consejo. Y el rastreo de metadatos de las comunicaciones realizadas a través de Internet se ha convertido en una de las grandes iniciativas que ha logrado aunar los esfuerzos del sector público y privado para combatir eficazmente la elaboración y la difusión de imágenes y vídeos de explotación sexual infantil a través de Internet.

El Centro Nacional para Menores Desaparecidos y Explotados (NCMEC - *National Center for Missing & Exploited Children*), fundado en los EE.UU. en 1984, lleva años alertando a las autoridades de que determinados usuarios han compartido fotos o videos con posible contenido pedófilo a través de las redes sociales, para lo cual se sirve del empleo de técnicas y herramientas OSINT y SOCMINT para el rastreo, identificación y localización de dichos archivos tales como *PhotoDNA* o *NeuralHash*.

PhotoDNA es una tecnología de identificación de imágenes desarrollada por Microsoft, en colaboración con el Dartmouth College en 2009, a partir de la creación de códigos hash únicos que sirven para prepresentar esa imagen. Este hash se calcula de tal manera que es resistente a las posibles alteraciones que se pueden realizar sobre dicha imagen (ej., el cambio de tamaño y formato del archivo). Posteriormente, dicho hash se compara con las bases de datos que las organizaciones y empresas participantes en esta iniciativa ya han identificado como ilícitas por su contenido de explotación sexual infantil (como el NCMEC o los integrantes del Proyecto VIC) y permite a las empresas prestadoras de servicios en internet denunciar a los usuarios que han difundido dichos archivos a través de sus plataformas.

Se utiliza en los propios servicios de Microsoft, incluidos Bing y OneDrive, pero también en otros prestadores de servicios tales como Facebook, Twitter, Adobe, así como en Gmail de Google, mientras que Apple ha decidido desarrollar una herramienta similar, llamada *NeuralHash*, para escanear las cuentas iCloud de sus usuarios en busca y captura de contenido pedófilo. De igual modo, Youtube también cuenta con sistemas automatizados de detección de videos con imágenes de explotación sexual infantil o de extremismo terrorista, gracias a la tecnología *fingerprinting* y a la detección y contraste de los *hashes* de los archivos subidos a dicha red. Desde junio de 2017, implementó tecnologías de aprendizaje automático para la detección de contenido extremista violento y de discurso de odio, basadas en la comparativa de los datos de videos ya revisados por decisiones humanas y eliminados de sus servidores[33].

Con ese mismo espíritu de servirse del escaneo automatizado de los metadatos para evitar la difusión de material contenedor de imágenes de explotación sexual infantil, en julio de 2021 se aprobó el Reglamento (UE) 2021/1232 del Parlamento Europeo y del Consejo de 14 de julio de 2021 por el que se establece una excepción temporal a determinadas disposiciones de la Directiva 2002/58/CE en lo que respecta al uso de tecnologías por proveedores de servicios de comunicaciones interpersonales

[33] Más información en: https://protectingchildren.google/#fighting-abuse-on-our-own-platform-and-services.

independientes de la numeración para el tratamiento de datos personales y de otro tipo con fines de lucha contra los abusos sexuales de menores en línea[34].

Dicho Reglamento establece una excepción temporal de tres años (hasta el 3 de agosto de 2024) a la protección dispensada por los artículos 5, apartado 1, y al artículo 6, apartado 1, de la Directiva 2002/58/CE, en materia de confidencialidad de las comunicaciones y los datos de tráfico, así como un marco especial en materia de los derechos de acceso, rectificación, etc., de los datos personales en virtud del Reglamento (UE) 2016/679, para que los proveedores de servicios en la Sociedad de la Información puedan utilizar, con respaldo legal, tecnologías específicas de contraste de *hashes* para imágenes y vídeos, y herramientas de IA para el análisis de datos de tráfico, con el único fin de detectar y retirar el material de abuso sexual de menores en línea, o el «embaucamiento de menores» con fines sexuales a través de medios tecnológicos, y denunciarlo a las autoridades policiales y a las organizaciones que actúen en interés público contra los abusos sexuales de menores.

El artículo 3 impone ahora a tales proveedores el deber de conservar de manera segura los datos de contenido y los datos de tráfico conexos, así como los datos personales generados mediante dicho tratamiento, cuando se haya detectado un archivo de un presunto abuso sexual a menores en línea, crear una firma digital única y no reconvertible («hash») de datos identificados de forma fiable como material de abuso sexual de menores en línea, bloquear la cuenta del usuario de que se trata, o de suspender o poner fin a la prestación del servicio a dicho usuario y denunciar sin demora el presunto abuso sexual de menores en línea a las autoridades policiales y judiciales competentes o a organizaciones que actúen en interés público contra los abusos sexuales de menores.

No obstante, también les exige que articulen mecanismos para permitir al usuario de que se trate recurrir ante el proveedor o interponer recursos administrativos o judiciales en asuntos relacionados con el presunto abuso sexual de menores en línea, y remitir a la Comisión Europea y a la Autoridad de control del Reglamento (UE) 2016/679 informes anuales[35] sobre el número de casos detectados de abusos sexuales de menores en línea, diferenciando entre material de abuso sexual de menores en línea y embaucamiento de menores, el número de casos en los que un usuario haya presentado una reclamación ante el mecanismo de recurso interno o interpuesto una acción judicial y el resultado de dichas reclamaciones y causas judiciales, el número y las tasas de errores (falsos positivos) de las diferentes tecnologías utilizadas, así como las medidas aplicadas para limitar la tasa de error y la tasa de error alcanzada, la política de conservación de los datos y las

[34] DOUE L 274, de 30 de julio de 2021.

[35] A falta de información sobre los datos del año 2022, contamos con datos para el periodo comprendido entre la entrada en vigor del Reglamento, el 3 de agosto de 2021, y el 31 de diciembre de 2021. El Informe de Google para Europa (Google Ireland Transparency Report Under Regulation (EU) 2021/1232, accesible en: https://transparencyreport.google.com/report-downloads?hl=en) indica que detectó automáticamente 331 cuentas de usuarios de la UE con material referido a abuso sexual infantil. De ellos, en 8 casos el usuario apeló internamente la decisión de Google de cancelar su cuenta y en ninguno de esos 8 casos se modificó la decisión inicial de cancelación. A nivel mundial (datos extraídos de https://transparencyreport.google.com/child-sexual-abuse-material/reporting), Google informa que detectaron un total de 3.282.824 de archivos con material de abuso sexual infantil en sus plataformas; inhabilitó 140.868 cuentas de usuarios y elaboraron un total de 1.620.122 hashes que añadieron a sus repositorios internos, además de compartirlos con el NCMEC para que puedan consultarlos otros proveedores. Twitter, por su parte (Twitter Report Under Article 3(1)(g)(vii) of Regulation (EU) 2021/1232, accesible en: https://blog.twitter.com/en_us/topics/company/2022/twitter-s-eu-submission-for-2021-1232) informó que había suspendido un total de 532.898 cuentas por violar su política de lucha contra la explotación sexual infantil.

salvaguardias de protección de datos aplicadas conforme al Reglamento (UE) 2016/679, y el nombre de las organizaciones que actúan en interés público contra los abusos sexuales de menores con las que se hayan compartido datos en virtud del presente Reglamento.

2.4. Rastreo de metadatos de la información transmitida a través de Internet y las Redes Sociales y Derecho Fundamental al secreto de la correspondencia

La conceptualización de los metadatos como "toda aquella información descriptiva sobre el contexto, calidad, condición o características de un recurso, dato u objeto que tiene la finalidad de facilitar su recuperación, autentificación, evaluación, preservación y/o interoperabilidad" conduciría a considerarlos igualmente como *auténticos contenidos de comunicaciones, por mucho que en no pocas ocasiones se encuentren alojados en cabeceras IP o DNS*[36]. El escrutinio de los metadatos de los contenidos compartidos o transmitidos por estas vías afecta al Derecho Fundamental a la privacidad del artículo 7 de la Carta de Derechos Fundamentales de la Unión Europea, y el consiguiente tratamiento (conservación y notificación a las autoridades) de la información obtenida y su posterior contraste con bases de datos de hashes sospechosos de contener pornografía infantil afecta igualmente al derecho a la protección de datos personales del art. 8[37].

Quizás por ello, el citado Reglamento (UE) 2021/1232 indica que esas actividades de escanéo de metadatos que llevan a cabo los proveedores para detectar abusos sexuales de menores en línea cometidos en sus servicios y denunciarlos "representan una injerencia en los derechos fundamentales al respeto de la vida privada y familiar y a la protección de los datos personales de todos los usuarios de servicios de comunicaciones interpersonales independientes de la numeración", y de ahí la necesidad de esta expresa disposición legal como excepción a la norma prohibitiva de los arts. 5.1, 6.1 y 15.1 de la Directiva 2002/58/CE y al régimen general del RGPD.

Pero también cabe la posibilidad de defender jurídicamente que dichas técnicas de escaneo de *hashes* y su comparativa automatizada con bases de datos de material ya declarado ilícito no dejan de ser sino una modalidad moderna o traslación al entorno digital de lo que siempre fueron «técnicas de inspección por motivos de seguridad pública u orden público» con el fin de detectar la presencia de productos prohibidos, que legalmente se consideraron siempre una excepción a la inviolabilidad de los envíos postales. El artículo 6.4 de la Ley 43/2010, de 30 de diciembre, del servicio postal universal, de los derechos de los ususarios y del mercado postal (Inviolabilidad de los

[36] Para RODRIGUEZ LAINZ [(2013) "En torno al concepto de comunicación protegida por el art. 18.3 de la Constitución". *Diario La Ley*, núm. 8142, de 5 de septiembre de 2013], "Si la cabecera incluye datos sobre titularidad o dominio, datación, volumen y naturaleza de la información, e identidades de destino y origen, evidentemente entrarán en el campo de los llamados datos de tráfico. Pero si tienen una finalidad ajena al tránsito mismo de la comunicación, cual sucede no solo respecto de metadatos que atienden a su finalidad etimológica de facilitar búsquedas de recursos, sino también aquella información añadida que nada tiene que ver con el contenido o su tráfico (v.g.r. información sobre ubicación geográfica del emitente asociada, siempre opcionalmente, a twitts, fotografías enviadas a través de determinadas aplicaciones tipo Apple o Android, o determinados servicios de valor añadido), deberán ser tenidos a todos los efectos como contenidos; puedan ser o no accesibles a los efectos de prestación del servicio correspondiente por las operadoras de telecomunicaciones".

[37] En el mismo sentido, RODRÍGUEZ LÁINZ, J. L. (2021) "Reflexiones sobre el tratamiento de datos personales por prestadores de servicios de comunicaciones vía internet para la lucha contra abusos sexuales de menores en línea en el Reglamento (UE) 2021/1232", *Diario La Ley*, núm. 9974, de 20 de diciembre de 2021.

envíos postales) ya preveía "el ejercicio de las facultades de control reconocidas legalmente a determinados funcionarios en el marco del ejercicio de sus funciones de inspección, como las sanitarias, aduaneras, de prevención de blanqueo de dinero o de seguridad o cualesquiera otras establecidas en la normativa sectorial, *con el fin de detectar la presencia de productos prohibidos*", y el artículo 12.3 del Real Decreto 1829/1999, de 3 de diciembre, por el que se aprueba el Reglamento por el que se regula la prestación de los servicios postales, en desarrollo de lo establecido en la Ley 24/1998, de 13 de julio, del Servicio Postal Universal y de Liberalización de los Servicios Postales, también advierte que "Ni los funcionarios de la Secretaría General de Comunicaciones del Ministerio de Fomento ni los de los servicios aduaneros podrán conocer el contenido de los envíos postales, debiendo respetar el derecho al secreto, la intimidad y la inviolabilidad de la correspondencia. *Se excepcionan los envíos que no contengan documentos de carácter actual y personal como la publicidad directa, los libros, los catálogos, las publicaciones periódicas, así como los paquetes postales en los que proceda su inspección por motivos de seguridad pública u orden público* y no se haya puesto de manifiesto expresamente que contienen objetos de carácter personal".

De igual modo que la inviolabilidad de la correspondencia no imposibilita que los envíos postales pueden escanearse y someterse a técnicas de detección de materiales tóxicos, explosivos y armas en su interior, también cabe sostener que el avance tecnológico facilita la utilización de herramientas para, de manera automatizada, contrastar factores o elementos (metadatos) indicativos de la posible existencia de un material expresamente prohibido por la legislación, así como por las específicas políticas internas de condiciones y términos de las empresas que prestan tales servicios.

IV. CONCLUSIONES

En lo más profundo de un *byte* puede estar la clave para determinar la culpabilidad o inocencia de un sujeto. Ya no es cuestión de cantidad de prueba incriminatoria; la tecnología ha sustituido la cantidad por la calidad. Si tres palabras del legislador pueden destruir bibliotecas enteras, un *byte* de información puede echar por tierra el testimonio de decenas de testigos.

En la investigación tecnológica del delito, por muy "pequeño" que resulte el dato o metadato del que se tiene conocimiento, la repercusión que ello puede tener sobre la afectación a los derechos fundamentales de los ciudadanos puede ser trascendental. La gran heterogeneidad de información digital creada y compartida a través de la red, así como la multiplicidad de sujetos que intervienen en la recopilación, la conservación, el tratamiento y el intercambio de datos, y de los lugares en donde resulta posible localizar dicha información, aumentan las posibilidades de las autoridades de localizar aquélla que pueda resultar clave en la identificación de un investigado y en la atribución al mismo de la conducta ilícita.

El carácter público o privado del entorno en donde se desenvuelva la conducta humana, así como la noción de lo que realizamos conscientemente en público (*knowing exposure*) deben ser objeto de reconsideración a la hora de valorar la posible existencia de una «expectativa razonable de privacidad», una vez que el *Big Data* permite recopilar a perpetuidad hasta la más mínima interacción humana en cualquier parte del mundo y el *Data Mining* y la IA habilitan a las autoridades a acceder, analizar y segmentar dicha

información en busca de cualquier elemento incriminatorio[38]. La recolección y almacenamiento de toda esa información procesada digitalmente ya no distingue si la misma fue generada con motivo de una interacción realizada en nuestros dormitorios o en mitad de un centro comercial.

Por ello, los criterios ponderativos de la proporcionalidad de la medida dirigida a acceder y recolectar dicha información no pueden ser objeto de examen aislado o desagregado del resto de actuaciones investigadoras. Es lo que, desde distintas instancias judiciales y a partir de distintas argumentaciones, se ha venido a calificar como la «teoría del mosaico» a la hora de examinar si ha habido o no una excesiva afectación sobre los derechos fundamentales del investigado: aunque el dato o información a adquirir, individualmente considerado, pueda considerarse inofensivo o mínimamente invasivo sobre la intimidad o cualquier otro derecho fundamental, es preciso examinar la totalidad de los actos de investigación llevados a cabo por las autoridades para tener una visión completa (mosaico) de la afectación llevada a cabo[39].

En España y en Europa también se ha llegado a la conclusión de que el examen de esas *migas de pan* que, agrupadas, conforman el entorno virtual del individuo, puede generar una importante intrusión en la vida privada de su titular, y que el acopio, análisis y cruce de los datos de tráfico o localización generados por el uso de las comunicaciones electrónicas constituyen una suculenta fuente de información que puede resultar muchísimo más intrusivo en la personalidad de un sujeto que el conocimiento del contenido de una conversación[40].

La protección de los menores, tanto en el mundo físico como en el virtual, debe considerarse una prioridad de cualquier país o gobierno, y las iniciativas antes examinadas en materia de control y retirada de material online que pueda contener imágenes y vídeos de explotación o abuso sexual infantil deben ser aplaudidas. Pero no me cabe la menor duda de que la temporal excepción introducida por el Reglamento (UE) 2021/1232 será objeto de prórroga o consolidación jurídica, así como también de que el específico fin justificativo de la misma —la lucha contra la explotación sexual infantil y el abuso sexual

---

[38] JOH, E. E. (2014) "Policing by Numbers: Big Data and the Fourth Amendment". 89 Wash. L. Rev. 35 (disponible en: https://ssrn.com/abstract=2403028) también cuestiona la actual interpretación judicial de la protección dispensada por la Cuarta Enmienda cuando se produce una reutilización de la información inicialmente recolectada para otros fines (*repurposing information*) por parte de la policía.

[39] KERR, O. S. (2012) "The Mosaic Theory of the Fourth Amendment", 111 *Michigan Law Review*, p. 311.

[40] Sobre la injerencia que puede significar el simple conocimiento de los datos de tráfico, sin acceder al contenido de las comunicaciones, vid. ORTIZ PRADILLO, J. C. (2021) "Big Data…", op. Cit. En España, la STC 173/2011, de 7 de noviembre, cuando el titular de un ordenador personal "navega por Internet, participa en foros de conversación o redes sociales, descarga archivos o documentos, realiza operaciones de comercio electrónico (…), está revelando datos acerca de su personalidad, que pueden afectar al núcleo más profundo de su intimidad, por referirse a ideologías, creencias religiosas, aficiones personales, información sobre la salud, orientaciones sexuales, etc., que (…) si se analizan en su conjunto, no cabe duda que configuran todos ellos un perfil altamente descriptivo de la personalidad de su titular, que es preciso proteger frente a la intromisión de terceros o de los poderes públicos, por cuanto atañen, en definitiva, a la misma peculiaridad o individualidad de la persona". De igual modo, el TJUE (vid. apartados 177 a 184 de la STJUE, Gran Sala, de 6 de octubre de 2020, *La Quadrature du Net*. Asuntos C-511/18, C-512/18 y C-520/18. EU:C:2020:791) ha sentenciado que el acceso a los datos conservados que informen de los lugares en que las comunicaciones tuvieron lugar debe considerarse una «injerencia grave» en la privacidad de las personas que, considerados en su conjunto, permiten extraer conclusiones muy precisas sobre la vida privada de las personas afectadas y proporcionan los medios para determinar el perfil de las mismas, *información tan sensible, a la luz del respeto de la vida privada, como el propio contenido de las comunicaciones*.

infantil en línea— será objeto de ampliación a otras conductas delictivas sumamente graves —v. gr., terrorismo o delincuencia organizada— hasta llegar a ser una medida legal más en materia de prevención, persecución y represión de "delincuencia grave", tal y como ya aconteció con la célebre Directiva 2006/24/CE en materia de conservación de datos.

Habrá que estar vigilantes, por tanto, al devenir legislativo en esta materia.

V. BIBLIOGRAFÍA

AGUILERA DÍAZ, V., SEISDEDOS, C. (2020) *Open Source INTelligence (OSINT): Investigar personas e Identidades en Internet*. Madrid: Oxword.

BALKIN, J. M. (2008) "The Constitution in the National Surveillance State", *Minnesota Law Review 93 (1)*, pp. 1-25.

ESTEBAN COBO, M. (2016) Herramienta para la extracción automática de metadatos en vídeos de dispositivos móviles. [Trabajo Fin de Grado]. https://eprints.ucm.es/id/eprint/38697/.

GINSBERG, J., MOHEBBI, M., PATEL, R. et al. (2009) "Detecting influenza epidemics using search engine query data". *Nature 457*, 1012–1014.

HAYS, C. L. (2004) "What Wal-Mart Knows About Customers' Habits". Noticia publicada el 14 de noviembre de 2004 en el periódico New York Times.

JOH, E. E. (2014) "Policing by Numbers: Big Data and the Fourth Amendment". 89 *Washington Law Review*, 35.

KERR, O. S. (2012) "The Mosaic Theory of the Fourth Amendment", 111 *Michigan Law Review* 311.

LLAMAS FERNÁNDEZ, M., GORDILLO LUQUE, J. M. (2007) "Medios técnicos de vigilancia", *Los nuevos medios de investigación en el proceso penal. Especial referencia a la tecnovigilancia,* Madrid: CGPJ, Cuadernos de Derecho Judicial, 2007-II, p. 236.

MANHEIM, K. M., KAPLAN, L. (2019) "Lyric, Artificial Intelligence: Risks to Privacy and Democracy", *21 Yale Journal of Law and Technology 106*.

NIEVA FENOLL, J. (2022) "El tránsito de la fe a la tecnología en el proceso penal", *Diario La Ley*, núm. 9986, de 11 de enero de 2022.

ORTIZ-PRADILLO, J. C. (2013) *Problemas procesales de la ciberdelincuencia*. Madrid: Colex.

ORTIZ PRADILLO, J. C. (2017) "Desafíos legales de las diligencias de investigación tecnológica", *El Proceso Penal. Cuestiones fundamentales*, (Coord. Olga Fuentes Soriano), Valencia: Tirant lo Blanch, pp. 303-316.

ORTIZ PRADILLO, J. C. (2021) "Big Data, Vigilancias Policiales y Geolocalización: nuevas dimensiones de los Derechos Fundamentales en el Proceso Penal", *Diario La Ley*, núm. 9955, de 18 de noviembre de 2021.

ORTIZ-PRADILLO, J. C. (2022) “Vigilancias policiales y utilización de dispositivos de seguimiento, localización y captación de la imagen”, *Reflexiones en torno al Anteproyecto de Ley De Enjuiciamiento Criminal de 2020* (Coordinadora Olga Fuentes Serrano). Valencia: Tirant lo blanch, pp. 813-881.

RODRIGUEZ LAINZ, J. L. (2013) “En torno al concepto de comunicación protegida por el art. 18.3 de la Constitución”. *Diario La Ley*, núm. 8142, de 5 de septiembre de 2013.

RODRÍGUEZ LÁINZ, J. L. (2021) “Reflexiones sobre el tratamiento de datos personales por prestadores de servicios de comunicaciones vía internet para la lucha contra abusos sexuales de menores en línea en el Reglamento (UE) 2021/1232”, *Diario La Ley*, núm. 9974, de 20 de diciembre de 2021.

SNOWDEN, E. (2019) *Vigilancia permanente*. Esther Cruz Santaella (Trad.), Barcelona: Editorial Planeta.

ZUBOFF, S. (2019). *The Age Of Surveillance Capitalism*. New York: PublicAffairs.