4

EL *BIG DATA* EN LA CIBERDEFENSA Y LA CIBERSEGURIDAD NACIONAL VERSUS EL DERECHO A LA PRIVACIDAD DEL CIUDADANO COLOMBIANO

El presente capítulo tiene por objeto realizar un estudio del *big data* en la ciberdefensa y ciberseguridad nacional vs. el derecho a la privacidad del ciudadano colombiano. Con ese propósito, se analizan y desarrollan los conceptos, los elementos, las fases y las técnicas del *big data* en el marco de la administración electrónica. Se realiza también un análisis del papel de los equipos multidisciplinarios y del uso del *big data*, los contextos de aplicación, la necesidad de análisis de información, el procesamiento de imágenes, voz y datos, el tratamiento de datos y determinados aspectos éticos y legales.

La metodología que se utiliza parte de explorar y comprobar los hechos desde una perspectiva jurídica, sigue los presupuestos de una investigación cualitativa y se diseña una teoría desde la interpretación de acciones y procesos sociales. En el marco de la utilización de la tecnologías de la información y las comunicaciones (TIC), en los últimos años han emergido herramientas de análisis masivo de información dentro de las propias administraciones. La combinación entre servicios, información, datos —muchos de estos privados—, herramientas y las nuevas técnicas de gestión de la información —como pueden ser las que utilizan *big data*— ocasionan situaciones que pueden poner en riesgo a gobiernos, administraciones,

ciudadanos, empresas, entre otros; por ende, su previsión debe ser un elemento en el marco del desarrollo de las competencias de cada administración.

Aproximaciones al *big data*

En Colombia, el Ministerio de Tecnología de la Información y las Comunicaciones (MinTIC, 2015a) ha señalado que "reconoce el valor de la información como herramienta para el fortalecimiento de sectores industriales, gubernamentales y académicos. La existencia de información disponible hace de su consecuente análisis un insumo cada vez más valioso en la toma de decisiones". Desde esa misma línea, ha referido que las tendencias mundiales en TIC evidencian la convergencia de tecnologías necesarias para el análisis de datos. Por eso, en reconocimiento prospectivo de las tendencias y los avances tecnológicos, crea Centros de Excelencia y Apropiación (CEA) con el fin de capitalizar el análisis de datos en sectores estratégicos. Los CEA en *big data analytics* focalizan las competencias de los recursos existentes para el desarrollo de estrategias cuya diferenciación recae en el análisis de información. Entre los objetivos del MinTIC está la creación de valor a partir de *big data anlaytics* para ciberseguridad, internet de las cosas y formulación de política pública (MinTIC, s. f.).

El análisis de grandes cantidades de datos y la generación de información que da valor añadido a la toma de decisiones de las organizaciones traen la necesidad de estudiar, desde la nueva perspectiva de derechos —muchos de estos fundamentales—, la intimidad, la privacidad, los derechos de autor, el derecho de acceso, entre otros, así como sus límites, que son elementos que han de pensarse antes de implementar estrategias de *big data*. Un ejemplo palpable de ello son las leyes de protección de datos de carácter personal, que como regla general han fundamentado el tratamiento desde la autorización previa, expresa e informada que da el titular de los datos a partir del consentimiento, cuya esencia es la libertad (Díaz, 2016).

De acuerdo con IBM (citado por Carrillo *et al.*, 2013, p. 8), *hay big* data si el conjunto de datos supera el terabyte de información y es sensible al tiempo; lo es además cuando mezcla datos estructurados con no estructurados. Su enfoque trata de buscar la manera de aprovechar estos datos, su combinación, su gestión y la aplicación de algoritmos predictivos de comportamiento; todo lo anterior,

con la finalidad de permitir la toma de decisiones que añadan valor al negocio. Es precisamente esa combinación de elementos, unos técnicos, otros jurídicos, la que nos lleva a realizar un análisis desde una nueva perspectiva.

La implementación de técnicas de *big data* no se da por sí sola, no tiene lugar desde la simple incorporación de tecnologías que permitan la práctica material *per se*; para ello es necesario realizar un diagnóstico integral de la misión de la organización, ya sea en el ámbito privado o en el público; luego procede una determinación del cumplimiento de esa misión desde la utilización de TIC, el tipo de datos que se maneja en el cumplimiento de la misión, la categorización de los datos, las implicaciones jurídicas del tratamiento de esos datos, la finalidad de dicho tratamiento, los aportes potenciales en la toma de decisiones si se cuenta con la información y el análisis de esos datos, la herramientas técnicas de análisis, las implicaciones jurídicas que trae el resultado del análisis de datos y, por último, el límite en cuanto al tratamiento y el análisis efectuados. Carrasco (2013) refiere los pasos para la implementación de *big data*:

a. Entender el negocio y los datos. Este primer paso exige un análisis detallado con las personas que hoy laboran y entienden los procesos y los datos que la empresa maneja.
b. Determinar los problemas y cómo los datos pueden ayudar a resolverlos.
c. Establecer expectativas razonables, es decir, definir metas alcanzables; esto se puede lograr si al implementar la solución de un problema, este no presenta alguna mejora y, por tanto, se debe buscar otra vía para hacerlo.
d. Cuando se inicia un proyecto de *big data,* es necesario trabajar en paralelo con el sistema que hoy está funcionando.
e. Al tratar de implementar un proyecto de *big data*, se debe ser flexible con la metodología y las herramientas; esto se debe a que ambas son recientes y pueden llegar a presentar problemas al ser implementadas.
f. Es importante mantener el objetivo de *big data* en mente, porque el proceso es pesado, máxime cuando los métodos y las herramientas que usan *big data* para el análisis de datos aún pueden presentar problemas. La idea es que se mantenga en mente la meta final del proyecto sin desanimarse pronto.

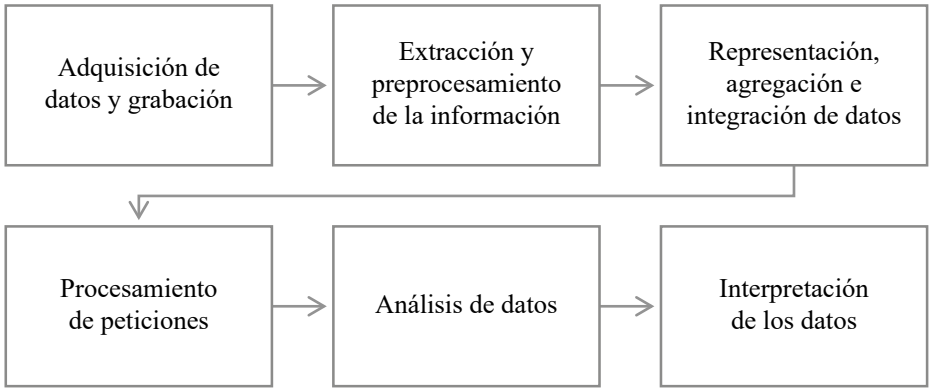
En relación con las dimensiones o los componentes del *big data*, en cuanto elementos esenciales y transversales a una gran cantidad de información, se

encuentran el volumen, la variedad, la velocidad y la veracidad de los datos. El volumen hace referencia a las cantidades masivas de datos que las organizaciones intentan aprovechar para mejorar la toma de decisiones en toda empresa, tanto privada como pública (Schroeck *et al.,* 2012, pp. 6-7). Asimismo, el volumen debe estar vinculado a las grandes cantidades de datos que llegan a las organizaciones (públicas o privadas) y que forman parte de la toma de decisiones estratégicas para el cumplimiento de su misión. De manera práctica, se puede afirmar que el volumen hace referencia al tamaño de la información (Grupo TRC, s f.).

Por su parte, la variedad tiene que ver con gestionar la complejidad de múltiples tipos de datos, incluidos los datos estructurados, semiestructurados y no estructurados. Entre los diferentes e innumerables tipos de datos se encuentran texto, datos web, tuits, datos de sensores, audio, video, secuencias de clic, archivos de registro y mucho más (Schroeck *et al.*, 2012, pp. 6-7). Uno de los grandes retos para el derecho será la delimitación de la finalidad para la cual los ciudadanos han autorizado el tratamiento de muchos de esos datos personales (algunos de ellos, de naturaleza sensible).

Los diferentes tipos de análisis siguen un conjunto de fases comunes que ayudarán en la toma de decisiones (Rayo, 2016). Las fases del *big data*, entendidas como aquellos pasos que deben seguir las empresas para la organización de la información, se muestran en la figura 1. Conviene resaltar que la fase final de interpretación de datos está directamente relacionada con el *data mining* o 'minería de datos', que se encarga de una multitud de tareas, como manipular, procesar, modelar, analizar y extraer la información que se necesite en función de resolver un problema determinado (Hop2croft, 2014).

Figura 1. Fases del *big data*



Fuente: Hop2croft (2014).

Aplicación del *big data* en la defensa y la seguridad nacional

Como se ha señalado, los datos (unos estructurados y otros no) de los que se pueden valer las organizaciones —incluido el Estado, desde los organismos de seguridad y defensa nacional— provienen de diversas fuentes y se encuentran en diferentes formatos. Aunque su uso en el sector de la seguridad nacional puede ser más que beneficioso, deberá estar limitado al cumplimiento de la finalidad legítima, pues su uso extralimitado puede ocasionar un gran perjuicio para los titulares de los datos. Desde la perspectiva de Carrillo *et al*. (2013, p. 44), el *big data* puede ser usado en diferentes entornos, pero ofrece grandes beneficios particularmente en el contexto de defensa y seguridad nacional, por lo cual será necesario precisar qué actividades presentes y futuras obtendrán resultados positivos en esta materia.

La defensa y la seguridad nacional tienen un enfoque preventivo, pues esto resulta menos costoso que invertir en daños que lesionen bienes jurídicos que se pretenden preservar; sin embargo, prevenir requiere decisiones de tiempo limitado y un alto nivel de síntesis en el número de datos y los factores involucrados. Por otro lado, las crisis actuales han visto aumentar su complejidad, todo ello debido a la globalización, los nuevos escenarios —con líneas difusas entre lo civil y lo militar—, los entornos intensivos en información con creciente mezcla de escenarios virtuales —p. e., ciberdefensa, económicos— y reales, etc. (Carrillo *et al*., 2013, p. 44).

La Política de Defensa y Seguridad para la Nueva Colombia 2015-2018 prioriza el cumplimiento de varios objetivos:

> Contribuir con capacidades, garantizar mayores y mejores niveles de seguridad ciudadana, contribuir a la modernización de la sociedad rural, combatir las nuevas y tempranas expresiones de crimen organizado que amenacen la seguridad y el funcionamiento transparente del Estado, garantizar la soberanía e integridad del territorio nacional, transformar y modernizar de forma continua el Sector Defensa, fortalecer la proyección internacional desde la cooperación bilateral, triangular y multilateral con los países aliados y estratégicos, poner al servicio del desarrollo nacional, comercial, industrial y agrícola las capacidades empresariales del Sector Defensa. Todo esto se podría materializar de mejor forma y de manera más eficaz desde la utilización de tecnologías, y de manera especial desde la utilización de técnicas de tratamiento y análisis de datos. (Ministerio de Defensa Nacional, 2015)

El *big data* ofrece alternativas en la solución de problemas existentes o emergentes. La aplicación del *big data* en materia de defensa y seguridad tiene por objeto captar y emplear grandes cantidades de datos con el fin de aunar sensores, percepción y decisión en sistemas autónomos, en clave de aumentar así el entendimiento de la situación y el contexto del analista y del agente del orden (Carrillo *et al*., 2013, p. 44). La inteligencia militar lucha contra el fraude y busca garantizar la seguridad ciudadana, el planeamiento táctico de misiones, la vigilancia y seguridad de fronteras, la lucha contraterrorista y contra el crimen organizado. Y, claramente, el *big data* puede emplearse en el ámbito de la defensa y la seguridad nacional. Es evidente que un Estado elabora una estrategia de seguridad y defensa nacional desde diferentes enfoques: desde el análisis de sus riesgos, en los distintos espacios estratégicos (tierra, mar y aire), para dar cumplimiento a los objetivos estratégicos nacionales (Ministerio de Defensa e Instituto Español de Estudios Estratégicos, 2012, p. 42).

La doctrina especializada ha señalado que la ciberdefensa es el conjunto de acciones u operaciones activas o pasivas desarrolladas en el ámbito de las redes, sistemas, equipos, enlaces y personal de los recursos informáticos y teleinformáticos de la defensa, a fin de asegurar el cumplimiento de las misiones o los servicios para los que fueran concebidos, a la vez que para impedir que fuerzas enemigas los utilicen para cumplir los suyos. Se ha planteado iniciar el proceso de ciberdefensa a través de la inteligencia informática, con el ciberespacio como

ambiente, para poder obtener los elementos descriptores que conformen la identificación de los escenarios y, a la vez, parametrizar las amenazas. Con ello será posible dimensionar los riesgos y permitir así el diseño de los instrumentos de defensa (Consejo Argentino para las Relaciones Internacionales [CARI], 2013, p. 9).

En este contexto, la ciberdefensa de basará en la utilización de las TIC para el cumplimiento de su fin; por ello, sus agentes deberán contar con instrumentos jurídicos que les permitan la realización de actividades de inteligencia en el ciberespacio. Por ello, la Ley 1621 de 2013 tiene por objeto establecer el marco jurídico para los organismos que llevan a cabo actividades de inteligencia y contrainteligencia; allí se establecen los límites y fines de las actividades de inteligencia y contrainteligencia, los principios que las rigen, los mecanismos de control y supervisión, la regulación de las bases de datos, la protección de los agentes, la coordinación y cooperación entre los organismos, los deberes de colaboración de las entidades públicas y privadas, entre otras particularidades.

La ciberdefensa se efectúa en términos de *defensa activa* y *defensa pasiva* del centro de operaciones, de los medios de información que posee la institución, con el fin de repeler los ataques cibernéticos que esta sufra, y aquí su arma rectora por disposición son las comunicaciones militares, que coadyuvan a la protección cibernética de la infraestructura crítica del país. La primera de ellas se puede explicar como una estrategia para adquirir una capacidad de defensa del ciberespacio, que combina la protección interior de los sistemas, la vigilancia permanente de redes sensibles y la respuesta rápida en caso de ataque, contrarrestando de este modo las amenazas ciberespaciales y garantizando acceso al ciberespacio. La defensa pasiva, por su parte, es la estrategia para la protección de los activos relacionados con los sistemas de información, a través de controles detectivos, correctivos y disuasivos que contrarresten las posibles amenazas (Fuerzas Militares de Colombia, 2015).

Ahora bien, se hace necesario abordar las aplicaciones específicas del *big data* en cada uno de los contextos señalados. Entre las aplicaciones específicas identificadas se encuentran: detección de intrusión física en grandes espacios o infraestructuras abiertas, criminología computacional, computación sobre información encriptada, análisis automático de vulnerabilidades de red, uso fraudulento de recursos corporativos y/o sensibles, identificación de anomalías, patrones

y comportamientos en grandes volúmenes de datos, inteligencia visual en máquinas, análisis de texto (estructurado y no estructurado), apoyo a la toma de decisión en tiempo real en contextos intensivos en datos, conciencia situacional, predicción de eventos, etc. (Carrillo *et al*., 2013, p. 46).

Las innovaciones de *big data* emergen en un momento en el que las organizaciones se enfrentan a nuevas problemáticas que se derivan de dos desafíos: a) disolución de los límites de las redes y b) extensión y apertura de las redes de datos de las organizaciones. Ello permite que socios, suministradores y clientes accedan a información de carácter corporativo a través de formas dinámicas para impulsar la innovación y la colaboración, lo cual, en último término, deviene en que las redes sean más vulnerables al mal uso y robo de información. Las aplicaciones y los datos corporativos son cada vez más accesibles por medio de, por ejemplo, servicios en la nube (Carrillo *et al*., 2013, p. 50).

El siglo de la sociedad de la información —también llamado siglo de la revolución de la información— ha traído, y sigue trayendo, nuevos riesgos para los derechos de los ciudadanos y para la seguridad de las naciones, dadas las nuevas formas e instrumentos de criminalidad, así como las renovadas dinámicas del activismo, el espionaje y el terrorismo. Con frecuencia, los ataques o fraudes no son detectados, sino que solo se perciben cuando el daño se ha materializado. Por esta razón, como lo señalan Carrillo *et al.* (2013), es pertinente buscar soluciones más ágiles basadas en evaluaciones dinámicas de riesgo. Para ello serán esenciales las operaciones de seguridad en tiempo real, así como el análisis de grandes volúmenes de datos; todo ello con el fin de garantizar seguridad.

Objetivos del *big data* en la aplicación de seguridad y defensa

En la más reciente Directiva del Parlamento Europeo y del Consejo de la Unión Europea (UE), del 6 de julio de 2016, relativa a las medidas destinadas a garantizar un elevado nivel común de seguridad de las redes y sistemas de información de la Unión, a fin de mejorar el funcionamiento del mercado interior, se dan varias prerrogativas:

a. Establece obligaciones para todos los Estados miembros de adoptar una estrategia nacional de seguridad de las redes y los sistemas de información.

b. Crea un grupo de cooperación para apoyar y facilitar la cooperación estratégica y el intercambio de información entre los Estados miembros, y desarrollar la confianza y seguridad entre ellos.
c. Crea una red de equipos de respuesta a incidentes de seguridad informática (en lo sucesivo, red de CSIRT, por sus siglas en inglés), con el fin de contribuir al desarrollo de la confianza y seguridad entre los Estados miembros y promover una cooperación operativa rápida y eficaz.
d. Estipula requisitos en materia de seguridad y notificación para los operadores de servicios esenciales y para los proveedores de servicios digitales.
e. Establece obligaciones para que los Estados miembros designen autoridades nacionales competentes, puntos de contacto únicos y CSIRT con funciones relacionadas con la seguridad de las redes y los sistemas de información.

La Ley de Inteligencia colombiana (Ley 1621 de 2013) establece de manera clara que la función de inteligencia y contrainteligencia es llevada a cabo por las dependencias de las Fuerzas Militares y la Policía Nacional —p. e., la Unidad de Información y Análisis Financiero (UIAF)— y por los demás organismos que faculte para ello la ley en mención (artículo 3). Por ello, el límite de la actividad de inteligencia está dado por el respeto a la Constitución, a la ley, a los derechos de los ciudadanos y a las normas del derecho internacional, para así asegurar los fines del Estado colombiano establecidos en el artículo 2 de la Carta Política: la vigencia del régimen democrático, la soberanía, la protección de las instituciones, hacer frente a amenazas como el terrorismo, el crimen organizado, el narcotráfico o el secuestro, y proteger los recursos naturales y los intereses económicos de la Nación.

Es precisamente sobre dichos contenidos que debe dársele sentido a lo que pretende la aplicación del *big data* en la ciberseguridad y defensa, a saber:

a. Disminuir el tiempo que los analistas emplean descubriendo ciberataques, al reunir y relacionar las fuentes de redes de datos dispares.
b. Aumentar la precisión, tasa y velocidad de detección de ciberamenazas a redes de computadores (Carrillo *et al*., 2013, p. 49).

En este orden de ideas, uno de los elementos fundamentales será determinar cuáles son aquellas cuestiones, desde la aplicación de análisis de datos, que deberían ser consideradas para las actividades de seguridad y defensa nacional, y que

incluye esas actividades. Carillo *et al.* (2013) han señalado que en los próximos años, a largo plazo, se dimensiona que en el análisis de información de *big data* el rompimiento en la mayor parte de productos de seguridad en redes incluye la monitorización de estas, la autenticación y autorización de usuarios, la gestión de identidades, la detección de fraudes y la gestión de gobernanza, riesgos y conformidad. A mediano plazo, se espera que las herramientas mejoren, al punto de permitir un número considerable de posibilidades de predicción avanzadas y controles automatizados en tiempo real. Esto será aplicado en las áreas relacionadas con protección de redes de telecomunicaciones, ciberdefensa, ciberseguridad y protección de infraestructuras críticas (Carrillo *et al*., 2013, p. 51).

Criminología en la aplicación del *big data*

La criminología computacional tiene un ámbito de aplicación directo en el *big data*, por cuanto analiza grandes cantidades de información relacionadas con actividades criminales. El uso de *big data* incrementa las probabilidades de neutralización de estas amenazas; en este sentido, se relaciona con diferentes técnicas como el cifrado de datos criminales, el análisis de agrupaciones, el aprendizaje de reglas de asociación para la predicción del crimen, el análisis de textos multilingües, opiniones y redes criminales, entre otras acciones (Carrillo *et al*., 2013, p. 51).

Desde otra perspectiva, la aplicación del *big data* en la criminología computacional consiste en el seguimiento de actividades sospechosas en redes, internet o en las páginas web de las entidades; de ahí que sea fundamental identificar qué información es recogida como soporte para la actividad criminal. Por otro lado, *big data* puede ser implementado con el fin de recopilar, almacenar y documentar pruebas de actividades ilícitas o maliciosas realizadas dentro y fuera de internet o de las entidades corporativas; esto último, en un sentido preventivo, aprovechando la capacidad de analizar, procesar y almacenar volúmenes considerables de información. Las áreas de aplicación utilizadas son las siguientes: la seguridad general, la lucha contraterrorista y contra el crimen organizado, la lucha contra la usurpación de identidad, la lucha contra el fraude, la lucha contra la explotación infantil y la pedofilia, etc. (Carrillo *et al*., 2013, p. 52).

Uso fraudulento de la información en el contexto del *big data*

Diferentes entidades y corporaciones han implementado el uso del *big data* en sus plataformas debido al gran número de usuarios y de recursos. Además del contenido de tipo sensible que puede tener esta información, se han dado unas pautas y reglas específicas, principalmente por el número considerable de sesiones concurrentes de acceso a estos recursos. De esta manera, gracias a las técnicas de *big data*, es posible analizar en tiempo real los datos de estas sesiones de acceso, identificando patrones de comportamiento y diferenciando usos permitidos de los recursos frente a otros, ya sean relacionados con el abuso de los recursos de información de la organización, o bien, con ataques cibernéticos (Carrillo *et al*., 2013, p. 52).

Un ejemplo de lo anterior es que el país, con más de 180.000 empleados de la Policía Nacional, se enfrentó al reto de correlacionar toda la información procedente de diferentes fuentes: desde cámaras de videovigilancia hasta llamadas al 123 y flujos de trabajo del personal en las calles. Para dar respuesta a este requerimiento, la Policía está inmersa en un proyecto de *big data* que permite transformar esa información en conocimiento. Así lo explicó Jairo Gordillo, director de TI de la Policía Nacional, durante el evento Datacenter Dynamics Converged (Villarrubia, 2014).

Se debe señalar que los ataques al bien jurídico de la información, los datos y los sistemas de información se efectúan desde el interior de la misma entidad y corporación, incluso por parte de los empleados, que aprovechan las fallas. Las técnicas de *big data* favorecen el desarrollo de herramientas frente a seguridad informática y gestión de conocimiento de grandes organizaciones (Carrillo *et al.*, 2013, p. 52).

La importancia de la conciencia situacional del *big data* en el contexto de la seguridad y defensa nacional de los Estados

Para que el *big data* se vuelva una herramienta contundente en el ámbito de la seguridad y defensa, es fundamental que se analice el contexto de la situación respecto del analista o combatiente, agente de orden, para efectos de esta investigación. Es preciso resaltar que, dada la necesidad de identificar amenazas en

situaciones complejas, contradictorias e inciertas, en las que hay una cantidad considerable de datos disponibles de fuentes abiertas, es determinante realizar acciones para su neutralización y la detección de la amenaza (Carrillo *et al*., 2013, p. 59).

La Organización de Estados Americanos (OEA) ha trabajado para fortalecer las capacidades de seguridad cibernética entre los Estados miembros desde principios de la década del 2000. El Programa de Seguridad Cibernética de la OEA apoya las iniciativas sobre la base de un análisis en profundidad y a partir de la comprensión de la magnitud de las amenazas; esto es, conocer el estado de la situación. La comprensión o el entendimiento contextual solo se pueden adquirir a través de la combinación de técnicas adecuadas —es decir, humano/máquina—, donde se aproveche la capacidad cognitiva de las personas para fundir y asimilar múltiples fuentes y tipos de información, en función de lograr nuevas perspectivas. Las técnicas de *big data* se aplican en dicho contexto, simplificando la exploración y el procesamiento de diversos datos; en específico, el *big data* puede aportar a través del procesamiento de sistemas de aprendizaje inmediato que procesen un lenguaje natural e inserten la representación semántica resultante en una base de conocimiento, en lugar de basarse en los actuales procesos costosos, que requieren de mucho tiempo para el aprendizaje de personas con diferentes áreas del conocimiento.

Hoy se ve posible unificar estas bases de datos con capacidades sensoriales como la "inteligencia visual" en máquinas, en las que se identifican no los objetos existentes en diferentes imágenes o videos, sino también las relaciones (acciones) entre ellos. De esta manera, se posibilita interpretarlos de manera inmediata y construir así una narrativa de información visual, que agrega resultados provenientes de la correlación en aspectos económicos, demográficos, patrones sociales, etc. (Carrillo *et al*., 2013, p. 59).

En el futuro se espera obtener información en varios frentes: en el descubrimiento de sucesos específicos, tanto en su planeamiento como una vez hayan ocurrido; en el develamiento de temas y conceptos desarrollados colaborativamente; en valores y creencias que motivan ciertos comportamiento de interés; en el análisis semántico y de inclinación en el apoyo a grupos y personas; en áreas de aplicación relativas al planeamiento táctico de misiones; en la toma de decisiones en tiempo real para las operaciones concernientes a defensa y seguridad;

en traducciones automáticas en un considerable número de idiomas y volumen. Sin embargo, este tipo de actividades trae aparejados riesgos para la intimidad de los ciudadanos; por ello, es esencial la delimitación de la finalidad para la cual se realizan las actividades mismas.

Pese a que en la actualidad se han producido avances destacados en la recuperación y el tratamiento de datos, la realidad no es tan promisoria; un claro ejemplo son los motores de búsqueda inteligentes, donde el progreso en la comprensión de la información no ha avanzado con el mismo ritmo. Una de las razones para este desacuerdo es el considerable aumento de datos no organizados relacionados tanto con el teatro de operaciones y el entorno directo del combatiente como con los escenarios de seguridad (Carrillo *et al*., 2013, pp. 59-60).

La creación de equipos multidisciplinares para crear un control efectivo en la aplicación de *big data*

El *big data* no es un sistema *software* que arroja informes de manera automática; tampoco consiste en un conjunto de sistemas informáticos que, una vez instalado y configurado, empieza a generar soluciones y aplicaciones relativas a la seguridad informática (Carrillo *et al*., 2013, p. 71). Por el contrario, se necesitan personas y profesionales para su operación; de ahí que sea fundamental la formación técnica y profesional en esta nueva área del conocimiento, que aborde las fuentes de datos, la tipología, la calidad, la naturaleza, la algoritmia estadística, la analítica, entre otros aspectos, desde la dimensión técnica, jurídica y práctica.

Carrillo *et al.* (2013) argumentan que este perfil de científico de datos implica conocer el contexto de complejidad del entorno de la aplicación. En las aplicaciones de seguridad y defensa es necesario señalar que la implementación de estos perfiles puede ser parte de la solución de la problemática ético-legal del *big data*. En suma, se busca incorporar el contexto de los datos de análisis, debido a que algunos de ellos son intrínsecamente inciertos —p. e., las señales GPS que rebotan entre los edificios—; también es necesario configurar un contexto de análisis que permita reducir la implementación de recursos (computación, almacenamiento, etc.), al hacer posible el planteamiento *big data* solo en aquellas fuentes que son relevantes para determinado contexto o en diferentes subconjuntos, de acuerdo con criterios de filtrado.

Actualmente se han creado diversas y numerosas herramientas informáticas; la mayoría de ellas forman parte de la evolución de plataformas *Business Intelligence*, un avance a la complejidad de visualizaciones. Entre los desafíos de la visualización de *big data* se encuentran la solución a la variedad de datos multidimensionales y aspectos espacio-temporales, el énfasis en análisis de información, la solución de problemas y la toma de decisiones.

La temporalidad, el análisis de escalas múltiples, la correlación y la causalidad son el gran reto; pero, más allá de esto, poder hacer frente al hecho mismo de que los criminales pueden igualmente identificar las correlaciones y anticiparse a las decisiones actuando de forma diferente a la planeada de forma inicial. Por ello, en el marco del equipo multidisciplinar, se deben incorporar expertos al respecto. Una de las características del *big data* es el análisis en tiempo real en *streams* de datos. Las capacidades existentes para procesar tal flujo de datos son tan intensivas en conocimiento y datos que necesitan razonamiento autorizado (Carrillo *et al*., 2013, p. 74).

La necesidad del Estado en el análisis de datos

En el sector público colombiano se ha avanzado en algunos pilotos, como los realizados por el Departamento Administrativo Nacional de Estadística (DANE) con el uso de *big data* en estadísticas oficiales incorporadas en sus estrategias, en el monitoreo de los Objetivos de Desarrollo Sostenible y en el Censo Nacional de Población y Vivienda. De igual manera ha avanzado el Ministerio de Hacienda, junto con el Departamento Nacional de Planeación (DNP), en una metodología a partir de datos de Google Trends, que analiza la frecuencia de términos de búsqueda para inferir actividad económica en ciertos sectores y obtener indicadores adquiridos anteriormente con estadísticas tradicionales (Infraestructura Colombiana de Datos Espaciales [ICDE], 2017).

Actualmente ha surgido la necesidad, por parte de los Estados, de analizar datos y percibir acciones. Esta actividad se realiza de manera poco automática, más bien tiende a ser lenta y se realiza frecuentemente por parte de las administraciones públicas con el fin de garantizar la seguridad nacional y la defensa del Estado; sin embargo, ello puede llegar a violentar libertades individuales de los ciudadanos, como se verá más adelante. La consultora Govwin Networks señala

que el Departamento de Defensa de los Estados Unidos gasta aproximadamente el 58,4% de todo el gasto federal en almacenamiento de datos, y la mayor parte proviene de la necesidad de analizar videos. Esos programas tienen como finalidad captar una significativa cantidad de imágenes y videos en tiempo real para el reconocimiento y toma de decisiones en el contexto de seguridad y defensa nacional (Carrillo *et al.*, 2013, p. 100).

Otro avance significativo en el uso del *big data* tiene lugar con la creación de la Biblioteca Virtual en Salud para la Vigilancia en Salud Pública, impulsada por el Ministerio de Salud y Protección Social, con características de un Repositorio Digital Institucional, definido, según la Red Alfa Biblioteca de Babel, como "un archivo electrónico de la producción científica de una institución, almacenada en un formato digital, en el que se permite la búsqueda y la recuperación para su posterior uso nacional o internacional" (Palacios *et al.,* 2014, pp. 81-82).

Por último, el uso estratégico del *big data* surge en el marco de arquitectura empresarial que promueve estándares, buenas prácticas e interoperabilidad para todo el sector público. Según Bohórquez (citado en Fernández, 2014, pp. 48-49), uno de los pilares de este marco, en términos de compras y contratación, es la preferencia por soluciones en la nube, en vez de infraestructura propia. Así, de acuerdo con el Programa de Compra Eficiente, la infraestructura tecnológica para *big data* estaría disponible a través de servicios en la nube que serían ofrecidos bajo acuerdos de precios aplicables a todo el sector público y estandarización en tecnología, calidad y variedad de proveedores.

Fuentes de imágenes y video

Existen diferentes tipos de video que pueden tener diversas funciones para asegurar la defensa y seguridad nacional. Entre estos se encuentran diferentes dispositivos con variadas funciones, a saber: a) UAV o drones (cuya función es capturar una significativa cantidad de videos para el reconocimiento de imágenes en situaciones hostiles); b) videos procedentes de los vehículos de exploración y reconocimiento terrestre, tanto en el espectro visible como en el infrarrojo; c) imágenes tomadas de satélites de vigilancia; d) videos de cámaras de vigilancia en lugares públicos; e) ubicación de cámaras de vigilancia en espacios privados, como hospitales y empresas; f) publicación y difusión en diferentes redes sociales, como

Facebook, Twitter, Youtube, blogs y otros lugares de la red (Carrillo *et al*., 2013, p. 101). Precisamente esas diversas fuentes nos llevan a entender que las actividades tendientes a garantizar la defensa y seguridad nacional deben estar amparadas en la Constitución y la ley, desde el respeto de los derechos de los ciudadanos —entre ellos, el derecho a la intimidad—, pero desde el entendimiento de la necesidad, idoneidad y proporcionalidad para garantizar la seguridad y defensa.

Procesamiento y uso de imágenes como herramienta para garantizar la defensa y seguridad de los Estados

La infraestructura *big data* frente al procesamiento de imágenes se ha aplicado progresivamente; pese a ello, esta herramienta le ha dado un aporte sustancial a la recolección de datos frente a la defensa y seguridad nacional, a través de la detección de movimientos y la detención en espacios no permitidos o en zonas de exclusión. Esta infraestructura implica el desarrollo de varias acciones: vigilancia de infraestructuras críticas, reconocimientos faciales en zonas determinadas, seguimiento y reconocimiento de objetivos, reconocimiento de comportamientos sospechosos en sitios públicos, identificación de actividades económicas o en zonas de conflicto, identificación de objetos abandonados, sospechosos, IED, etc. En el ámbito militar, estas herramientas tienen gran relevancia y utilidad, particularmente frente a los distintos sistemas existentes y desplegados. Entre los más importantes se pueden mencionar: herramientas de visualización, seguimiento de objetivos (ISTAR), herramientas de visualización del estado operacional de campo de batalla, herramientas de ayuda en torno a decisiones de carácter militar, etc. (Carrillo *et al*., 2013, p. 101).

Por otro lado, la defensa requiere de ciertas habilidades y herramientas como son los programas I+D, que se pueden lograr a través de la implementación de diferentes tecnologías de procesamiento de imágenes e infraestructura *big data*. En el mercado se pueden encontrar diferentes alternativas, pues los fabricantes ofrecen productos de *big data*; por su parte, las empresas de ingenieros ofrecen sistemas de procesamiento cada vez más inteligentes. El uso de estas dos herramientas para lograr la capacidad requerida no es tarea fácil.

En 2000, la STO organizó un congreso en Canadá con el nombre de IST.020/rws-002: "Multimedia Visualization of Massive Military Data Dases", que causó

gran interés frente a la visualización y presentación de datos, así como en relación con la vigilancia y función de grandes cantidades de datos y matemáticas de procesamiento. Vale señalar que desde hace casi dos décadas ya se plantea la necesidad de desarrollar actividades tanto investigativas como tecnológicas tendientes a la incorporación de proyectos de analítica de datos para las actividades de seguridad y defensa nacional. Posteriormente, la STO ha organizado, a través de su panel IST (Information Systems Technologies), diferentes grupos temáticos y de trabajo relacionados con la fusión de datos. Aunque el eje principal no es el procesamiento de imágenes, en el objeto de análisis se encuentran las fuentes heterogéneas que analizan estudios de la OTAN (Carrillo *et al*., 2013, p. 103).

Uno de los programas implementados por la DARPA para garantizar la defensa y seguridad nacional es el Mind's Eye, utilizado en el gobierno del exmandatario Barack Obama. Este programa es empleado con el fin de "organizar, acceder y descubrir información útil de las grandes cantidades de datos de los que disponen su administración". El resultado de los datos procesados permite anticipar el actuar de grupos que quieran realizar actos hostiles contra el pueblo norteamericano. El objetivo del programa Mind's Eye consiste en otorgarle a un sistema autónomo terrestre (UGV) una cámara inteligente que sea capaz de describir la escena que está visualizando. Esto se hace a través de verbos que puedan describir la escena. Además, podrá aprender nuevos conceptos a partir de su experiencia visual. La implementación de este programa ayudará sustancialmente en la minimización de riesgos, pues el combatiente no tendrá la necesidad de desplazarse (Carrillo *et al*., 2013, p. 104).

Existe otro programa denominado Insight, cuyo objetivo es mitigar los costos de supervisar las oleadas de datos que pueden llegar a los combatientes en un momento determinado, con la finalidad de obtenerse en tiempo casi real. Para ello, este programa está trabajando un sistema de inteligencia, reconocimiento y seguimiento (ISR) en el que actúa el *big data* de la mano de imágenes que provienen de los sensores ISR ubicados en el campo de batalla (Carrillo *et al*., 2013, p. 104).

En la defensa y seguridad nacional, el uso de imágenes es muy demandado debido a que estas se presentan como una herramienta que permite mejorar los procesamientos de imágenes en los siguientes aspectos: flexibilizar las infraestructuras para ejecutar soluciones a medida fácilmente; facilitar el procesamiento masivo en paralelo, dando potencia de cálculo a los sistemas; suministrar

información externa adicional con la cual poder fusionar información, mejorando su calidad y haciendo que esté compuesta de datos fidedignos y confiables (Carrillo *et al*., 2013, p. 105).

Para el aprovechamiento del *big data* en el ámbito de defensa y seguridad nacional serán necesarias las infraestructuras respectivas, así como los grandes volúmenes por analizar, lo que puede significar en un principio que solo los países desarrollados, con los recursos necesarios, puedan aprovechar toda la amalgama de posibilidades que implica el *big data*. Sin embargo, cuando los costes se empiecen a hacer más accesibles, y se desarrollen más aplicaciones y herramientas, estas soluciones ganarán terreno en los sistemas que no puedan ser considerados como *big data*, y de igual manera serán implementados en países en vía de desarrollo (Carrillo *et al*., 2013, p. 105). Asimismo, se puede llegar a aplicar el *big data* en la guerra electrónica, cuyo fin sería reducir o impedir el empleo del espectro electromagnético; la guerra electrónica se valdrá del espectro como campo de batalla.

En este punto es pertinente señalar las clases de acciones desplegables en el contexto de guerra electrónica, a saber: sincronización del momento de emisión con otros receptores situados en diferentes lugares para localizar el emisor (Angle of Arrival [AOA]); búsqueda y captación de cada frecuencia en la que no solo existe ruido; interceptación y mantenimiento de la frecuencia que emite para conocer sus parámetros; identificación y cotejo de dichos parámetros con emisores esperados o conocidos; localización; entre otros (Carrillo *et al*., 2013, p. 109).

Aspectos legales y éticos en el procesamiento de información *big data*

Uno de los retos que enfrenta la aplicación del *big data* está relacionado con la utilización y explotación de datos, debido a que este procesamiento puede transgredir aspectos legales y éticos como son el derecho a la privacidad de los ciudadanos. Por ello, este aspecto es relevante en los ámbitos de la seguridad y la sanidad (Carrillo *et al*., 2013, p. 71). La creciente utilización de tecnologías por las administraciones públicas trae consigo la necesidad de determinar si la infraestructura de tecnologías y la información propiamente dicha en que se funda la e-administración tienen una categoría especial de protección. Desde ya se debe decir que no existe norma colombiana que considere que el daño a la

infraestructura tecnológica de la administración pública sea considerado tan de alto nivel y, por ende, sea fundamental su protección.

Diferentes organizaciones de carácter gubernamental y no gubernamental implementan el *big data* como herramienta de procesamiento de información; no obstante, son las entidades públicas particularmente las que reciben y acceden a información de carácter sensible. Por esta razón, es necesario implementar nueva legislación que contemple este tipo de situaciones, ya que en la actualidad la legislación no cubre todo el tipo de posibilidades y usos de captación y análisis del *big data* (Carrillo *et al.*, 2013, p. 71).

El derecho de seguridad y defensa versus el derecho de privacidad: una mirada desde la jurisprudencia de la Corte Constitucional colombiana

El *big data* puede constituir una herramienta fundamental en la búsqueda de información, particularmente en lo concerniente a la defensa y seguridad de los Estados; sin embargo, su análisis y procesamiento puede llegar a vulnerar derechos fundamentales como la privacidad, la honra, el buen nombre y el debido proceso. Por esta razón, la Corte Constitucional ha analizado jurisprudencialmente dichas problemáticas, como se verá a continuación.

En la Sentencia T-277 de 2015 se analiza cómo con el *big data* las personas pueden acceder a todo tipo de información; sin embargo, se ha creado una figura denominada *Habeas data*, que tiene como pretensión crear la posibilidad de conocer, rectificar y actualizar las informaciones que se tienen sobre una persona en los bancos de datos públicos o privados. Las innovaciones de *big data* emergen en un momento en el que las organizaciones se enfrentan a nuevas problemáticas que se derivan de dos desafíos: disolución de los límites de las redes y extensión y apertura de las redes de datos de las organizaciones. Con ello se permitiría a socios, suministradores y clientes acceder a información de carácter corporativo a través de formas dinámicas para impulsar la innovación y la colaboración; y, a la final, ello podría derivar en que las redes sean más vulnerables al mal uso y el robo de información. Las aplicaciones y los datos corporativos son cada vez más accesibles por medio de servicios en la nube, por ejemplo.

El señor Roberto Eladio Espriella Fernández instaura una acción de tutela contra Ecopetrol S.A. con fundamento en lo indicado anteriormente, debido a que

la empresa lo señaló de pertenecer a grupos al margen de la ley (Sentencia T-022 de 2017). La empresa ejerce su derecho de defensa manifestando que la información publicada por ella el 19 de febrero de 2015 se refiere a hechos que ocurrían para ese momento. Por otro lado, señala que no puede hacerse responsable del sinnúmero de publicaciones periodísticas que surgieron a partir de esto.

De esta manera, el Juzgado Treinta Civil del Circuito concede las pretensiones del señor Espriella, pues evidentemente se vulneraban sus derechos a la honra, intimidad y buen nombre. Sin embargo, la decisión fue revocada por la Sala Primera Civil de Decisión del Tribunal Superior de Bogotá, razón por la cual la Corte entra a analizar el caso. Lo hace confirmando la decisión de primera instancia y señalando que la publicación, en efecto, violaba los derechos fundamentales del demandante anteriormente relacionados, ya que la información publicada por la Empresa no tenía un carácter público, solo le concernía a las partes, es decir, a la entidad y al trabajador. La decisión adoptada por la Corte se fundó, además de lo preceptuado en la Constitución, en la transgresión al Código de Ética, el Reglamento Interno de Trabajo y al Código Sustantivo del Trabajo.

Por otro lado, en la Sentencia T-277 de 2015 la Casa Editorial El Tiempo publicó una noticia en la cual se vinculaba a una persona con el delito de trata de personas, situación que vulneró los derechos de la accionante a la honra, el buen nombre y el debido proceso. No obstante, la editorial fue exonerada en la investigación por parte de la Fiscalía, pues había operado la prescripción de la acción. Pese a lo anterior, esta información seguía apareciendo en motores de búsqueda como Google y en las bases de datos de la editorial, situación que estaba afectando a la accionante y a su familia. Por esta razón, la accionante agotó la vía administrativa mediante derecho de petición ante la Casa Editorial El Tiempo; sin embargo, esta última se rehusaba a eliminar esta información de sus bases de datos argumentando que dicha información era veraz e imparcial. Por esta razón, la accionante acude a la acción de tutela, ya que para ella existe una doble discriminación: en principio se le adjudicaba un delito en el que ella no fue vencida en juicio y posteriormente esta editorial la señala públicamente como vinculada a estos delitos, sin rectificar o eliminar la información contenida, pese a que había operado la prescripción.

Para la Corte existe una contraposición de derechos relativa a la libertad de expresión vs. los derechos a la honra, el buen nombre y el debido proceso.

Explica que el derecho a la libertad de expresión se refiere a la libertad de informar y recibir información veraz e imparcial; sin embargo, esto plantea unas problemáticas en torno a la aplicación de la red como herramienta que garantiza este derecho a la información. Así, establece que en la red —al ser descentralizada— los mensajes y los contenidos producidos se transmiten de forma tal que la revisión o censura previa de contenidos por parte de una autoridad central es una tarea compleja. En principio, esta parece ser una de las innovaciones y habilidades de internet, pues sin lugar a dudas lo hace un entorno libre, lo que posteriormente puede presentar ciertos retos en aspectos sensibles como los relativos al control de contenidos tendientes a afectar derechos como la intimidad, la honra, la imagen y el buen nombre de las personas.

La Corte señala en la Sentencia T-277 de 2015 el derecho de *Habeas data* del que gozan las personas, y refiere constitucionalmente el artículo 15 de la Carta Política, donde se consagra que las personas "tienen derecho a conocer, actualizar y rectificar las informaciones que se hayan recogido sobre ellas en bancos de datos y en archivos de entidades públicas y privadas. En la recolección, tratamiento y circulación de datos se respetarán la libertad y demás garantías consagradas en la Constitución".

Lo anterior permite establecer que el derecho de *Habeas data* implica una doble vía: el conocimiento de información y la rectificación. Por eso, en el presente casi la Corte le ordena a la Casa Editorial El Tiempo la actualización de la información sobre los hechos que relacionan a la accionante con el delito de trata de personas, y que de igual manera se informe que esta no fue vencida en juicio. Por otro lado, le ordena a El Tiempo que por medio de la herramienta técnica "robots.txt", "metatags" u otra similar neutralice el acceso a la noticia "Empresa de trata de blancas" a partir de la mera digitación del nombre de la accionante en los motores de búsqueda en la red (Sentencia T-277 de 2015).

Con el fin de evitar el uso indebido de datos tanto por entidades públicas como privadas, se han propuesto diferentes proyectos de ley. Un claro ejemplo de estas iniciativas fue el Proyecto de Ley 184 de 2010, ley estatutaria que pretendía la protección de datos personales en las bases de datos, con el fin de garantizar el derecho a la intimidad de los ciudadanos. Sin embargo, dicho proyecto de ley fue demandado en la Sentencia C-748 de 2011, ya que para el accionante tenía vicios en el procedimiento y vulneraba los principios a la unidad de materia.