Universitat d'Alacant
Universidad de Alicante
Facultat de Dret
Facultad de Derecho

FACULTAD DE DERECHO

GRADO EN CRIMINOLOGÍA

TRABAJO DE FIN DE GRADO

CURSO ACADÉMICO [2020-2021]

TÍTULO:

EL BIG DATA COMO HERRAMIENTA DE PREVENCIÓN DE LA DELINCUENCIA

AUTOR:

MIGUEL ÁNGEL JIMÉNEZ HERNÁNDEZ

TUTOR ACADÉMICO:

MARTA GARRIDO MACÍAS

RESUMEN

Este trabajo consiste en una extensa revisión bibliográfica dirigida a conocer como las tecnologías que utilizan Big Data pueden ser herramientas útiles para la prevención del crimen. Para ello, se han explicado varios estudios, cada uno con un enfoque distinto, los cuales ponen a prueba algoritmos basados en Big Data para predecir el riesgo delictivo. Como conclusión se ha obtenido que este tipo de tecnologías, utilizadas respetando todos los límites éticos y jurídicos, son una herramienta más eficaz a la hora de predecir la incidencia delictiva que se producirá en un espacio-tiempo determinado que los métodos utilizados tradicionalmente. Con ello se demuestra que el Big Data es una arma útil y necesaria en la lucha por conseguir una prevención situacional óptima.

Palabras clave: *prevención del delito, teorías de la pena, formas alternativas de prevención, prevención situacional, Big Data, data mining y machine learning, legislación en protección de datos y valores éticos.*

ABSTRACT

This project consists of an extensive bibliographic research focus on knowing how Big Data technologies can be useful tools for crime prevention. For this, several studies, each one with a different approach, which test algorithms based on Big Data to forecast crime risk, have been explained. As a conclusion, it has been obtained that this type of technology, used respecting all ethical and legal limits, is a more effective tool when it comes to predict the crime incidence that will occur in a given space-time than the methods that traditionally are used. This shows that Big Data is a useful and necessary weapon in the fight to achieve an optimal situational prevention.

Keywords: crime prevention, theories of punishment, alternative forms of prevention, situational prevention, Big Data, data mining and machine learning, data protection legislation and ethical values.

ÍNDICE

1. INTRODUCCIÓN

La delincuencia es un elemento que siempre ha convivido con el ser humano, ahí donde imperaba un ordenamiento jurídico, había alguien dispuesto a transgredirlo. En este contexto, las Fuerzas y Cuerpos de Seguridad del Estado (FSCE), siempre han cobrado un papel indispensable en nuestra sociedad, encargadas en todo momento de procurar el cumplimiento de la ley, y velar por la seguridad de todos los ciudadanos. Actualmente, nos encontramos en una increíble e imparable revolución tecnológica, especialmente desarrollada gracias a los avances científicos, y de la que diferentes ámbitos como el empresarial, el sanitario o el de transporte se están, en buena cuenta, beneficiando. Con esto en mente, se hace imperiosa la necesidad de abordar la manera en que los medios empleados por las distintas policías se van adaptando y acompasando al nivel tecnológico en el que se encuentra la sociedad hoy en día (del mismo modo que han hecho el resto de sectores).

Una de estas innovaciones tecnológicas que ha revolucionado el mundo en los últimos años, es el Big Data (datos masivos). Como durante este trabajo se va a explicar, estas inmensas cantidades de información, almacenadas, procesadas y analizadas de la manera correcta, y con una metodología y unos fines criminológicos concretos, pueden convertirse en un gran aliado a la hora de prevenir el crimen.

A continuación, se procederá a explicar los objetivos principales que se han tratado de alcanzar con este texto, así como la metodología utilizada para ello.

1.1. OBJETIVOS

El presente trabajo tiene como objetivo principal analizar la manera en que la tecnología conocida como Big Data va introduciéndose en el mundo de la "guerra" contra el crimen, situando el foco en su modalidad proactiva. En otras palabras, se pretende estudiar cómo este tipo de tecnología puede ayudar a las diferentes instituciones a prevenir la delincuencia, explorando para ello, su funcionamiento técnico, sus múltiples usos, así como sus virtudes y defectos.

Con esto en mente, en primer lugar, se describirá el marco contextual en el que nos encontramos; explicando el papel que tiene la prevención del delito dentro de las

políticas criminales; observando el camino recorrido y, señalando sus diferentes modos de actuación, empezando por el Derecho Penal y finalmente abordando los distintos medios extrapenales existentes. Además, se realizará una exposición de las diferentes teorías criminológicas que apoyan la premisa preventivo-criminal, así como de las distintas hipótesis que sustentan la idea de que, metodologías como las que propone el Big Data, pueden ser muy eficaces en secundar el mencionado objetivo.

En segundo lugar, se conceptualizará el término Big Data y se desarrollarán sus características y elementos principales, buscando así facilitar la comprensión de este tipo de tecnología y el entendimiento de la misión principal de este trabajo.

No obstante, estas implementaciones tecnológicas no están libres de costes, pues en la mayoría de los casos, en aras de la seguridad pública, se vulneran multitud de derechos y libertades individuales. Por ello, como objetivo final, en este texto se estudiará, desde un enfoque ético-jurídico, la controversia que rodea la inclusión del Big Data en la agenda de la política criminal.

1.2. METODOLOGÍA

La metodología que se ha utilizado en este trabajo consiste en una extensa revisión bibliográfica centrada tanto en explicar la conceptualización, contextualización y teorización de los términos relevantes, como en profundizar en las distintas implicaciones que tiene el Big Data como herramienta de prevención de la delincuencia, así como sus controversias ético-jurídicas. Toda la búsqueda bibliográfica se ha realizado a través de bases de datos especializadas que cuentan con una gran variedad de textos científicos y jurídicos. Las plataformas utilizadas (situadas por orden de información proporcionada) han sido las siguientes: *Google Académico, Dialnet, ScienceDirect, Scielo, ResearchGate, y Proquest*. Cabe mencionar que algunos de los documentos ha sido posible conseguirlos por medio de la Biblioteca de la Universidad de Alicante y especialmente gracias a la herramienta RedUA.

Para indagar dentro de las plataformas académicas y ubicar los textos que eran relevantes para el objeto de estudio se han utilizado las siguientes palabras clave: *prevención del delito, teorías de la pena, formas alternativas de prevención, prevención*

situacional, Big Data, data mining y machine learning, legislación en protección de datos y valores éticos.

Además, para reforzar esta revisión se ha acudido a los materiales docentes proporcionados por el grado de Criminología de la Universidad de Alicante, en especial los facilitados por las asignaturas de "Prevención y Tratamiento de la Delincuencia", "Derecho Penal: Parte General", "Sociología de la Desviación" y "Teorías Criminológicas". Estos han sido de gran utilidad como modo de contrastar información, a la vez que han sido fuente de autores y teorías.

Todos los textos consultados han sido analizados y de ellos se ha extraído la información más relevante para los distintos apartados de este estudio. Hay que señalar que en castellano había abundancia de textos jurídicos y sociológicos relacionados con la prevención de la delincuencia y sus teorizaciones, pero, sin embargo, en lo que respecta al Big Data y sus implicaciones en el mundo de la seguridad, la mayoría de los documentos han sido proporcionados por fuentes de anglosajonas. Esto es una evidencia más de la cantidad de trabajo e investigación que aún es necesaria en esta materia.

2. MARCO CONTEXTUAL Y TEÓRICO. IMPORTANCIA DE LA PREVENCIÓN DEL DELITO

Entender la importancia que tiene la prevención del delito exige entender el fenómeno criminal como una concatenación de factores sociales e individuales generados por un extenso abanico de elementos y circunstancias que afectan e influyen en la vida de los individuos a razón del tiempo. Así, el delito tendría su génesis en la unión de diferentes factores de toda índole sobre un sujeto (o sujetos) determinado (estos dependerán de las características del hecho y del individuo). Y es aquí donde encuentra su justificación la prevención del delito. Conseguir comprender los diferentes tipos de causas que influyen en los comportamientos criminógenos, puede dar lugar a la elaboración de una serie de programas y estrategias dirigidos especialmente a bloquear la aparición de los diversos factores de riesgo antes de que se materialicen como delitos (Shaw, 2011).

De este modo, la conceptualización del término "prevención de la delincuencia" ofrecida por la ONU en su resolución 2002/13 del Consejo Económico y Social, en la página 64, versa de la siguiente manera: "*las estrategias y medidas encaminadas a reducir el riesgo de que se produzcan delitos y sus posibles efectos perjudiciales para las personas y la sociedad, incluido el temor a la delincuencia, y a intervenir para influir en sus múltiples causas*". Es palpable la importancia que tiene planificar estrategias de prevención dirigidas a evitar dichas infracciones y sus respectivos efectos perjudiciales. Sin ello, no solo sería imposible proporcionar un alto grado de protección sobre la población, sino que sería imposible paliar el miedo social al delito y promover un sentimiento de seguridad entre la ciudadanía, elementos indispensables para el desarrollo sostenible de los países (ECOSOC, 2002).

Además, la ONU (2002), en la misma resolución, otorga la responsabilidad de llevar a cabo dicha prevención al "gobierno", el cual se tiene que encargar de promover políticas sociales, económicas, criminales, educativas y sanitarias enfocadas al tratamiento de la delincuencia. Esto pone de manifiesto una vez más la suma importancia que debe adquirir la prevención criminológica dentro de las medidas que adopta el Estado. Desde hace ya bastante tiempo, no viene siendo válida como única respuesta contra la delincuencia el modelo reactivo, es decir, actuar una vez se ha cometido el delito, sino

que se exige a todo órgano de poder que, además, vuelque sus esfuerzos en combatir la delincuencia de una manera proactiva (antes de que se cometan los delitos). Solo así se puede lograr que la ciudadanía tenga un grado de seguridad notorio y palpable propio de toda sociedad desarrollada.

No satisfecha con eso, la ONU también deja clara la disruptiva que existe con el enfoque clásico, el cual reservaba de manera exclusiva los derechos de perseguir la delincuencia al terreno punitivo estatal. Así, evidencia la importancia de aplicar una estrategia multidisciplinar a la hora de combatir el crimen que aborde todos los espectros sociales. Por ello, a continuación, se analizará la importancia de la prevención, no solo desde sus orígenes (la pena), sino también desde sus diferentes ámbitos de actuación alejados del sistema penal.

2.1. LA PENA COMO MÉTODO DE PREVENCIÓN

El análisis de la pena desde un punto de vista preventivo debe separarse en dos apartados distinguidos para así evitar una posible confusión de conceptos e ideas. En primer lugar, es preciso explicar la función que cumple la pena en nuestra sociedad a través de las distintas teorías que abordan la materia, con el fin de dilucidar si la pena tiene un fin preventivo o, por el contrario, solo cuenta con una finalidad retributiva y de castigo. En segundo lugar, una vez analizadas si las finalidades de la pena son preventivas o no, es necesario estudiar en qué medida ocurre esta presumible prevención y si es el método más eficaz o, por el contrario, necesita ser complementado con otras técnicas alternativas alejadas del conglomerado penal.

2.1.1. La función de la pena

Todo Estado social y democrático de derecho debe tener como obligación indispensable asegurar la protección efectiva de todos los miembros de su comunidad. Es aquí, en el ámbito de la seguridad, donde el Estado, como poseedor exclusivo del *Ius Puniendi*, otorga esa responsabilidad tan fundamental al Derecho Penal. De un modo claro, no se puede completar la definición de seguridad sin otorgar un papel clave a la prevención de la delincuencia. Por lo tanto, un Derecho Penal constituido en el seno de un Estado social y democrático de derecho debe orientar la pena hacía su función

preventiva, aunque no de un modo estrictamente único, y siempre con arreglo a los principios de proporcionalidad, culpabilidad y exclusiva protección de bienes jurídicos, (Puig, 1982).

Tal es así que, en la actualidad, y a lo largo de la historia, el método de prevención de la delincuencia más utilizado y en el que se han depositado más esperanzas, incluso por Estados no democráticos, y a mi parecer de una manera equivocada, es el punitivo Estatal. De este modo, la pena, entendida como la autoconstatación propia del poder represivo del Estado (Ramírez y Mallafré, 1980), no solo cumpliría una función retributiva, como abogaban Kant y Hegel al defender la función de la pena desde un punto de vista absolutista, sino que, además, de un modo intencionado o no, cumpliría un fin preventivo (Teoría de la unión) (Angulo y López, 2001).

Esta teoría, también denominada ecléctica o mixta, pretende conciliar y unificar bajo un mismo espectro las clásicas teorías absolutas[1] y relativas[2], es por ello, por lo que no sorprende que dentro de la misma convivan diferentes vertientes. No obstante, todas

[1] Estas, defienden la aplicación de la pena con un exclusivo fin retributivo, es decir, la imposición de un mal por un mal cometido. Hay que hacer hincapié en que, como se ha mencionado anteriormente siguiendo la doctrina de Puig (1982), los absolutistas no tienen cabida en un Estado social y democrático de derecho pues uno de los fines de este, sería la ansiada búsqueda del estado de bienestar, lo que conllevaría necesariamente un notable grado de seguridad, imposible de conseguir olvidándose de la prevención de los delitos. Por lo tanto, como bien expresaba Roxin (1976), no se puede concebir un Derecho Penal que imponga sus penas con el objetivo de proteger los bienes jurídicos, sin que para llevar a cabo esta tarea "*prescinda de toda finalidad social".* Sin embargo, esta teoría, aunque criticada en su mayoría por un amplio sector de la doctrina (Durán-Migliardi, 2011), también cuenta con reseñas positivas en su intento de contribución al ámbito del Derecho Penal. De esta manera, el hecho de que la pena sea entendida como la retribución de un mal causado en la búsqueda de una justicia efectiva, obliga al Legislador a que esa retribución sea estrictamente proporcional al daño cometido, en tanto en cuanto, la pena solo sería justa y por lo tanto legítima si se adecua a la culpabilidad del autor (principio de proporcionalidad). Como tal, el ejercicio del Derecho Penal y con él, el de las penas, supone un límite al ius puniendi del Estado (Marín de Espinosa-Ceballos et al., 2016).

[2] Estas, surgen de una manera reactiva a las absolutistas: dotando de una finalidad ulterior a las penas, eliminando la concepción clásica de que la pena es un fin en sí misma, y dando protagonismo a la punición como medio de prevención encargada, no tanto de impartir justicia, sino simplemente de proteger a la sociedad (Cárdenas, 2004). En este caso, la pena entendida como un mal aplicado a una persona concreta encontraría su fundamento en conseguir fines de valor positivo para la sociedad, como la disuasión de futuros delitos (Rodríguez, 2019). Estas teorías influyen en la prevención del delito de diferentes maneras ya que, son dos las áreas o aspectos sobre los que la prevención puede incidir por medio de la pena. Así, se puede distinguir entre las teorías de la prevención general y las teorías de la prevención especial. Las primeras se pueden definir como la "*función y fin de la pena que se dirige a evitar que los ciudadanos, en general, cometan delitos*". Y, por otro lado, la prevención especial se explica cómo: "*función y fin de la pena y de la medida de seguridad que se dirige a evitar que el sujeto infractor cometa nuevos delitos*" (Diccionario panhispánico del español jurídico [DPEJ], 2020). Por lo tanto, la clave para diferenciar entre ambas teorías relativas reside en el sujeto o sujetos (el objetivo) a quien se orienta la prevención delincuencial. Mientras que la teoría de la prevención general va dirigida al conjunto de la población, la teoría de la prevención especial se orienta únicamente hacía aquellos individuos que han cometido ya un crimen.

ellas coinciden en que, en palabras de Puig (2003), se admite la retribución del mal causado siempre y cuando esté orientada a la protección de la sociedad, es decir, tenga además fines preventivos. Y, del mismo modo, el objetivo de la prevención es aceptado, en tanto en cuanto la pena se encuentre sujeta a los límites que proporciona la justa retribución del daño (principio de proporcionalidad y culpabilidad).

El debate y las críticas surgieron pues, entre las diferentes corrientes que se fueron formando ya que discrepaban entre que aspecto merece más importancia. Ante esto, surgió una idea de pensamiento bien distinta, protagonizada por Roxin y Schmidhäuser, que abogaba por que la pena no cumple un fin único, sino que es definida como un concepto ambivalente que reparte el peso que otorga a cada aspecto en función de la fase en la que nos encontremos (Puig, 2003). Por ejemplo, plantean dar protagonismo a las teorías de la prevención general en la fase de intimidación sobre todo el conjunto de la sociedad. Posteriormente, una vez cometido el delito, en la fase de sentencia e imposición de una pena, deben prevalecer siempre los principios de proporcionalidad y culpabilidad derivados de las pretensiones absolutistas. Por último, en la fase de la ejecución de la pena, se exige dar prioridad a la prevención especial para así buscar la resocialización efectiva del delincuente.

Todo esto permite admitir que, de la pena, sin restar valor a su finalidad retributiva, sí que se deduce un fin intrínsecamente preventivo. Así, la bibliografía estudiada ha dejado claro de qué modo se manifiesta esta prevención, tanto para que el conjunto de la población se abstenga de cometer delitos (prevención general) como para que los que ya los han cometido no vuelvan a reincidir (prevención especial). Esto explica y justifica que haya sido el método más utilizado durante años.

2.1.2. Eficacia preventiva de la pena

A continuación, una vez se ha dejado claro que la pena sí que tiene una función preventiva y, teniendo en cuenta la referencia ya mencionada con anterioridad acerca de la exigencia de la ONU (2002)[3] sobre la implementación complementaria de medidas

[3] Véase ECOSOC (2002). Medidas para promover la prevención eficaz del delito. Resolución 2002/13 del Consejo Económico y Social.

extrapenales de toda índole para combatir el crimen; se pretende ahora analizar en qué grado la pena es un mecanismo eficaz y útil en la prevención de la delincuencia.

Los defensores del sistema penal como principal mecanismo de prevención asumen la siguiente premisa: "a mayor gravedad de las penas menor es la delincuencia". Parten de la idea de que una mayor *"imposición del orden"* tiene un mayor grado de poder intimidatorio sobre la población, logrando que se abstengan de cometer delitos (Fernández-Vega, 2017). Sin embargo, sin negar la posible prevención general negativa que puede conseguirse a raíz de estas medidas, los estudios muestran que esto no es así.

Para entender mejor este postulado, es importante traer a colación las bases criminológicas en las que se apoya, respecto a las cuales Pires (2007) hace un análisis muy certero. En primer lugar, este alegato se basa en la idea de la individualización del ser humano como un sujeto racional, que acorde a su pensamiento decide que acciones llevar a cabo. Es decir, se apoya en la teoría de que el delito es un problema ajeno a la comunidad, en el que solo influye la estructura individual del sujeto. Teniendo esto en cuenta, en segundo lugar, los partidarios del método punitivo consideran que este pensamiento racional que lleva al sujeto a decidir qué acciones realiza se basa en una ponderación de costes y beneficios (teoría de la elección racional). El delincuente a la hora de cometer el hecho valoraría los costes que conlleva realizar esa acción (la pena) y los beneficios que se reportarían de la misma. Así, si en este cálculo, la balanza se inclina para el lado de los beneficios se perpetraría el crimen, pero si se inclina en lado contrario se abstendría de hacerlo. Por lo tanto, la solución pasa por aumentar los costes de cada acción delictiva, es decir, aumentar la severidad de las penas para que así estos siempre superen a los beneficios.

No obstante, autores como Dubé (2012) expusieron que la ponderación de costes y beneficios no es la única racionalización que se lleva a cabo a la hora de valorar si se delinque o no. De este modo, puso de manifiesto una variable más, "*la teoría del riesgo*", según la cual, en función del riesgo, el coste se infravalora en mayor o menor medida a la hora de tomar la decisión, es decir, pese a que el coste está presente, este no se asume como algo real y fáctico ya que no es algo seguro, sino que simplemente se percibe como una posibilidad. Este contrapeso debilita la importancia que el delincuente otorga al coste, desnivelando la balanza a favor del beneficio. Por lo tanto, según esta teoría, por más que aumentes la pena, esta se seguirá percibiendo en función del riesgo o de las posibilidades

de ser castigado que el autor asocie (criterio subjetivo) a cada acción, por lo que no tendrá el efecto disuasorio esperado. Además, se plantea la posibilidad de que estos no sean los únicos factores que influyen en la decisión, dando paso también a posibles valoraciones morales sobre la legitimidad del sistema o sobre la forma en que éste está organizado (Ortiz de Urbina, 2004). Si a ello se le suman las extensas teorías criminológicas que abogan por que la criminalidad no es un factor relacionado exclusivamente con el individuo, sino un problema comunitario influido tanto por factores psicológicos como sociológicos (teorías sociológicas del crimen)[4] (Pérez, 2011), obtenemos una posible explicación acera de por qué el modelo punitivo en exclusiva no es tan eficaz a la hora de prevenir la delincuencia.

En la misma línea, el penalista Sergi Cardenal Montravelta (2015), establece dos requisitos para que la pena tenga eficacia como arma disuasoria. En primer lugar, que los costes del delito sean superiores a los beneficios del mismo. Y, en segundo lugar, el grado de conocimiento que se tenga sobre la cuantía de las penas y la consideración de tener esa información en cuenta a la hora de valorar. El primero de los requisitos, como se ha mencionado anteriormente, no se consigue legislando un incremento de la severidad penal, ya que este no conlleva necesariamente el aumento correspondiente del coste. Con respecto al segundo requisito, y pese a que el propio Montravelta (2015) asume que no importa tanto el conocimiento sobre la sanción real sino la idea que el delincuente tenga sobre la posible pena, pudiendo ésta afectar de manera positiva o negativa al coste; es obvio que un aumento en la penalidad de los delitos no podrá tener efecto en la ponderación de costes y beneficios si éste no cala previamente en la conciencia de la sociedad. En este asunto, Lazo (2018) desvela la clara *"barrera endógena*" que existe entre la norma y el ciudadano, provocada a su vez por una sobreproducción legislativa[5] y por fallos en la publicidad de esta[6].

[4] Destacan autores como Ferri y Lacassagne, véase Pérez, J. A. (2011). La explicación sociológica de la criminalidad. *Derecho y cambio social*, *7*(22), Recuperado de http://www.derechoycambiosocial.com/revista022/explicacion_sociologica_de_la_criminalidad.pdf .

[5] El Código Penal español ha sufrido 32 reformas legislativas desde 1995. Véase Mellón, J. A., Jiménez, G. A., y Rothstein, P. A. (2017). Populismo punitivo en España (1995-2015): presión mediática y reformas legislativas. *Revista española de ciencia política*, (43), 13-36.

[6] Si bien las normas se publicitan en el Boletín Oficial del Estado y es responsabilidad del ciudadano informarse del contenido de dichas leyes, en la práctica esto no suele ser lo habitual sobre todo entre la población lega en Derecho.

Prueba de toda esta teorización es el trabajo empírico realizado por Smith et al. (2002)[7], en el cual se analizaron 117 estudios, con una muestra de 442.741 delincuentes, sobre la influencia que tienen los diferentes tipos de condena en la reincidencia. Las conclusiones reflejan que las condenas más duras no se relacionan con unas tasas de reincidencia más reducidas (Garrido, 2010).

Con todo esto, se puede concluir que el endurecimiento de las condenas no es una medida viable para reducir la delincuencia. Por lo que vemos lógico aceptar que la prevención general obtenida por medio del Derecho Penal está limitada, es decir, hay un cupo de población que se abstiene de cometer delitos gracias al conglomerado penal, pero sigue habiendo otro porcentaje de sujetos a los que este tipo de medidas preventivo-punitivas no afecta por mucho que se endurezcan las penas. Esto, abre la puerta al planteamiento de opciones alternativas, alejadas del sistema punitivo y que sí que puedan aportar valor en la lucha por la prevención del delito.

2.2. FORMAS ALTERNATIVAS DE PREVENIR LA DELINCUENCIA

Bien es sabido que la prevención del delito no debe agotarse en el Derecho Penal. Para que una sociedad cuente con un plan estructurado y organizado de acción (no reacción) contra la delincuencia es necesario que las medidas abarquen ámbitos de todo tipo. Pero antes, es necesario realizar un inciso acerca del concepto de delincuencia y las teorías del comportamiento. De acuerdo con estas, los individuos se ven influenciados en su conducta a través del contexto social, ambiental y físico que experimentan (Kitchen y Schneider, 2007). Por lo que se entiende que, poner el foco en todos estos elementos externos, actuará de un modo beneficioso en la conducta proactiva de la población. Así, el ECOSOC (2002, p.64) no centra el foco en el sistema penal a la hora de prevenir la delincuencia, sino que, destaca la importancia de que las medidas de prevención del crimen se centren en "*promover el bienestar de las personas y fomentar un comportamiento favorable a la sociedad mediante la aplicación de medidas sociales, económicas, de salud y de educación*". Del mismo modo, destaca la necesidad de hacer hincapié en distintos enfoques preventivos como son: el diseño ambiental, la prevención

[7] Véase Smith, P., Gendreau, P., & Goggin, C. (2002). The effects of prison sentences and intermediate sanctions on recidivism: General effects and individual differences. Ottawa, ON: Solicitor General Canada.

de las situaciones que propician los delitos y los programas de tratamiento con delincuentes.

El enfoque económico y de desarrollo social como medio para reducir la delincuencia se basa en la idea de que todo el entorno familiar, educativo, económico, sanitario y social influyen en el comportamiento de los individuos. Todos estos ámbitos, dependiendo de sus características, pueden ser considerados o bien como factores de riesgo, o bien como factores de protección. Así, primero reconociendo los factores de riesgo y luego aplicando medidas destinadas a minimizarlos se reducirían las probabilidades de que un sujeto delinca. En especial, como factores de riesgo asociados a este enfoque destacan: la personalidad del individuo, las influencias familiares (maltratos, influencias antisociales, etc.), las condiciones de vida (mala situación económica, pobreza), la educación (abandono escolar, poca motivación), el grupo de iguales (influencias antisociales) y las oportunidades laborales (precariedad, dificultades para encontrar empleo) (Sozzo, 2000). De esta manera, resultan de gran ayuda mecanismos que reconozcan estos factores de riesgo y que actúen sobre ellos, paliándolos y provocando como resultado una situación mucho más prosocial para el individuo.

Prueba de ello es un estudio denominado "*High/Scope Perry Pre-School Project"* desarrollado en 1962 por el psicólogo estadounidense David P. Weikart, en el que se partió de una muestra de 123 niños afroamericanos de entre 3 y 11 años que contaban con factores de riesgo relacionados con el ámbito familiar y educativo. Los menores fueron divididos en dos grupos, uno formado por 58 alumnos que recibiría un programa de desarrollo infantil especializado, y otro de 65 alumnos que formaría el grupo de control. Los resultados indicaron que, de los jóvenes que recibieron el programa de educación, solo contaban con detenciones un 7%, mientras que de los sujetos del grupo de control, un 20% habían sido arrestados alguna vez. Esto es solo una muestra de los resultados preventivos que se pueden obtener invirtiendo en políticas de índole social, económico y educativo, entre otras.

En segundo lugar, el diseño ambiental busca aumentar la sensación de seguridad y prevenir de los delitos que tengan su génesis en el ordenamiento urbanístico de las ciudades. Esto se debe a que, en palabras del arquitecto Oscar Newman (1995), esta organización territorial ha provocado la creación de espacios de exclusión sometidos a la delincuencia y a la marginación social. Así, políticas urbanísticas como una correcta

configuración de usos, donde el espacio es utilizado en distintas materias y con variedad de horarios, provoca una mayor interacción social y facilita un entorno con altos niveles de presencialidad. Esto evita el abandono y deterioro del espacio, factor claro de protección frente a la delincuencia. Asimismo, un área que cuente con mejor visibilidad supone un control preventivo ya que aumenta la posible vigilancia de la zona evitando comportamientos delictivos. Del mismo modo, los territorios que provocan un apego más intenso por parte de sus ciudadanos son más respetados por sus habitantes, por lo tanto, incentivar el desarrollo de medidas orientadas a conseguir un mayor afecto puede provocar un desplazamiento de la delincuencia (Fernández-Vega, 2017).

En tercer lugar, la ONU exige centrar los esfuerzos en prevenir las situaciones que propician los delitos, lo que se ha denominado por la doctrina criminológica como prevención situacional. Éstas, tienen como objetivo reducir las oportunidades delictivas en un área concreta. Es por ello por lo que están estrechamente relacionadas con las teorías ambientales o ecológicas, una de las maneras de reducir las situaciones propicias de criminalidad consiste en alterar las características del entorno urbano, como se ha explicado anteriormente. Además, es importante mencionar, para así no caer en glorificaciones ni ideas equivocadas, que estas políticas tienen un objetivo concreto: actuar sobre la situación, no sobre la persona, es decir, las medidas se diseñan en gran parte para delitos de carácter oportunista y despersonalizados (Ariza, 1998).

Esta manera de prevenir la delincuencia basa sus premisas en dos teorizaciones muy importantes. Por un lado, en la teoría de la elección racional en la que se concibe al criminal como un ser racional, que tras un proceso de pensamiento objetivo y voluntario contempla la clásica ponderación de costes y beneficios y opta por delinquir (con todos los matices que ello conlleva y que fueron explicados con anterioridad). Y, por otro lado, a esta idea se le une la teoría de las actividades rutinarias, según la cual, el delito es fruto de una situación determinada en la que confluyen un ofensor motivado, una potencial víctima y la ausencia de un guardián capaz (Sozzo, 2000). Por lo tanto, si tenemos en cuenta todos estos factores: 1) la elección de delinquir se basa en una “elección racional” y, 2) los delitos surgen cuando se da una situación determinada; se pueden idear medidas de prevención extrapenales que vayan dirigidas a incrementar los costes de cometer un ilícito penal y a minimizar al máximo la aparición de esas oportunidades delictivas.

Con esta base teórica se han categorizado cinco objetivos a los que tienen que ir dirigidos las distintas técnicas de prevención situacional. Así, acorde a la tabla establecida por Cornish y Clarke en 2003, encontramos las siguientes estrategias (Summers, 2009; ONU, 2011):

- Aumentar el esfuerzo, que el delito sea más difícil de cometer o que lo parezca (Vehículos antirrobo, controlar accesos y salidas, etc.);
- Aumentar el riesgo, que el delito sea más fácil de detectar (aumentar el número de guardianes, mejorar la iluminación, etc.);
- Reducir las ganancias, que el delito aporte menos beneficios (contenedores de tinta roja, deshabilitar móviles robados, etc.);
- Reducir la incitación, los elementos emocionales que pueden conducir al delito (evitar disputas, neutralizar la presión de grupo, etc.); y,
- Suprimir las excusas, aumentando los sentimientos de culpabilidad y vergüenza al delinquir y facilitando la elección de acciones alternativas (campañas de tráfico, fijar prohibiciones, etc.).

Por último, caben destacar las medidas de prevención de la delincuencia extrapenal para sujetos que ya han pasado por el propio sistema punitivo. Ante las extensas evidencias que reflejan las carencias del sistema penal como medio de prevención especial, surgen nuevos tratamientos y programas dirigidos a la resocialización efectiva del penado. La continua búsqueda de respuestas ha fomentado el desarrollo de estudios que han demostrado la existencia de programas de intervención que llegan a reducir la reincidencia en un 20% frente al grupo de control (Garrido, 2010)[8].

En la actualidad y cada vez con más frecuencia según van trascurriendo los años, todas estas formas de prevención, en especial la prevención situacional, empiezan a contar con un aliado común que ejerce un apoyo cualitativo y cuantitativo importante. Con esto me refiero a la irrupción de las tecnologías en nuestra sociedad y, como no, en el mundo de la lucha contra la delincuencia el cual no acostumbra a quedarse atrás. De esta manera, estamos siendo testigos de cómo poco a poco van aportando su servicio a la ansiada búsqueda de seguridad pública. Así, utilizando las tecnologías en nuestro beneficio

[8] Un ejemplo de ello es el estudio realizado por los criminólogos Vicente Garrido y Santiago Redondo sobre la eficacia de los programas de tratamiento en delincuentes sexuales. Véase Illescas, S. R. (2006). ¿Sirve el tratamiento para rehabilitar a los delincuentes sexuales? *Revista española de investigación criminológica*, *4*, 1-22.

aumentamos los mecanismos de protección y de detección de ilícitos penales, a la vez que, influimos en el aumento del riesgo y la disminución de los beneficios. En general, estas tecnologías agilizan, simplifican y mejoran la labor preventiva, donde antes tenía que estar un vigilante ahora puede ubicarse una cámara de seguridad, esto conlleva un incremento sustancial de la capacidad preventiva sin un aumento proporcional de los medios personales empleados (con mismos medios humanos se puede mejorar la eficacia preventiva). Por ejemplo, un avance reseñable es la incorporación de Sistema de Información Geográfica (SIG), los cuales suponen una revolución y un gran salto en la gestión de la seguridad ya que son capaces de trabajar con inmensas capacidades de información (Iancu, 2016).

En conclusión, podemos afirmar que las medidas de prevención extrapenales superan en multitud de ocasiones la prevención generada por el sistema penal, no se trata con esto de superponer una a otra, sino de que actúen como herramientas complementarias en la búsqueda de la seguridad ciudadana. Un buen plan de prevención debe contener como base fundamental una estrategia socioeconómica estable, con políticas inclusivas que permitan a la sociedad desarrollarse en su proyecto individual de vida de un modo prosocial; unas apropiadas y constantemente innovadoras técnicas, tanto situacionales como ambientales, que reduzcan las oportunidades delictivas antes de que materialicen; unos correctos programas de reinserción social que disminuyan la reincidencia; y, un Derecho Penal alejado del populismo punitivo y dirigido a resocializar al delincuente. Con este contexto presente, se pondrá el foco sobre como las tecnologías innovadoras como el Big Data pueden aportar al mundo de la prevención del delito, en especial al sector de la prevención situacional. De este modo, se abordará en que consiste esta herramienta y como puede ser útil para la prevención de la delincuencia, así como su posible encaje ético y jurídico.

3. CONCEPTUALIZACIÓN DEL BIG DATA

La tecnología ha irrumpido en nuestras vidas como un "elefante en una cacharrería", cambiando, alterando y afectando todo lo que hay a su alrededor. Los que han vivido esa irrupción lo considerarán un proceso lento y gradual, pero si lo comparamos con la historia de nuestra humanidad, ésta, la tecnológica, ha sido la revolución más rápida e influyente de todos los tiempos. Lo más impactante de todo es que parece no tener fin, el progreso que se va adquiriendo año tras año no muestra indicios de que en ningún momento se vaya a estabilizar, tal es así que las tecnologías que usábamos hace diez años ahora están obsoletas, anticuadas o son inservibles. Cada vez es menor la gente que no se adapta a esta era tecnológica, ya no se hace tan fácil de encontrar a alguien que se resista a "modernizarse" y a acomodar su vida de acuerdo con los nuevos avances tecnológicos.

Así lo refleja un estudio elaborado por el Cisco (2019) que estima que el número de dispositivos móviles conectados a Internet será de 1.5 por persona para 2022. Si la población a principios de 2021 era de más de 7.7 miles de millones de personas según el *U.S. Census Bureau*[9], para 2022 habrá más de 12.3 miles de millones de dispositivos conectados a Internet emitiendo datos. Toda esa ingente cantidad de dispositivos generarán en 2022 la friolera de 77.5 *exabytes* de datos al mes, lo que supone multiplicar por más de 6 la cifra que se alcanzó en 2017 al mes (12 *exabytes*). Para hacernos una idea de lo que eso implica, en 0,5 *exabytes* cabría una biblioteca digital de todos los libros que se han escrito a lo largo de la historia en cualquier idioma[10]. Por ejemplo, un estudio realizado en 2021 calculó que en un minuto había 200.000 personas escribiendo *tweets*, se enviaban 197.6 millones de correos electrónicos, se subían a Instagram 695.000 fotos y se escribían más de 69 millones de mensajes de texto[11]. Esto es solo una muestra de la cantidad de datos que se generan por parte de las personas al interactuar en Internet, pero no son los únicos, ya que también las propias tecnologías generan datos mediante sensores

[9] Véase https://www.census.gov/
[10] Véase Forecast, G. M. D. T. (2019). Cisco visual networking index: global mobile data traffic forecast update, 2017–2022. *Update*, *2017*, 2022. Enlace al estudio: https://davidellis.ca/wp-content/uploads/2019/12/cisco-vni-mobile-data-traffic-feb-2019.pdf
[11] Estudio elaborado por la consultora Cumulus Media. Véase https://www.allaccess.com/merge/archive/32972/infographic-what-happens-in-an-internet-minute

que transmiten información en tiempo real[12] (señal de GPS), así como datos de telecomunicaciones y biométricos (Barranco, 2012; Galimany, 2014).

Ante esta abrumadora masa de datos, las herramientas convencionales (bases de datos) han sido ineficaces y han devenido en obsoletas. Es por ello por lo que como solución a este problema aparece el Big Data, dirigido al almacenamiento, tratamiento y transferencia de todos estos datos que, por su complejidad, volumen, velocidad de crecimiento y demás características no pueden ser almacenados, gestionados ni analizados mediante los medios tradicionales (Méndez 2020). El Big Data no es otra cosa que "*la colección de tecnologías y estrategias capaces de capturar y analizar, de forma económica, grandes volúmenes de datos provenientes de múltiples fuentes heterogéneas a una alta velocidad*" (Puyol, 2014, p.482). De este modo, todos estos datos que se generan no se pierden, ni se eliminan, sino que van a parar a inmensos centros de datos (*data centers)* formados por un gran número de servidores con discos duros de altísima capacidad. A modo de ejemplo hay que mencionar que el *Grupo Fractalia* (2016)[13] situaba en más de 2.000 el número de *data centers* disponibles en el mundo, de los cuales 36 están ubicados en España.

Ha sido necesaria una inversión tan grande en este entramado tecnológico debido a las complejas características comunes que cumplen tanto los datos objeto de almacenamiento y análisis, como el propio objetivo del Big Data. De este modo, hay autores que emplean el término de las tres "*uves*" (volumen, velocidad y variedad) (Laney, 2001) y otros que apuntan a que son cinco las "*uves"* que caracterizan a este tipo de datos (se añade veracidad y valor) (Puyol, 2014). Sin desmerecer a Douglas Laney, en este trabajo se optará por seguir la tipificación que aporta el Magistrado Javier Puyol Moreno en comparación con la valoración de distintos autores. Así, las características son las que siguen:

1) Volumen: Esta es la característica que más se asocia al Big Data, de ahí su nombre (macrodatos). Como se ha mencionado anteriormente, la cantidad de datos que se generan en el mundo superan ya ampliamente a la cantidad de datos que generaría una

[12] A esta tecnología se le denomina "Internet de las cosas" y se estimó que en 2012 había 30 millones de sensores con un crecimiento del 30% al año, lo que deriva en más de 100 millones de sensores interconectados actualmente. Véase Barranco Fragoso, R. (2012). ¿Qué es Big Data? *IBM Developer*. Obtenido de https://developer.ibm.com/es/technologies/data-science/articles/que-es-big-data/

[13] Véase Grupo Fractalia (2016). Big Data: ¿Dónde se almacena tanta información¿ [Blog]. Recuperado de https://fractaliasystems.com/big-data-donde-se-almacena/

biblioteca digital con todos los libros que se han escrito en la historia. Este volumen no para de crecer año tras año hasta tal punto que el 90% de todos los datos creados han sido en los últimos años (Firican, 2017). Por ello es necesario una estrategia y unas infraestructuras tecnológicas adecuadas que sean capaces de guardar todo este volumen de información.

2) Velocidad: Muy relacionada con la característica anterior, la velocidad hace referencia a la elevada frecuencia con que los datos van siendo generados, almacenados y analizados. En el mundo en el que vivimos se requiere una constante instantaneidad que hace imprescindible tener todos los datos generados disponibles al instante para ser evaluados y utilizados. Toda esta velocidad de actuación resulta imposible para los sistemas tradicionales, así un ejemplo de la frecuencia con la que actúan aplicaciones como Google en la actualidad sería cuando predice las palabras que vamos a escribir en su buscador en cuestión de segundos (Galimany, 2014).

3) Variedad: Quizá una de las características que más diferencia al Big Data de las herramientas convencionales. Para entender mejor esto es necesario primero matizar la diferencia entre datos estructurados y no estructurados. Los primeros son datos bien definidos en su longitud y formato y tienen una estructura fija, es decir, son fáciles de analizar para los sistemas antiguos (por ejemplo: los resultados de un cuestionario). Los segundos no tienen un formato único, y es difícil estructurarlos y situarlos en modo de tabla, es decir son datos más de carácter cualitativo, por ejemplo, archivos de audio, video, texto, etc. (Hoferek, 2019). Las bases de datos tradicionales no tienen ningún problema analizando datos estructurados, es decir, datos provenientes de una sola fuente y de manera organizada, sin embargo, en la actualidad se requiere un análisis de datos más potente que provenga de fuentes muy distintas y heterogéneas. Así, el Big Data se desmarca siendo capaz de almacenar y analizar múltiples tipos de datos, ya sean estructurados o no estructurados (Galimany, 2014).

4) Veracidad: Esta característica se refiere a la fiabilidad de los datos recogidos, la confianza que se le otorgue a la información indicará el grado de validez de los análisis y resultados obtenidos (Hoferek, 2019). Esta veracidad va enfocada sobre todo a la fuente de donde se obtuvieron los datos, el punto de mira se centra en la metodología empleada para su recogida (Firican, 2017). Para conseguir este alto grado fiabilidad se utilizan

métodos de limpieza de datos (fusión de datos[14] y matemáticas avanzadas), sin embargo, es imposible hacer un cribado totalmente efectivo ya que hay variables que cuentan con una "*imprevisibilidad inherente"* como por ejemplo el clima o la economía (Puyol, 2014). Es por ello, por lo que hay que aprender a lidiar con ese porcentaje de datos de dudosa fiabilidad y adaptarse a él, "*la necesidad de reconocer y abordar esta incertidumbre es una de las características distintivas de Big Data"* (Paredes-Moreno, 2015, p.43). Un ejemplo de falta de veracidad sería aceptar por buenos los datos obtenidos a través de una cuenta de Instagram con un millón de seguidores, pero de los cuales el 75% son *bots*[15].

5) Valor: Esta es una de las características más relevantes, hasta tal punto que el valor se configura como el fin ulterior que se pretende conseguir con todo este conglomerado tecnológico. Ninguna de las anteriores características tendrá ningún sentido si no se acaba obteniendo un valor, ya sea económico o social, de los datos captados. Así, Nocetti (2017) afirma que gestionar y almacenar información es importante, pero esto, como tal, no proporciona ventaja alguna, la clave reside en el valor que se puede obtener a partir de esos datos. Como tal, mediante el análisis de todos los datos recogidos se pretende extraer unas conclusiones que otorguen valor a la empresa o institución que lo demande. Del mismo modo, Colmenarejo (2018) afirma que los datos, de forma aislada, no tienen importancia o beneficio alguno, pues de ellos no podemos sacar conjeturas fiables, es cuando son analizados en común cuando de verdad cobran un valor representativo. Autores como Solove (2007) denominan "*agregación*" al proceso de extraer conclusiones (valor) acerca del perfil de un usuario partiendo de los datos obtenidos sobre él. En la misma línea, para denominar a este proceso de extracción de valor a partir de los datos en "bruto" se utiliza el término anglosajón "*Knowledge Discovery in Databases"* (KDD) (Santos et al., 2006). Y es justo ahí donde el Big Data destaca y se consolida como unas de las nuevas oportunidades tecnológicas capaces de mejorar el rendimiento y la toma de decisiones de los diferentes sectores implicados (Puyol, 2014).

[14] La fusión de datos es *"el proceso de detección, asociación, correlación, estimación y combinación de datos en varios niveles, que provienen de diferentes fuentes, como: sensores, bases de datos, bitácoras, observaciones, señales e incluso decisiones*", con el objetivo de conseguir una mayor precisión en los datos (Muñoz, et al., 2017, p.34).

[15] En español "robots", hace referencia a la idea de que no hay personas reales detrás de cada cuenta en las redes sociales, sino que son perfiles falsos manejados por un *software*. Se estima que más del 10% de las redes sociales y un 62% del tráfico de internet está siendo generado por *bots* (Varol, et al., 2017).

Hay autores, incluso, que tipifican cinco elementos característicos más, llegando hasta un total de diez (“*uves”)*. Uno de estos autores es el Director de Gobernanza de Datos e Inteligencia Empresarial de la *University of British Columbia*, George Firican (2017), el cual establece, además de los ya mencionadas, la visualización, la volatilidad, la vulnerabilidad, la validez y la variabilidad, como características intrínsecas de los datos relacionados con el Big Data.

Esta tecnología ha hecho posible que la información digital (estructurada y no estructurada) sea fácil de almacenar, procesar, distribuir y transmitir. La complejidad de analizar y sacar conclusiones de datos no estructurados superaba la capacidad de procesamiento de las anteriores tecnologías. Así, el Big Data, a su vez, se ve identificado como el gran almacén, capaz de albergar una cantidad ingente de información de todo tipo, el cual, además, se vale de herramientas complementarias como la minería de datos (*data mining*) o el *machine learning* en su búsqueda por extraer valor y, en definitiva, un sentido y un propósito a toda esa información (KDD). No podemos entender el Big Data sin este tipo de conceptos y elementos que otorgan verdadero significado a toda esta tecnología. En palabras de Suárez (2019), “poder no lo tiene quien tiene simplemente un banco de datos, sino quien sabe y conoce cómo sacarles el mayor provecho”. Para que se entienda mejor, el KDD es la búsqueda de conocimiento a través de los datos, pero para conseguir este fin son necesarias una serie de pautas entre las que podemos incluir la preparación de los datos (“*Data Warehouse”*), su limpieza y su interpretación. Entre estos pasos es donde podemos ubicar las estrategias de *data mining* y *machine learning* (Santos et al., 2006).

De esta manera, se puede definir el *data mining* como un proceso de descubrimiento de correlaciones, perfiles y tendencias a través del análisis de los datos disponibles en los *data warehouse* utilizando algoritmos sofisticados y tecnologías de reconocimiento de patrones y redes neuronales, entre otras técnicas avanzadas de análisis de datos (Pérez-López y Santín-González, 2007). Si a esta idea le añadimos los conceptos de Inteligencia Artificial (IA) y aprendizaje automático tenemos como resultado el *machine learning*, el cual consiste en un perfeccionamiento automático mediante *IA* de los algoritmos creados para extraer patrones de los datos, en función de lo que van aprendiendo de cada procesamiento y operación. Esto incrementa el nivel de efectividad y precisión a raíz de la experiencia del algoritmo (Blum, 2003).

Haciendo una síntesis de ideas podríamos concluir que el objetivo del Big Data no es otro que ser capaz de almacenar, tratar, organizar y analizar una inmensa cantidad de datos fiables de toda índole a una gran velocidad, con el fin de extraer de los mismos una serie de conclusiones, análisis o patrones que permitan tanto comprender el entorno como adelantarse a él. Gracias a ello hay empresas por todo el mundo que consiguen entender mejor a sus potenciales clientes y son capaces de predecir cómo se va a configurar el mercado en los próximos tiempos. Por ejemplo, Google elabora perfiles comerciales de todos sus usuarios a través del estudio de los correos electrónicos que se escriben y se reciben desde su plataforma (información no estructurada). A ello le suman toda la información referente al contenido de las páginas webs que el usuario visita y toda la información personal y demográfica que obtienen, por ejemplo, mediante los archivos compartidos o a través de la ubicación de los dispositivos. De este modo, Google almacena, ordena y analiza toda esta información y consigue un "*retrato robot del consumidor*" que le permite influir en sus "*hábitos comerciales"* (agregación o KDD) (Harcourt, 2014)[16].

Además de estos beneficios privados que pueden obtener las grandes empresas mercantiles, también se pueden beneficiar de la tecnología del Big Data sectores de todo tipo como el de la salud, el de la energía y la sostenibilidad, el del transporte y el de la seguridad (el cual se tratará más adelante), entre otros (Parlamento Europeo, 2017a). En la misma línea, la Comisión Europea en 2014, señala como la información a través de los datos y sobre todo a través del Big Data está dando lugar a innovaciones en la tecnología, propiciando el desarrollo de nuevas herramientas y habilidades que sirven para un incremento del bienestar social a todas las escalas (Monleón-Getino, 2015).

[16] Véase "Governing, Exchanging, Securing: Big Data and the production of a digital knowledge", Public Law and Legal Theory Working Paper Group, Columbia Law School, 2014, pp. 4-5. Citado en Sancho-López, M. (2018). El derecho al olvido en el Big data: nuevos retos para la protección de la privacidad (Tesis Doctoral, Universitat de Valencia).

4. BIG DATA Y PREVENCIÓN DE LA DELINCUENCIA

4.1. NOCIONES PREVIAS

Desde hace ya varios años se viene haciendo notar la importancia de una revolución tecnológica en el ámbito de la seguridad pública, en especial en la prevención delincuencial (Vilalta-Perdomo, 2017). En contraposición, el sector privado siempre se encuentra en constante innovación, buscando insistentemente el avance y desarrollo tecnológico que le haga prevalecer y destacar frente a sus competidores. Esa es la razón por la que el Big Data ha encontrado tanta acogida en el mundo empresarial, prueba de ello es que tanto Google y Facebook como otros gigantes tecnológicos cuentan con sus propios centros de almacenamiento (data centers) y análisis de Big Data (Grupo Fractalia, 2016).

No obstante, en Europa, y mucho antes en Estados Unidos[17], ya se viene notando una cierta escalada progresiva en lo que al uso de Big Data en temas de seguridad se refiere (Cinelli y Gan, 2019). Y aunque de manera más reservada y a menor escala, tanto la Administración como, más en concreto, los Cuerpos y Fuerzas de Seguridad del Estado empiezan a sacar provecho también de sus propias remesas de datos. Así, tienen en su poder todo tipo de información relativa al historial delictivo de una zona determinada, el grado de presión policial, censos de población, sus características sociodemográficas y estadísticas judiciales y carcelarias, entre otros tipos de datos de toda índole (Vilalta-Perdomo, 2017). Incluso, tienen la capacidad de llevar a cabo nuevas metodologías de extracción directa de información sobre la población a través de la espontaneidad de las redes sociales, las cuales pueden ofrecer un reflejo de las características de una comunidad concreta.

Toda esta información permite sacar conclusiones y patrones estadísticos acerca de los comportamientos e interacciones que pueden tener los distintos individuos dentro de la comunidad (Van`t-Wout et al., 2019). Ya existen evidencias de como el análisis de

[17] En el año 2002 el FBI anunció que se valdría de datos comerciales, hábitos y preferencias de los ciudadanos para conseguir descubrir, a través de un perfil socio-criminológico, a potenciales terroristas antes de que pudieran atentar (Félix, 2002). Véase Fernández-Zalazar, D. C., y Guralnik, G. E. (2017). El fenómeno de data mining. Efectos psicosociales y en la subjetividad. In *IX Congreso Internacional de Investigación y Práctica Profesional en Psicología XXIV Jornadas de Investigación XIII Encuentro de Investigadores en Psicología del MERCOSUR*. Facultad de Psicología-Universidad de Buenos Aires.

todos estos datos ha dado como resultado una mayor eficacia en las políticas públicas a través de la prevención de los delitos (Vilalta-Perdomo, 2017). A este proceso, protagonizado por las Fuerzas y Cuerpos de Seguridad del Estado, se le denomina como "*predictive policing*" y consiste, en palabras de Perry et al. (2013), en "*la aplicación de técnicas de análisis, en particular técnicas cuantitativas, para identificar objetivos potenciales que requieren la intervención policial, además de prevenir delitos o resolver crímenes pasados mediante pronósticos estadísticos"* (p.1-2) . Este concepto descansa sobre el principio de repetición, según el cual los individuos repiten una conducta siempre y cuando esta les haya sido beneficiosa; por lo que el Big Data, al ser capaz de analizar grandes cantidades de datos relacionados con la población y sus incidentes delictivos, es una herramienta eficaz para extraer patrones sobre las características de los futuros actos criminales y así favorecer que una respuesta proactiva ante la delincuencia (Cinelli y Gan, 2019).

Además, a este término, algunos autores le añaden que es necesario una especie de "dimensión *mitológica"*, según la cual debe prevalecer la idea de que las estadísticas y el análisis de grandes cantidades de datos ofrecen unos conocimientos superiores sobre la realidad criminal que permiten extraer conclusiones inequívocas que antes eran imposibles (Castellanos, 2019). Sin embargo, como se verá posteriormente, esto ha dejado de ser una cuestión de "fe mitológica" para pasar a ser un hecho demostrado científicamente.

Con el avance del tiempo, cada vez tenemos más capacidad para extraer un mayor "groso" de datos e información de los distintos acontecimientos sociales. Como se ha expresado anteriormente, el volumen de datos creados por el ser humano está en un increíble aumento exponencial año tras año. De esta tendencia alcista la Criminología no se ha quedado fuera. Por ejemplo, la Agencia Nacional de Mejoras Policiales (NPIA) del Reino Unido tuvo cerca de 9.2 millones de registros criminales en 2009; dato, que no es grande si lo comparamos con la cantidad de información criminológicamente relevante que día tras día se va expulsando a través de las redes sociales (Williams et al., 2017). Lo mismo sucede con el resto de "*uves*", la velocidad con que los datos se van produciendo, su variedad (redes sociales, archivos criminales, censos de población, etc.), su veracidad y la imperiosa necesidad de extraer valor de toda esa información hace que el Big Data este cada día cobrando más protagonismo.

Habiendo dejado claro la importancia que cobra esta tecnología en la lucha contra el delito es preciso discernir entre los dos modos en los que el Big Data, a juicio de Chan y Bennet-Moses (2016), puede beneficiar a la Criminología por medio del *predictive policing*. En primer lugar, el Big Data se puede utilizar como una herramienta de investigación complementaria que ayude al conjunto de metodologías y teorías ya existentes en su intento de comprender y desengranar el fenómeno criminal (*Descriptive Analitics)*. Así, de este modo, esta nueva implementación tecnológica, entre sus muchas fuentes de información, es capaz de almacenar, procesar y analizar, por ejemplo, el contenido que se va publicando en redes sociales o la información relacionada con las telecomunicaciones. Esto supone una mejora tanto cualitativa como cuantitativa sobre los tradicionales métodos de extracción de información de la población (encuestas de victimización, historias de vida, etc.), ya que permite obtener datos en tiempo real y testimonios de colectivos normalmente no representados[18], lo que se traduce en un incremento de la veracidad de los resultados obtenidos (Williams et al., 2017). Prueba de la validez de este tipo de métodos de investigación es, por ejemplo, un estudio elaborado por Tumasjan et al. (2010) en el que se midió en Twitter el sentimiento que se expresaba hacía cada uno de los candidatos a las elecciones generales de Alemania. Se concluyó que los resultados obtenidos a través de esta fuente de información eran igual de precisos que los obtenidos a través de las encuestas y sondeos tradicionales.

En segundo lugar, siguiendo la doctrina de Chan y Bennet (2016), el Big Data, valiéndose de técnicas como el *data mining* y el *machine learning*, puede ser utilizado para desarrollar algoritmos capaces de analizar toda esa gran cantidad de datos y enfocarlos en elaborar técnicas de prevención criminal (*Predictive Analitics)*. Los algoritmos matemáticos son de gran ayuda para identificar patrones criminológicos relevantes entre las distintas variables, de tal modo que se consiga obtener "nuevo conocimiento para la elaboración de políticas de seguridad" (Van`t-Wout et al., 2019). Este tipo de métodos va dirigido tanto a crear mapas criminales basados en "puntos calientes" que ayuden a la policía a realizar una mejor prevención situacional como a

[18] Con las encuestas de victimización tradicionales siempre hay sectores poblacionales que tienden a quedar marginados o no representados, provocando así una desconexión entre la realidad y los resultados criminológicos. A través del análisis de las nuevas tecnologías, por ejemplo, a través de las redes sociales, es posible acceder a esos individuos que antes no expresaban su parecer en este tipo de metodologías. Véase Williams, M. L., Burnap, P., y Sloan, L. (2017). Crime sensing with big data: The affordances and limitations of using open-source communications to estimate crime patterns. *The British Journal of Criminology*, *57*(2), 320-340.

crear perfiles personales basados en la peligrosidad criminal y las probabilidades de reincidencia de los condenados. Como a lo largo de este texto se verá, en la actualidad, el Big Data muestra un mayor grado de eficacia en la prevención situacional, ambiental y despersonalizada, mostrando serias debilidades en el campo del perfilamiento criminal.

Estos dos espectros en los que esta tecnología muestra su utilidad no son contrapuestos, es más, son complementarios, así, se manifiestan como las dos piezas de un puzle que al combinarse pueden llevar al desarrollo de programas de prevención de la delincuencia muy eficaces e innovadores (*predictive policing)*. Además, se pretende conseguir la automatización del sistema de tal modo que, vaya adaptándose a las nuevas actualizaciones socio-criminológicas, es lo que se conoce como el ciclo del Big Data o ciclo de inteligencia (Clark, 2019). Una vez se ha obtenido, procesado y analizado la información, las conclusiones son evaluadas y se produce una transición hacia nuevos problemas a los que dar solución.

Por último, es preciso mencionar una idea que en la mayoría de los análisis sobre la materia tiende a pasar desapercibida. Esta tiene que ver con los factores que facilitan la inclusión de este tipo de tecnologías en las Fuerzas y Cuerpos de Seguridad del Estado. En esta línea, Villalobos-Fonseca (2020) destaca la cultura de cambio, la capacidad de innovación policial y las habilidades de resiliencia y de trabajo en equipo como elementos necesarios para una correcta cohesión del Big Data en las instituciones policiales. Del mismo modo, se ha averiguado que los cuerpos de policía que "*tienen un modelo o filosofía policial muy clara*", valorando esta herramienta como un medio para conseguir un fin, tienen un mayor grado de éxito a la hora de incorporar estas tecnologías en las labores de prevención (Villalobos-Fonseca, 2020).

4.2. ESTUDIOS EMPÍRICOS

A continuación, con el fin de demostrar el posible impacto que puede tener el Big Data como herramienta contra la delincuencia, se expondrán una serie de estudios científicos en los que se investiga la utilidad y eficacia de este tipo de tecnologías a la hora de predecir la incidencia criminal en un área concreta. De este modo, se han dividido los proyectos seleccionados en tres categorías según la materia en la que están enfocados: prevención situacional, prevención especial y redes sociales.

4.2.1. Prevención Situacional

A) Predictive Police Patroling (P3-DSS)[19]

Este estudio ha sido llevado a cabo por Miguel Camacho Collados, Inspector del Cuerpo Nacional de Policía (CNP) y consiste en la elaboración de un algoritmo (P3-DSS) formado por una Unidad de Preprocesamiento de Datos (DPPU), una Unidad de Predicción del Riesgo de Criminalidad (CFRU) y una Unidad de Optimización de los Sectores de las Patrullas Policiales (PSOU). El proyecto tiene el objetivo de implementar "*una política de patrullaje predictivo para aumentar la presencia de los agentes en las zonas donde más se necesitan, y así reducir la probabilidad de ocurrencia del delito"* (Collados, 2016, p.6).

Para la investigación se utilizó el Distrito Central de Madrid compuesto por seis barrios distintos, con una población de 150.000 habitantes. Se recopilaron los registros criminales referentes al delito de robo (105.755 incidentes) entre los años 2008 y 2012, ya que es uno de los delitos más cometidos en España y uno de los principales objetivos del Cuerpo Nacional de Policía. A su vez, se valieron de los Sistemas de Información Geográfica (SIG) del CNP que integran los sucesos delictivos sobre un mapa geográfico de la ciudad, además de la localización de las patrullas de policía. Para la distribución de las zonas se divide el área seleccionada en cuadrículas o celdas.

Con todos estos datos se pone en marcha el sistema P3-DSS. En primer lugar, cobra protagonismo la Unidad de Preprocesamiento de Datos encargada de averiguar la *matriz C* en función del espacio y el tiempo (C = total de delitos denunciados en el espacio-tiempo seleccionado). De tal modo que, cada celda obtiene un grado de incidencia delictual, acorde a si en la misma han ocurrido más o menos delitos de robo. A continuación, es turno de la Unidad de Unidad de Predicción del Riesgo de Criminalidad, la cual, al analizar la *matriz C*, elabora un modelo de predicción delictiva que estima el riesgo de incidencia criminal para cada cuadrícula mediante series de tiempo, lo que se denomina comúnmente como un "mapa de calor". Es importante

[19] Estudio disponible en el siguiente enlace: https://digibug.ugr.es/bitstream/handle/10481/44557/26134081.pdf;jsessionid=38DAF079A0A27506D243406E9CF0EF6C;jsessionid=38DAF079A0A27506D243406E9CF0EF6C?sequence=6

mencionar que se les da menos importancia a los datos antiguos (2008) que a los más actuales, los cuales van soportando cada vez más relevancia (2012).

Para comprobar que la calidad de los datos predictivos del sistema CFRU, se analizan todos los años desde 2008 a 2011, con el objetivo de comprobar los datos reales del último año (2012) con la estimación otorgada por el algoritmo. De este ejercicio se obtiene que el algoritmo solo aporta un coeficiente de error de 1,73 robos por turno policial en cada estimación. Esto demuestra que la predicción de la delincuencia mediante este método es extremadamente exacta. Además, se verifica que este algoritmo es más eficaz que el método anterior utilizado por el Cuerpo Nacional de Policía para las estimaciones de criminalidad.

Posteriormente, entra en actuación la Unidad de Optimización de los Sectores de las Patrullas Policiales, la cual en función de los niveles de riesgo criminal recomienda una distribución de las patrullas personalizada para cada celda del mapa. Este algoritmo busca ubicar de manera eficiente a los agentes policiales dentro de un área determinada, para así paliar de la mejor manera posible la criminalidad que se ha calculado para ese espacio-tiempo. Para este estudio, se considera que un sistema de organización de patrullas policiales es eficiente cuando las áreas divididas son compactas, la carga de trabajo entre los agentes es homogénea y las patrullas pueden apoyarse entre ellas de una zona a otra. Para ello, dentro de esta unidad, se utiliza el sistema de Problema Multi-Criterio de División de Distritos Policiales (MC-PDP) que calcula las variables de un zoneo de patrullas eficiente para estimar el mejor reparto de los agentes en función de la incidencia delictiva de cada cuadrícula. En este estudio se calculó que para los datos utilizados, este algoritmo obtiene entre un 10,4% y un 11,97% de media (según el sistema de patrullas utilizado) más de eficacia que los medios tradicionales usados para asignar el diseño de patrullas.

Por lo tanto, se corroboró que existe una mejora significativa en la eficacia cuando se aplica el modelo P3-DSS, por lo que se demostró que es una herramienta muy útil como mecanismo de prevención de la delincuencia ayudando a las Fuerzas y Cuerpos de Seguridad del Estado a mejorar sus medidas de prevención situacional mediante el uso eficaz de los diseños a la hora de patrullar. De este modo los autores demuestran que el algoritmo genera rápidamente configuraciones de patrulla, que son más eficientes que las actualmente empleados por el Cuerpo Nacional de Policía (Collados, 2016).

B) Predicción Criminal a través de Datos Móviles y Demográficos[20]

Este estudio fue llevado a cabo por Bogomolov et al. (2014) a raíz de haber participado en la competición del "*Datathon for Social Good*", organizado por *Telefónica Digital*, *The Open Data Institute* y el *MIT*, durante el *Campus Party Europe* de 2013.

Gracias a esto recibieron lo que ellos denominaron como "*Smartsteps"* (Pasos Inteligentes), que no es otra cosa que una serie de datos referentes al comportamiento humano, estructurados de forma anónima en franjas de una hora, por un tiempo de tres semanas; y, procesados gracias a la actividad de las redes de antena de telefonía móvil situadas en la ciudad de Londres. Esta información fue proporcionada en un mapa dividido en cuadrículas o celdas. De esta manera, se reportó información relativa a la cantidad de personas que se encontraban en cada celda en función del tiempo, además de sus características personales (si son residentes, trabajadores o visitantes, así como su sexo y edad). Aparte de esto, también recibieron los registros criminales de diciembre 2012 y enero 2013, con información geográfica de donde ocurrió el hecho, el departamento de policía involucrado y su tipología delictiva. Por último, se les proporcionó datos sociodemográficos de la población de Londres, repartidas a un nivel *LSOA* (poblaciones de entre 1000 y 1500 habitantes). Estos incluían información relativa al nivel de estudios, migración, raza, procedencia, patrimonio, trabajo, ingresos, precios de viviendas, espacios verdes, esperanza de vida y tasa de natalidad, entre otros datos demográficos relevantes.

Con todo esto, los investigadores se plantearon comprobar si, a través de los datos proporcionados, se puede crear un sistema de predicción de la delincuencia basado en una metodología de puntos calientes que estime la criminalidad que habrá en los próximos meses. Para ello, se analizaron y procesaron los datos, obteniendo un mapa de puntos calientes, que indicaba: celda de alta criminalidad cuando se daban más de cinco delitos y celda de baja criminalidad cuando se daban cinco o menos incidentes (se utilizó el número 5 porque corresponde con la mediana de los delitos cometidos). A continuación, se extrajeron las variables que presentaban mayor correlación con la delincuencia y se

[20] Estudio disponible en el siguiente enlace: https://arxiv.org/pdf/1409.2983.pdf

introdujeron los datos en el algoritmo seleccionado para que elaborara un modelo predictivo de las celdas que serían puntos calientes dentro de un mes.

Los resultados mostraron que el algoritmo matemático es capaz de predecir con una precisión de casi el 70% si una cuadrícula será una zona de alta criminalidad o no en el próximo mes. También se demostró que introducir los datos relativos a los "*Smartsteps"* produce una mejora en la eficacia de la predicción criminal de un 7%. Esto concuerda con el hecho de que el top-20 de variables que más se correlacionan con el delito están relacionadas con la tecnología "*Smartsteps"*, minimizando la importancia de los datos relativos a las características sociodemográficas. De este modo, se encontró que las variables que más influían eran: el porcentaje de gente residente que se encontraba en cada celda y el porcentaje de diversidad de estos (residentes, trabajadores, hombres, mujeres, jóvenes, adultos, etc.). A mayor número de residentes en un área y mayor porcentaje de diversidad, mayor es el riesgo de incidencia criminal.

Esto demuestra que a través del análisis de grandes cantidades de información se pueden elaborar algoritmos capaces de predecir aproximadamente la delincuencia de una zona concreta, incrementado así, las posibilidades de realizar medidas eficaces en materia de prevención criminal.

C) Predicción Criminal a través de Datos de Taxis y PDI´s[21]

Los datos demográficos y la información geográfica han sido las variables más comúnmente utilizadas para elaborar modelos predictivos. Sin embargo, el desarrollo de nuevas tecnologías capaces de almacenar, procesar y analizar masivas cantidades de datos (Big Data) ofrece la posibilidad de explorar nuevas perspectivas para entender el proceso criminal (Wang et al., 2016). En este estudio desarrollado por Wang et al., (2016) se proponen dos variables complementarias para predecir la incidencia criminal, usando como modelo la ciudad de Chicago en Estado Unidos.

Estos dos factores alternativos son, por un lado, el flujo de viajes realizados por los taxis, que funcionaría como un medidor de la cantidad de movimiento de personas entre las distintas áreas de la ciudad; y, por otro lado, los Punto de Interés (PDI´s), que aportan información sobre los servicios que ofrece una determinada zona. Los viajes en

[21] Estudio disponible en el siguiente enlace: https://dl.acm.org/doi/pdf/10.1145/2939672.2939736

taxi cobran importancia, bajo la hipótesis de que los territorios, no solo los que colindan sino los que tienen mucho flujo de personas, se influencian mutuamente en las tasas de delincuencia. Mientras que los PDI´s, al ser un indicativo de las características de un área concreto, pueden ser útiles para elaborar un perfil del barrio que sea de ayuda a la hora de predecir la delincuencia, comprobando así, que factores se correlacionan más con la misma. Estos fueron clasificados de acuerdo con las siguientes categorías: comida, vivienda, viajes, arte y entretenimiento, ocio y aire libre, educación, vida nocturna, acontecimientos profesionales, tiendas y eventos.

A estas variables, se le añade un registro criminal de la ciudad de Chicago que aporta información del día y el lugar en el que se perpetraron los ilícitos, así como el tipo delictivo al que pertenece, desde el año 2001 hasta el 2015, sumando un total de 5.856.414 incidentes. Estos, son incorporados a un Sistema de información Geográfica (SIG) y se obtiene un mapa de calor dividido por distritos. Además, se recopiló información sobre los datos demográficos de las distintas zonas a través de las siguientes características: número de población, densidad, tasas de pobreza, tasas de desigualdad, estabilidad residencial y diversidad étnica. Con todo esto, el objetivo de este estudio fue el de predecir la tasa de delincuencia de una zona determinada, en función de los datos proporcionados por las áreas colindantes y las frecuentemente comunicadas, así como a través de los PDI´s y los datos sociodemográficos.

Del procesamiento y análisis de todos los elementos se obtuvo que los factores demográficos más correlacionados con la delincuencia son las tasas de pobreza y desigualdad, mientras que la diversidad étnica se correlaciona positivamente con bajos niveles de criminalidad. Del mismo modo, se averiguó que los acontecimientos profesionales o de carácter laboral eran los que más correlación mantenían con los incidentes delictivos. Sorprendentemente, el factor vida nocturna no mostró apenas correlación con la criminalidad. En esta línea, el estudió también concluyó que existe una correlación positiva entre los territorios colindantes, de tal manera que si las áreas cercanas tienen una criminalidad alta, es más probable que la zona objetivo también la tenga. Por último, se obtuvo que existe una correlación positiva entre el flujo de viajes de taxi y la incidencia criminal de una zona. Así, se demostró que cuanto más tráfico de taxis se recibe de zonas conflictivas, mayor tasa de delincuencia se soporta.

En el mejor de los casos, analizar estas variables alternativas (flujo de taxis y POI´s) disminuye un 17,6% el porcentaje de error sobre la predicción del crimen. Esto, certifica que abrir el abanico de factores y variables sobre los que realizar un análisis conlleva un mayor porcentaje de éxito a la hora de estimar la tasa delictiva. Y, no hace falta decir, que los métodos tradicionales encuentran serias dificultades a la hora de realizar esta tarea, por lo que el Big Data se configura como una herramienta imprescindible en todo modelo que busque la prevención situacional por medio de la predicción delictiva.

4.2.2. Prevención especial

Habiendo evidenciado la importancia del Big Data para la elaboración de estrategias de prevención situacional, es imposible no preguntarse si esta efectividad predictiva tendría el mismo resultado a la hora de estimar las probabilidades de reincidencia de las personas ya condenadas. Para responder a esta pregunta es preciso abordar un estudio elaborado por Van´t-Wout et al. (2019)[22], en el cual se propuso estimar las posibilidades de reincidencia de individuos que ya cuentan con un historial de detenciones, mediante el análisis de Big Data. Debido a que este estudio solo pone el foco en las personas detenidas y no condenadas hay que tomar con cierta cautela los resultados. Aun así, la representación del texto servirá para esbozar un modelo de predicción de reincidencia del comportamiento humano, que permitirá hacerse una idea de si esta tecnología tiene cabida en el sector de la prevención especial. Para la explicación de este estudio se empleará el término reincidencia para la acción de volver a ser detenido pese a que no corresponde, pues esta solo ocurriría una vez se ha sido condenado.

Para la elaboración del proyecto se recopilaron los informes de detenciones realizadas en la Región Metropolitana De Santiago de Chile desde el año 2009 hasta enero de 2018, con un total de 777.724 registros correspondientes a 332.609 individuos. Además, se obtuvieron sus datos biográficos (sexo, edad, y familiares con antecedentes) por medio de la Policía de Investigaciones. A continuación, a cada sujeto se le separó la última detención del resto, de tal modo que el algoritmo en base al resto de detenciones

[22] Estudio disponible en el siguiente enlace: https://repositorio.uc.cl/xmlui/bitstream/handle/11534/28903/Propuestas%20para%20Chile%202018.pdf?sequence=1#page=48

tuvo que predecir la probabilidad de reincidencia. Así, se empleó la última de detención a modo de verificación de la efectividad del algoritmo. Como fundamento para el algoritmo, se emplearon variables relativas al historial delictivo (frecuencia, número de detenciones y tipos de delitos) y a los datos biográficos.

Los resultados muestran que el algoritmo es capaz de predecir cuando un sujeto no va a ser detenido con un 91% de acierto. Sin embargo, a la hora de predecir si un sujeto va a ser detenido, el porcentaje baja hasta un peligroso 63%. Esta, parece una cifra no tan débil, no obstante, no es suficiente para ser aplicada en un modelo de predicción de reincidencia. Debido a que esto es una predicción individualizada, al contrario que la prevención situacional, el error puede conllevar un elevado menoscabo social y personal para el individuo que la sufre. De este modo los autores concluyen que: "*en términos de predicción, los resultados del estudio indican una limitada capacidad predictiva a partir de los datos disponibles"* (Van´t-Wout et al., 2019, p.67).

Por lo tanto, no se puede responder de manera afirmativa a la pregunta planteada con anterioridad. Hasta la fecha, la tecnología del Big Data no se ha afianzado como un método de prevención especial eficaz. Así, se ubica su utilidad significativa en el campo de la prevención situacional donde sí se ha demostrado su valor predictivo.

4.2.3. Redes Sociales

Las nuevas tecnologías relacionadas con el Big Data, capaces de almacenar y analizar grandes cantidades de información, abren las puertas al uso estadístico y analítico de uno de los campos que más está creciendo en cuanto a volumen en el mundo de Internet: las redes sociales. Esto incluye, por supuesto, a la Criminología, donde desde hace unos años se viene estudiando la posibilidad de predecir las tasas de criminalidad a través del análisis de las redes sociales. En esta línea, es interesante mencionar el proyecto llevado a cabo por Bendler et al. (2014)[23] en el que se estudió la posible existencia de una correlación entre la frecuencia de mensajes publicados en Twitter en un área determinada y su tasa delictiva.

[23] Estudio disponible en el siguiente enlace: https://nnw.org/sites/default/files/INVESTIGATINGCRIME-TO-TWITTERRELATIONSHIPSINURBANENVIRONMENT.pdf

Este estudio se llevó a cabo en la ciudad de San Francisco en Estados Unidos y se recogieron datos de registros criminales desde agosto de 2013 hasta noviembre de 2014. Las mismas fechas se utilizaron para recopilar los *tweets* a través de Twitter API. Los datos fueron introducidos en Sistema de Información Geográfica (SIG) para ubicar en un mapa los delitos cometidos y la geolocalización de los tweets. A continuación se realizó una Regresión de Poisson entre los datos y se obtuvieron evidencias de la correlación entre el incremento de mensajes en Twitter y el descenso de la actividad criminal, en función del espacio de tiempo de una hora. Sobre todo se demostró que los mensajes en esta red social se correlacionaban con tipologías delictivas específicas como los robos. Esto puede deberse a que Twitter proporciona la localización desde donde se manda un mensaje, por lo que permite averiguar la población real que se encuentra en una zona determinada (*"población ambiente"*)[24]. Así, es comprensible que cuando más desolada se encuentre una zona mayor es la probabilidad de que ocurran robos (menos *tweets* implican menos personas en un área, lo que se relaciona con una mayor ausencia de guardianas capaces por lo que deriva en una menor prevención situacional).

Posteriormente, para comprobar la eficacia predictiva de estos datos, se dividieron los mismos en dos grupos, uno de entrenamiento (los primeros 28 días) para el algoritmo y otro de control para verificar los resultados. Además, para verificar el valor de analizar las redes sociales para predecir la delincuencia, se extrajo, por un lado, una estimación de la criminalidad utilizando las variables relacionadas con Twitter y, por otro lado, una sin emplearlas. Las conclusiones reflejaron que el éxito de la predicción criminal para el delito de robo ascendía al utilizar las redes sociales como indicadores, de un 76% a un 81%. De este modo se demostró que el uso de herramientas capaces de almacenar, procesar y analizar grandes cantidades de datos (Big Data), en este caso relativos a las redes sociales, tiene un gran valor y eficacia para explicar y predecir el fenómeno criminal.

En la misma línea, autores como Williams et al. (2017)[25] realizaron estudios similares y concluyeron que el uso de variables alternativas, como la frecuencia de mensajes escritos en Twitter, aumenta la capacidad de éxito de los modelos de estimación

[24] Véase también Malleson, N., y Andresen, M. A. (2015). The impact of using social media data in crime rate calculations: shifting hot spots and changing spatial patterns. *Cartography and Geographic Information Science*, *42*(2), 112-121.

[25] Estudio disponible en el siguiente enlace: https://orca.cf.ac.uk/87031/7/azw031.pdf

delictiva. No obstante, este proyecto busco dar un paso más allá y analizó, no solo la frecuencia con la que los *tweets* se mandaban sino el contenido de los mismos. Para ello, formularon una hipótesis basada en la teoría de las ventanas rotas, que alegaba que si el contenido de los *tweets* indicaba una cierta tendencia a la degradación del barrio, este, sería un indicativo de futuras tasas de delincuencia. Pese a que esta hipótesis fue refutada[26], se consiguió desarrollar una metodología capaz de analizar el contenido de 8,4 millones de *tweets*. Esto abre la puerta a que, diferentes teorías basadas en hipótesis alternativas se beneficien de este tipo de algoritmos capaces de interactuar con millones de fuentes de datos provenientes de las redes sociales y, extraigan de ellos conclusiones encaminadas a entender y predecir, con un mayor grado de acierto, la criminalidad de una zona concreta.

4.3. EJEMPLOS PRÁCTICOS

Aunque no de un modo hegemónico, estos estudios empíricos se han ido materializando a través de diferentes programas de análisis y predicción de la delincuencia. Esta tendencia está a la alza, así, solo en Europa desde el año 2000 se han multiplicado por diez el número de programas utilizados por los distintos países, entre los que destacan Alemania, Reino Unido y Francia (Cinelli y Gan, 2019). El continente americano tampoco se queda atrás, países como Estados Unidos, México (Vilalta-Perdomo, 2017) y Chile (Van´t-Wout et al., 2019) han sido punteros en el desarrollo de metodologías tecnológicas de prevención de la delincuencia. Del mismo modo, al revisar la bibliografía existente se puede apreciar como China está dedicando muchos medios a estudios de tecnologías de Big Data con fines de control criminal[27]. Por razones evidentes, no se procederá a explicar todos los ejemplos existentes, sino que se expondrá el programa que más impacto y relevancia ha tenido durante estos años, además, del que por alusión y pertenencia geográfica nos compete.

[26] Esta hipótesis no pudo ser probada debido a que influyen varios factores a la hora de utilizar la teoría de las ventanas rotas como un medidor fiable. Dependiendo de las características socioeconómicas del barrio se tiene una sensibilización distinta a la degradación de las zonas, por lo que no esta teoría flaquea a la hora de ser empleada como indicador de la criminalidad entre distintas áreas (Williams et al., 2017).

[27] Por ejemplo: Feng, M., Zheng, J., Ren, J., Hussain, A., Li, X., Xi, Y., y Liu, Q. (2019). Big data analytics and mining for effective visualization and trends forecasting of crime data. *IEEE Access*, *7*, 106111-106123.

A) Predpol / Geolitica[28]

Predpol, ahora denominado Geolitica, es un proyecto que se desarrolló a raíz de una investigación en común entre el Departamento de Policía de Los Ángeles, la Universidad de California (UCLA) y la Universidad de Santa Clara, este buscaba utilizar los datos criminales para algo más que fines descriptivos. De esta manera, se elaboró un algoritmo matemático capaz de predecir las zonas, en relación con el tiempo, en donde existe un mayor riesgo de que se cometa un incidente delictivo. Todo ello, a través de los datos exclusivamente de registros criminales referentes a tipología delictiva y a fecha y lugar de los incidentes; es importante mencionar que no utiliza ningún tipo de información demográfica, social o personal. El algoritmo se sirve de un Sistema de Información Geográfica (SIG) que divide el mapa del área en cuadrículas de 150 metros cuadrados aproximadamente. Para cada celda, asigna un nivel de riesgo criminal en función del análisis de los datos mencionados anteriormente. Esto informa a los agentes de policía, no solo de las zonas que deben patrullar sino del tiempo que deben permanecer en cada una de ellas para procurar una mayor prevención situacional (Predpol, 2021).

Actualmente, Geolitica ha sido implementado en más de cincuenta departamentos de policía estadounidenses y otras tantos en Reino Unido (Smith, 2018). En Kent, Inglaterra, se realizó una evaluación interna del programa en el año 2014, en ella se concluyó que Geolitica, en aquel momento Predpol, doblaba las posibilidades de predecir la localización de un ilícito criminal frente a los métodos de inteligencia tradicionales (Cinelli y Gan, 2019). En el mismo año, en la ciudad de Atlanta, Estado Unidos, se registró que, en el periodo de 90 días que había estado el algoritmo en funcionamiento, se disminuyó la delincuencia entre 8% y un 9%, especialmente en los delitos de robos de coches y en viviendas. Por otro lado, las tasas de criminalidad aumentaron un 8% en las zonas donde no se aplicó el sistema predictivo (Turner et al., 2014).

Sin embargo, la totalidad de distritos en Reino Unido, y desde hace relativamente poco el Departamento de Policía de Los Ángeles[29], han decidido dejar de utilizar el programa debido al elevado desembolso económico que conlleva, pues el software es proporcionado por una empresa privada. En su defecto, los cuerpos de policía están optando por desarrollar sus propios programas internos de análisis y predicción de la

[28] Enlace a página web: https://www.predpol.com/ y https://geolitica.com/

[29] Noticia disponible en el siguiente enlace: https://www.buzzfeednews.com/article/carolinehaskins1/los-angeles-police-department-dumping-predpol-predictive

delincuencia a través de la tecnología Big Data (Cinelli y Gan, 2019). Un claro ejemplo de ello es el sistema P3-DSS[30], mencionado con anterioridad, que está en vías de desarrollo en España a través del Cuerpo Nacional de Policía.

B) EuroCop Pred-Crime[31]

En España aún no se ha llevado a la *praxis* ningún programa o estrategia de predicción de la delincuencia por medio de tecnologías relacionadas con el Big Data. No obstante, sí que existe un sistema que aspira, junto al algoritmo del CNP P3-DSS, a colaborar con las policías de todo el país en la lucha por la prevención situacional. Así, hay que mencionar a la aplicación Pred-Crime, que fue desarrollada en 2014 por la empresa española *EuroCop Security Systems* en colaboración con la Universidad Jaume I de Castellón. Esta consiste en un sistema de "*tratamiento de datos masivos vinculados a delitos […], basado en un modelo espacio-temporal e información geográfica de mapas de calor; que utiliza modelos y algoritmos matemáticos y que permitirá la predicción y prevención de los delitos"* (EuroCop, 2015).

De este modo, el software recoge información de registros criminales, encuestas y procesos de participación ciudadana, además de otros datos descriptivos espacio-temporales, y los emplea en elaborar, a través de su algoritmo matemático, mapas de calor de zonas con mayor riesgo delictivo (EuroCop, 2015). También, como novedad con respecto al programa anterior, Pred-Crime permite la incorporación del análisis de redes sociales, factor que como se ha visto anteriormente aumenta el valor de los informes preventivos. Actualmente solo existen evidencias de que este modelo haya sido puesto a prueba en las Policías Locales de Rivas Vaciamadrid y de Castellón (Cinelli, 2019).

[30] Véase Collados, M. C. (2016). *Statistical analysis of spatio-temporal crime patterns: Optimization of patrolling strategies* (Doctoral dissertation, Universidad de Granada).
[31] Enlace a página web: https://www.eurocop.com/sistemas-de-eurocop/analisis-y-prediccion-del-delito/

5. ENCAJE ÉTICO Y JURÍDICO DEL BIG DATA PARA PREVENEIR LA DELICNUENCIA

El Big Data, esas ingentes cantidades de datos destinados a ser almacenados, procesados, organizados y analizados, con el objetivo ulterior de obtener patrones del comportamiento humano sobre el modo de actuar de las masas y más concretamente de los individuos. Su importancia cobra sentido en la medida en que, a través de esta herramienta se pueden tanto describir y explicar conductas humanas, como anticiparse a ellas (predecir). Sin embargo, hay una razón más por lo que esta tecnología es útil para el fin propuesto. Todos los datos son generados directa o indirectamente por cada individuo. Es decir, esta información digital, originada por el ser humano, informa de las cualidades, características o preferencias (localización, ideología, gustos, rasgos físicos, patrimonio, elementos demográficos, etc.) propias de cada sujeto y pertenecen, en definitiva, a su forma de ser en contacto con los medios sociales disponibles. Además, se obtiene información de los administradores de servicios públicos (centros de salud, colegios, etc.) y datos identificativos proporcionados por los consumidores al contratar ciertos productos (domicilio, correo electrónico, teléfono, etc.); todos ellos suelen estar disponible por voluntad de los usuarios (Valls-Prieto, 2018). Y es ahí, donde reside su valor como innovación tecnológica, ya que consigue acceder a la esfera personal del individuo, a su "historial de actividad", donde desarrolla su vida social, para extraer información y así conseguir elaborar programas de descripción y predicción de su conducta.

A lo largo de este texto, en los estudios y ejemplos planteados, se ha visto como los algoritmos matemáticos se valían de información respectiva a registros criminales, redes sociales, características demográficas (población, raza, renta, empleabilidad, educación, etc.), Puntos de Interés disponibles, situación geográfica y flujos de movimiento. Todas estas variables, que han servido para elaborar modelos predictivos de la criminalidad, han sido fruto de los datos generados por miles de individuos objeto de estudio. No obstante, algo que no conoce todo el mundo es que, la ciudadanía es dueña y propietaria indiscutible de sus propios datos personales y que, en consecuencia, existen una serie de normas y regulaciones encaminadas a proteger que se cumpla esta relación de titularidad.

A continuación, se expondrá cual es la normativa vigente en materia de protección de datos que, tanto las empresas privadas como las instituciones públicas deben cumplir para poder utilizar el Big Data como herramienta de prevención de la delincuencia. Seguido, se analizará el debate ético y moral existente en la ponderación de los derechos implicados, con la clásica disputa análoga entre libertad y seguridad.

5.1. REGULACIÓN LEGISLATIVA

Comenzando por un planteamiento exterior, a nivel europeo, la *Carta de los Derechos Fundamentales de la Unión Europea* (CDFUE) (2012) dicta en su artículo 8 el derecho de "*toda persona a la protección de los datos de carácter personal que la conciernan*". Del mismo modo, se añade en el punto segundo que los mismos "*se trataran de modo leal, para fines concretos y sobre la base del consentimiento de la persona afectada o en virtud de otro fundamento legítimo previsto por la ley".* Además, se asegura de dejar claro que "*toda persona tiene derecho a acceder a los datos recogidos que le conciernan y a su rectificación"*, así como establece que estas normas deben controlarse por una "*autoridad independiente"*. Bajo la potestad que otorga este precepto, se ha formulado posteriormente un abundante cuerpo legislativo y jurisprudencial dirigido a desarrollar más en profundidad el contenido de este Derecho (Sancho-López, 2018).

La normativa creada para cumplir este objetivo fue el Reglamento (UE) 679/2016 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos del Diario Oficial de la Unión Europea (en adelante Reglamento (UE) 679/2016). Sin embargo, y pese a que esta regulación trata de dotar de significado al concepto de "*Habeas Data"*[32], de tal manera que los ciudadanos cuenten con mayores oportunidades de control sobre sus datos personales (Sancho-López, 2018); no tiene competencia, según indica su artículo 2 d), en materia de tratamiento de datos "*por parte de las autoridades competentes con fines de prevención, investigación, detección o enjuiciamiento de infracciones penales, o de ejecución de sanciones penales, incluida la de protección frente a amenazas a la seguridad pública y su prevención"* (Reglamento

[32] El "*habeas data"* se configura cómo el derecho de todo ciudadano a la intimidad y libertad informática, otorgando al usuario el poder de la autodeterminación de sus propios datos, de tal manera que tenga la facultad de acceder (art. 15), conocer (art. 15), rectificar (art. 13), cancelar (art. 19), suprimir (art. 17), limitar (art. 18) y controlar (art. 20) el uso de los mismos (Diccionario panhispánico del español jurídico [DPEJ], Reglamento (UE) 679/2016).

(UE) 679/2016); delegando dicha regulación a la Directiva 2016/680 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos (en adelante Directiva 2016/680).

Primero de todo, hay que aclarar quienes son los sujetos objeto de esta regulación, así la Directiva 2016/680 en su artículo 3 establece que se entiende por "*autoridades competentes"*, tanto las de carácter público (FCSE, Juzgados, etc.) como cualquier otro organismo que cuente con la confianza y el respaldo de los anteriores, para el tratamiento de datos personales con los fines a los que alude la citada Directiva (Geolitica, Pred-crime, etc.). En segundo lugar, es preciso mencionar que tipos de datos son susceptibles de ser tutelados por este ordenamiento, es decir, sobre qué clase de información recaerá el "escudo" del derecho a la protección de datos. En relación con esto, la presente Directiva en la consideración 21 rige que la información objeto de amparo será toda "*la relativa a una persona física identificada o identificable*"[33], excluyendo de la ecuación a los datos anónimos, ya sea por origen o porque han sido resultado de un proceso de anonimización.

Autores como Valls-Prieto (2018) declaran que esta protección jurídica abarca un espectro amplísimo, sobre todo en tecnologías como el Big Data, donde la correlación de distintas variables y fuentes de información puede llevar a la identificación de una persona. Coincido con dicho escritor en que, a criterio personal, los esfuerzos deben centrarse en regular que se consiga una efectiva anonimización de los datos que cumpla con todas las garantías, en lugar de promover una visión abolicionista que deje "cojo" al derecho de seguridad pública. Sin embargo, discrepo en que este proceso sea tan difícil de lograr, a modo de ejemplo, atendiendo a los estudios mencionados, datos como las estadísticas criminales, los Punto de Interés, la situación geográfica, los flujos de movimiento, algunas características sociodemográficas (población, empleabilidad,

[33] El artículo 3 de la Directiva 2016/680 señala que se considera "*persona física identificable a toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un identificador, como por ejemplo un nombre, un número de identificación, unos datos de localización, un identificador en línea o uno o varios elementos propios de la identidad física, fisiológica, genética, psíquica, económica, cultural o social de dicha persona*".

educación, etc.), o incluso datos relativos a las redes sociales (la población ambiente a través de la frecuencia de los mensajes) sí que son, y han sido, posibles de anonimizar.

Continuado con el análisis de la Directiva 2016/680, la misma, en su artículo 10, hace una distinción relevante mediante la cual otorga un mayor nivel de protección al tratamiento de categorías especiales de datos, entre los que se incluyen los relativos al *origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, así como el tratamiento de datos genéticos, datos biométricos […], datos relativos a la salud o a la vida sexual.* Estos, solo podrán ser utilizados cuando sea estrictamente necesario, lo autorice el Derecho de la Unión o el de un Estado miembro o hayan sido publicados por el titular de los mismos.

En cuanto al procesamiento de los datos por parte de las autoridades competentes (artículo 4), y en sintonía con lo que promulga la CDFUE, la Directiva 2016/860 indica que estos deben ser "*tratados de manera lícita y leal"*, recogidos con un fin determinado y legítimo, adecuados (principio de necesidad) y relevantes en relación con el fin propuesto (principio de proporcionalidad). Incluso, se deben marcar plazos para la eliminación de la información una vez haya pasado un tiempo determinado o fijar un control que estipule si siguen cumpliendo los principios que legitimaron el tratamiento en origen (artículo 5).

Los modelos predictivos que hemos abordado en este texto emplean fórmulas matemáticas y sistemas automatizados (*data mining* y *machine learning*) para extraer patrones y conclusiones a partir de los datos recogidos. A este respecto, la Directiva 2016/680 se manifiesta, en su artículo 11, indicando que estas decisiones basadas únicamente en algoritmos automatizados estarán prohibidas siempre y cuando, al utilizar los datos de carácter personal, se produzcan "*efectos jurídicos negativos para el interesado o le afecten significativamente*", con la excepción de que esta medida esté autorizada por el Derecho de la Unión o de un país miembro y que se garanticen todos los derechos y libertades del sujeto. Esto deja en una situación comprometida a modelos algorítmicos predictivos basados en la elaboración de perfiles con fines de prevención especial (por ejemplo, el estudio elaborado por Van´t-Wout et al., 2019). Así, no quedan incluidas bajo esta regulación los que tratan datos de carácter anónimo y se centran en medir las estimaciones del delito de una manera situacional (por ejemplo, P3-DSS).

Por último, es relevante hacer mención al artículo 41 de la citada regulación europea, el cual regula la creación por parte de los Estados miembros de una "*autoridad de control"* encargada de "*proteger los derechos y las libertades fundamentales de las personas físicas en lo que respecta al tratamiento de sus datos personales*". De esta manera, entre sus funciones están, entre otras (artículo 46): supervisar y hacer cumplir los preceptos de la mencionada Directiva; promover la sensibilización acerca de los riesgos, normas y garantías del tratamiento de información; asesorar en materia de protección de datos; y, investigar y tratar las reclamaciones planteadas a raíz del incumplimiento de esta legislación (Directiva 2016/680).

Estas normativas europeas, tanto el CDFUE y el RGDP como la Directiva 2016/680, inspiraron la creación de una propia normativa española con una estructura y contenido muy similar. En palabras de Durán (2019) la Ley Orgánica de Protección de Datos y garantías de derechos digitales (LOPD) es una adaptación de la regulación europea y, en consecuencia, no puede contradecir lo dictaminado desde el antiguo continente, por lo que simplemente se limita a especificar algunos elementos técnicos. Por ello, dadas las similitudes, no se considera necesario iniciar un nuevo análisis de este texto legislativo. Solo cabe mencionar, que España, siguiendo directrices europeas, en esta ley regula la creación de una autoridad competente en materia de protección de datos denominada Agencia Española de Protección de Datos (AEPD) encargada de supervisar que se cumplen las directrices europeas (art 44 y ss. LOPD, de 5 de diciembre).

5.2. PONDERACIÓN DE DERECHOS IMPLICADOS Y CONSIDERACIONES ÉTICAS

Antes de entrar en valoraciones éticas y ponderaciones de distintos bienes jurídicos, es necesario realizar una aclaración acerca de qué derechos están en juego en este contexto. Es habitual relacionar el uso del Big Data o el incumplimiento de las leyes de protección de datos, con la vulneración del artículo 18.1 de la Constitución Española (CE), el cual tutela entre sus líneas el Derecho Fundamental a la intimidad y a la privacidad. Frente a esta premisa, autoras como Sancho-López (2018) exponen un planteamiento distinto dirigido hacía el empoderamiento del derecho a la protección de datos personales, entendido como un Derecho Fundamental en sí mismo. Así, como bien lo explica la autora, fue el propio Tribunal Constitucional el encargado de otorgarle ese

valor en función del artículo 18.4 de la CE, recalcando además la importancia de los conceptos de libertad informática y "habeas data".

En el otro lado de la balanza, se encuentra el derecho de todo ciudadano a que el Estado le permita desarrollar su vida en unas condiciones protagonizadas por la seguridad pública (art. 149.1.29ª CE). Del mismo modo, toda persona tiene derecho a que su seguridad ciudadana sea promovida a través de elementos institucionales como las Fuerzas y Cuerpos de Seguridad del Estado (art.140.1 CE).

Dicho esto, es aquí donde se produce el eterno debate o conflicto entre seguridad y libertad (protección de datos). ¿Es ético o moral que se utilicen datos personales para elaborar perfiles de afinidad delictiva? O al contrario, ¿es ético o moral que las FCSE, no eviten la comisión de futuros delitos por no poder acceder a los datos de la población? Encontraremos diferentes respuestas en función del sujeto a quién se interrogue, dependiendo, incluso, del país en el que se pregunte. Esto nos hace recordar el importante papel que juega la cultura en la moral de los individuos. Así, por ejemplo, si preguntamos hace 30 años a las personas de la época, si están de acuerdo con que se pase su equipaje por un escáner de Rayos-X y se les someta a cacheos sin ninguna justificación previa en los aeropuertos, su respuesta sería claramente negativa. Sin embargo, hoy en día, si eliminámaramos esta normativa y se suprimiera esta medida que atenta contra el derecho a intimidad de todos los pasajeros, muchos de ellos viajarían con miedo e inseguridad.

Por la tanto, es importante mantenerse racional y objetivo en este asunto, y no dejar que tu juicio se incline hacia ninguno de los dos extremos de manera desmesurada. La solución a esta dialéctica, desde mi punto de vista, pasa por encontrar el equilibro entre los dos derechos, hallar el punto medio donde se mantenga el valor de ambos, sin provocar un detrimento excesivo en el opuesto. Esta no es una solución novedosa ni revolucionaria. No obstante, y pese a que en la actualidad los Derechos individuales están muy consagrados, la tendencia histórica, siempre ha ido encaminada a favorecer el espectro de la seguridad, antes que el de la libertad.

Por lo tanto, y para que se entienda de un modo didáctico, podemos visualizar esta problemática como un edificio, donde el suelo estuviese formado por las libertades individuales, y el techo por el derecho a la seguridad. Como es lógico, el techo siempre tiende a caer, siendo necesaria una serie de vigas o pilares que mantengan la estructura. Como pilar principal se encuentra el Derecho, que busca el control basándose en la

legitimidad. Son múltiples los ordenamientos jurídicos, sobre todo en España, plagados de garantías inviolables (principios de proporcionalidad, necesidad, legalidad, igualdad, etc.), que permiten que no se haga un uso desmesurado y descontrolado de las actuaciones en materia de seguridad. En este aspecto, el pilar estaría formado por las regulaciones legislativas que hemos visto en el punto anterior. Sin embargo, esta columna por sí sola no basta, prueba de ello ha sido la existencia de regímenes dictatoriales donde imperaba el derecho penal de autor, amparado bajo el marco normativo.

De esta manera, como soporte complementario, se encuentra la Filosofía, encargada de buscar el control por medio de la ética y la moral. Como se ha mencionado previamente, la moral varía según el lugar y el periodo en el que nos encontremos, por lo que es preciso analizarla de acuerdo con las características concretas del espacio-tiempo en el que vivimos. Actualmente, la Unión Europea se encargó de recoger, a través de la Resolución del Parlamento Europeo, de 16 de febrero de 2017, relativa a las recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica, "*una buena expresión de lo que constituyen los principios éticos esenciales"* (Hueso, 2019, p.37). En consecuencia, se clarifican una serie de indicaciones que deben respetar todas las normativas que regulen temas relacionados con la robótica, la Inteligencia Artificial, y de manera análoga con el Big Data. Estos principios, dictaminados en los puntos 12 y 13, son los siguientes: "*beneficencia, no maleficencia, autonomía, justicia y transparencia, así como […] los principios consagrados en la Carta de los Derechos Fundamentales*" (Parlamento Europeo, 2017b, p.8).

Para el estudio de estos principios se seguirá la doctrina señalada por Lorenzo Cotino Hueso (2019), Catedrático de Derecho Constitucional de la Universidad de Valencia. En primer lugar, el principio de beneficencia implica que toda medida en la que se utilice el Big Data debe ir orientada a "hacer el bien", a buscar siempre el beneficio de la humanidad, y a obtener una sociedad más justa, segura y sostenible (Hueso, 2019). No se debe perder el foco de la problemática que se intenta solventar con el Big Data, que no es otra que conseguir una mayor prevención criminológica. De este modo, se debe evitar toda tentación que trate de desviar la atención sobre los fines benévolos originalmente planteados.

En segundo lugar, el principio de no maleficencia implica que, mediante el uso de esta tecnología no se provoquen "*daños físicos, psicológicos, financieros o sociales"*

(Hueso, 2019, p.37ed). Por lo que, de manera indirecta, los modelos matemáticos que den resultados individualizados e identificables deben de limitarse al máximo debido al alto coste personal y social que pueden suponer. En tercer lugar, el principio de autonomía conlleva la imperiosa necesidad de mantener un componente de dependencia en las decisiones tomadas por los algoritmos (Hueso ,2019). Es decir, no se debe otorgar todo el poder a estas nuevas tecnologías, el ser humano siempre debe tener "la última palabra". Esto supone aceptar la necesidad de que, a pesar de que los algoritmos predictivos deducen el índice de criminalidad de una zona, haya siempre una persona encargada de decidir si esa decisión es proporcionada, fiable y se acoge a todas las garantías.

En cuarto lugar, la justicia, en este caso no relacionada con el Derecho sino con la idea de igualdad, hace referencia a evitar todos los elementos que puedan provocar una situación de discriminación o estigmatización social (Hueso, 2019). Con esto en mente, el legislador europeo ofreció una protección distintiva a los datos de carácter especial (definidos con anterioridad), y que pueden generar situaciones discriminatorias y desfavorables entre las personas. Por último, enmarcado como un principio que engloba al resto, se encuentra el principio de transparencia que desarrolla la posibilidad de entender y comprender como funcionan los algoritmos para poder regular con decisión que se cumplan el resto de principios. Así, se pretende evitar la opacidad en el funcionamiento de los mismos de tal forma, que se habilite la posibilidad de "rendir cuentas" en caso de ser necesario (autoridad de control) (Hueso, 2019).

Hay que mencionar que esto, únicamente se trata de una representación de los principios éticos reseñados por la Unión Europea, sin que indique ningún tipo de exclusividad y con plena conciencia de que estos principios, pueden no agotarse en la mencionada recomendación europea.

A modo de síntesis, siempre y cuando se cumplan la totalidad de principios jurídicos y éticos, se podrá concluir que las herramientas de Big Data empeladas para la prevención de la delincuencia son "correctas" y "legítimas" para el desempeño de tal función. Por lo que cabe recordar, la importancia de establecer medios de control eficaces, capacitados para exigir el cumplimiento de todo este cuerpo ético-jurídico, de tal manera que, en busca de un objetivo benévolo y lícito, no se consiga un resultado inmoral y dañino.

6. CONCLUSIÓN

El objetivo principal de este trabajo era analizar el papel del Big Data como una herramienta de lucha contra el crimen. Con esto en mente, y a raíz de la revisión bibliográfica efectuada, se han podido apreciar las enormes posibilidades con las que cuentan las nuevas tecnologías del Big Data, a la hora de ayudar a prevenir la delincuencia, especialmente en su modalidad situacional, frente a otras de carácter psicosocial. Antes, solo podíamos contar con el juicio de experiencia proveniente de los altos mandos policiales, encargados de dirigir las patrullas. Sin embargo, a lo largo del texto se ha visto, como el análisis de Big Data ha demostrado un mayor grado de acierto prediciendo las zonas más proclives a albergar un incidente delictivo. Por lo que es innegable el potencial que tiene esta tecnología, como herramienta capaz de abrir una nueva puerta hacía una prevención situacional más eficaz y exitosa.

No obstante, este beneficio que se puede extraer del Big Data no viene sin coste alguno. Estos resultados se obtienen gracias al análisis de millones de datos generados por la ciudadanía, lo que puede suponer una cierta injerencia en su vida privada. Actualmente, Europa cuenta con una normativa extensa que regula la actividad de las FCSE a la hora de recopilar y analizar datos de carácter personal. Aunque es cierto que en este contexto legislativo el uso de datos anónimos (como los empleados en estrategias situacionales) es legítimo, es importante no relajarse y exigir que no se levante el pie del acelerador, promulgando que los textos legislativos mantengan una actitud renovadora que asegure que, con el paso del tiempo, no se perjudique ningún Derecho Fundamental y que en todo momento los valores éticos y morales sean protagonistas en la prevención criminal.

Teniendo en cuenta el potencial de esta tecnología, no resulta muy verosímil que actualmente en España solo se cuente con un proyecto ya desarrollado, aunque aún no operativo (Pred-Crime), y otro en proceso de prueba (P3-DSS). Del mismo modo, se hace palpable la diferencia que existe en cuanto al interés y relevancia que muestra nuestro país en el uso del Big Data si nos comparamos con países del continente americano. Se debe seguir trabajando y aportando medios para conseguir la implementación de todas estas herramientas tecnológicas en el día a día de las FCSE, solo así se podrá avanzar en el camino hacía una sociedad más segura y justa.

7. BIBLIOGRAFÍA

Angulo, M., y López, C. (2001). Teoría de la pena, Constitución y Código penal. *Derecho Penal Criminología, 22(71),* 55-68.

Arcila-Calderón, C., Barbosa-Caro, E., y Cabezuelo-Lorenzo, F. (2016). Técnicas big data: análisis de textos a gran escala para la investigación científica y periodística. *El profesional de la información*, *25*(4), 623-631.

Ariza, J. J. M. (1998). El control social del delito a través de la prevención situacional. *Revista de derecho penal y criminología*, 2, 281-326.

Barranco-Fragoso, R. (2012). ¿Qué es big data? *IBM Developer*. Obtenido de https://developer.ibm.com/es/technologies/data-science/articles/que-es-big-data/

Bendler, J., Brandt, T., Wagner, S., y Neumann, D. (2014). Investigating crime-to-twitter relationships in urban environments-facilitating a virtual neighborhood watch. *Proceedings of the European Conference on Information Systems (ECIS) 2014*, Tel Aviv, Israel.

Bernal-del Castillo, J. (2013). Prevención y seguridad ciudadana. La recepción en España de las teorías criminológicas de la prevención situacional. *Revista de Derecho Penal Y Criminología, 09,* 267-304.

Blum, A. (2003). "Machine learning theory". *FOCS 2003 Procs of the 44th Annual IEEE Symposium on foundations of computer science*. Washington DC: IEEE Computer Society, (pp. 2-4)

Bogomolov, A., Lepri, B., Staiano, J., Oliver, N., Pianesi, F., y Pentland, A. (2014). Once upon a crime: towards crime prediction from demographics and mobile data. *Proceedings of the 16th international conference on multimodal interaction* (pp. 427-434).

Cárdenas-Ruiz, M. (2004). Las teorías de la pena y su aplicación en el Código Penal. *Derecho Y Cambio Social*, 1(02). Recuperado de http://www.derechoycambiosocial.com/revista002/pena.htm

Castellanos, J. F. (2019). *Criminología y Big data: El futuro de la investigación criminológica (Vol.37).* México DF, México: Instituto Nacional de Ciencias Penales.

Chan, J., y Bennett-Moses, L. (2016). Is big data challenging criminology? *Theoretical criminology*, *20*(1), 21-39.

Cinelli, V. (2019). Prevención del crimen y predicción de delitos: ¿en qué punto está España? [Blog]. Recuperado de: https://blog.realinstitutoelcano.org/prevencion-del-crimen-y-prediccion-de-delitos-en-que-punto-esta-espana/

Cinelli, V., y Gan, A. M. (2019). El uso de programas de análisis predictivo en la inteligencia policial: una comparativa europea. *Revista de Estudios en Seguridad Internacional*, *5*(2), 1-19.

Clark, R. M. (2019). *Intelligence analysis: a target-centric approach*. Washington, Estados Unidos: CQ press.

Collados, M. C. (2016). *Statistical analysis of spatio-temporal crime patterns: Optimization of patrolling strategies* [Doctoral dissertation, Universidad de Granada].

Colmenarejo, R., (2018). Ética aplicada a la gestión de datos masivos. *Anales de La Cátedra Francisco Suárez,* 52, 113-129.

Conde, F. M., y Arán, M. G. (2010). *Derecho penal. Parte General,* 8ª edición. Valencia, España: Tirant lo Blanch.

Constitución Española. Boletín Oficial del Estado, 29 de diciembre de 1978, núm. 311.

Diccionario panhispánico del español jurídico [DPEJ]. (2020). Definición. Dpej.rae.es. Recuperado de: https://dpej.rae.es/lema/prevenci%C3%B3n-general

Directiva 2016/680 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos. Diario Oficial de la Unión Europea L119/89, de 4 de mayo de 2016.

Dubé, R. (2012). La théorie de la dissuasion remise en question par la rationalité du risque. *Canadian Journal of Law and Society*, 27, 1-29.

Durán-Domínguez, A. (2019). *Implicaciones éticas y legales en el uso del big data*. (Trabajo fin de grado). Universidad de Sevilla, España

Durán-Migliardi, M. (2016). La prevención general positiva como límite constitucional de la pena: Concepto, ámbitos de aplicación y discusión sobre su función. *Revista de derecho (Valdivia)*, *29*(1), 275-295.

Durán-Migliardi, M. (2011). Teorías absolutas de la pena: origen y fundamentos: conceptos y críticas fundamentales a la teoría de la retribución moral de Immanuel Kant a propósito del neo-retribucionismo y del neo-proporcionalismo en el derecho penal actual. *Revista de filosofía*, *67*, 123-144.

ECOSOC (2002). Medidas para promover la prevención eficaz del delito. Resolución 2002/13 del Consejo Económico y Social.

EuroCop. (2015). *Pred-Crime.- Sistema para la Predicción y Prevención del Delito.* Disponible en: https://www.eurocop.com/sistemas-de-eurocop/analisis-y-prediccion-del-delito/

Feng, M., Zheng, J., Ren, J., Hussain, A., Li, X., Xi, Y., y Liu, Q. (2019). Big data analytics and mining for effective visualization and trends forecasting of crime data. *IEEE Access*, 7, 106111-106123.

Fernández-Vega, E. (2017). El control y la prevención del delito como objeto de la criminología. Miscelánea Comillas. Revista de Ciencias Humanas y Sociales, 75(146), 171-194.

Fernández-Zalazar, D. C., Guralnik, G. E. (2017). El fenómeno de data mining. Efectos psicosociales y en la subjetividad. *IX Congreso Internacional de Investigación y Práctica Profesional en Psicología XXIV Jornadas de Investigación XIII Encuentro de Investigadores en Psicología del MERCOSUR*. Facultad de Psicología-Universidad de Buenos Aires.

Firican, G. (2017). The 10 Vs of Big Data. UPSIDE where DATA means BUSINESS. Recuperado de https://tdwi.org/articles/2017/02/08/10-vs-of-big-data.aspx

Forecast, G. M. D. T. (2019). Cisco visual networking index: global mobile data traffic forecast update, 2017–2022. *Update*, *2017*, 2022.

Galimany, A., (2014). *La creación de valor en las empresas a través del Big Data* (trabajo fin de grado). Universidad de Barcelona, España.

Garrido-Genovés, V. (2010). La prevención de la delincuencia en Europa y en España: Los retos pendientes. *Revista de Derecho Penal Y Criminología*, *03*, 377-408.

Grupo Fractalia (2016). Big Data:¿Dónde se almacena tanta información? [Blog]. Recuperado de https://fractaliasystems.com/big-data-donde-se-almacena/

Hoferek, S. R. (2019). *El derecho a la intimidad, la protección de datos personales y el big data a la luz del ordenamiento jurídico argentino* (trabajo fin de grado). Universidad Siglo XXI, Argentina.

Hueso, L. C. (2019). Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el Derecho. *Revista catalana de dret públic*, (58), 29-48.

Iancu, A. (2016). *Nuevas tecnologías, policía y prevención del delito* (trabajo fin de grado). Universitat Jaume I, España.

Kitchen, T., y Schneider, R. H. (2007). *Crime prevention and the built environment*. Routledge.

Lazo, A. D. (2018). *El endurecimiento de las penas no disminuye la acción delictiva.* (trabajo fin de grado). Universidad de San Martín de Porres, Lima, Perú.

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META group research note*, *6*(70), 1.

Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. *Boletín Oficial del Estado*, *294*, de 06 de diciembre de 2018, BOE-A-2018-16673.

Marín de Espinosa-Ceballos, E., Esquinas-Valverde, P., Zugaldía-Espinar, J., Moreno-Torres-Herrera, M., y Ramos-Tapia, M. (2016). *Lecciones de derecho penal Parte General* (3rd ed). Valencia, España: Tirant lo Blanch.

Malleson, N., y Andresen, M. A. (2015). The impact of using social media data in crime rate calculations: shifting hot spots and changing spatial patterns. *Cartography and Geographic Information Science*, 42(2), 112-121.

Meliani, L. (2018). Machine Learning at PredPol: Risks, Biases, and Opportunities for Predictive Policing. RC TOM Challenge 2018. Harvard Business School, Boston, Estados Unidos.

Mellón, J. A., Jiménez, G. A., y Rothstein, P. A. (2017). Populismo punitivo en España (1995-2015): presión mediática y reformas legislativas. *Revista española de ciencia política*, (43), 13-36.

Méndez, A. F. (2020). *Creación de procedimiento de big data para el Ministerio de Justicia y del Derecho.* [Monografía]. Repositorio Institucional UNAD. https://repository.unad.edu.co/handle/10596/34035

Monleón-Getino, A. (2015). El impacto del Big-data en la Sociedad de la Información. Significado y utilidad. *Historia y comunicación social*, *20*(2), 427-445.

Montraveta, S. C. (2015). ¿Eficacia preventiva general intimidatoria de la pena? *Revista Electrónica de Ciencia Penal y Criminología, 17-18*, 1-44.

Muñoz, J., Molero-Castillo, G., y Benítez-Guerrero, E. (2017). Método de fusión de datos de fuentes heterogéneas para mantener la consistencia de datos. *Res. Comput. Sci.*, 139, 33-46

Newman, O. (1995). Defensible space: A new physical planning tool for urban revitalization. *Journal of the American Planning Association, 61*(2), 149-155.

Nocetti, F. G. (2017). Ciencia de datos y big data. *Nexos: Sociedad, Ciencia, Literatura*, 472. Obtenido de https://www.nexos.com.mx/?p=31892

Ortiz de Urbina, I. (2004). Análisis económico del derecho y política criminal. *Revista de Derecho Penal y Criminología, 2*, 31-73.

Paredes-Moreno, A. (2015). Big Data: Estado de la cuestión. *International Journal of Information Systems and Software Engineering for Big Companies (IJISEBC)*, *2*(1), 38-59.

Parlamento Europeo (2017a): Resolución del Parlamento Europeo, 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))

Parlamento Europeo. (2017b). Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL)).

Pérez-López, C., y Santín González, D. (2007). *Minería de datos. Técnicas y herramientas: técnicas y herramientas*. Madrid, España: Paraninfo.

Pérez, J. A. (2011). La explicación sociológica de la criminalidad. *Derecho y cambio social*, *7*(22). Recuperado de http://www.derechoycambiosocial.com/revista022/explicacion_sociologica_de_la_criminalidad.pdf

Perry, L. Mclnnis, B., Price, C. Smith, C., y Hollywood, S. (2013). Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations, Santa Monica, California: RAND Corporation.

Pires, A. (2007). Une "utopie juridique" et politique pour le droit criminel moderne? *Criminologie, 40(2),* 9-18.

PredPol. (2021). *The Three Pillars of Predictive Policing*. Dsiponible en: http://www.predpol.com/law-enforcement/. Acceso el 30 de abril de 2021.

Puig, S. M., (1982). *Función de la pena y teoría del delito en el Estado social y democrático de Derecho*. Barcelona: Bosch.

Puig, S. M., (2003). *Introducción a las bases del derecho penaI*. Montevideo: B de f.

Puyol-Moreno, J. (2014). Una aproximación a Big Data. *Revista de Derecho UNED, 14*, 471-505.

Ramírez, J. B., y Mallafré, H. H. (1980). Pena y estado. *Papers: revista de sociologia*, 13, 97-128.

Reglamento (UE) 679/2016, de 27 de abril, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos. Diario Oficial de la Unión Europea, nº L119/1, de 4 de mayo de 2016.

Rodríguez-Horcajo, D. (2019). Pena (Teoría de la). *EUNOMÍA. Revista En Cultura De La Legalidad*, 16, 219-232.

Roxin, C. (1976). Sentido y límites de la pena estatal. En Luzón Peña D. M. (trad.) *Problemas básicos del derecho penal* (pp. 11-36). Reus.

Sancho-López, M. (2018). El derecho al olvido en el Big data: nuevos retos para la protección de la privacidad (Tesis Doctoral, Universitat de Valencia).

Santos, J. C. R., Ruiz, R., y Gilbert, K. (2006). Minería de datos: Conceptos y tendencias. *Inteligencia Artificial: Revista Iberoamericana de Inteligencia Artificial, 10 (29), 11-18.*

Shaw, M. (2011). *Manual sobre la aplicación eficaz de las Directrices para la prevención del delito*. Naciones Unidas, ONU.

Smith, M. (2018). Can we predict when and where a crime will take place? *BBC News.* Recuperado de https://www.bbc.com/news/business-46017239

Smith, P., Gendreau, P., y Goggin, C. (2002). The effects of prison sentences and intermediate sanctions on recidivism: General effects and individual differences. Ottawa: Solicitor General Canada.

Solove, D. J. (2007). I've got nothing to hide and other misunderstandings of privacy. *San Diego L. Rev.*, *44*, 745-772.

Sozzo, M. (2000). Seguridad urbana y tácticas de prevención del delito. *Cuadernos de jurisprudencia y Doctrina Penal*, *10*, 17-82.

Suárez, S. Y. S. (2019). Big Data: Un paso hacia el futuro. *Revista Neuronum*, 6(1), 169-172.

Summers, L. (2009). Las técnicas de prevención situacional del delito aplicadas a la delincuencia juvenil. *Revista de derecho penal y criminología*, 1, 395-409.

Tumasjan, A., Sprenger, T., Sandner, P., y Welpe, I. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. *Proceedings of the International AAAI Conference on Web and Social Media, 1(4),* 175-185.

Turner, G., Brantingham, J., y Mohler, G. (2014). Predictive policing in action in Atlanta, Georgia. *The Police Chief, 81*, 72-74.

Umaña-Hernández, C. E. (2015). Prevenciones sobre la prevención: algunas consideraciones desde la criminología. *Política criminal y "prevención".* Universidad externado de Colombia. doi:10.4000/books.uec.1139.

Unión Europea. (2012). Carta de los Derechos Humanos de la Unión Europea del 26 de octubre, 2012/C 326/02.

Valdés, F. V. (2018). *Big Data: cómo afecta a la privacidad de los ciudadanos* (trabajo fin de master). Universidad Oberta de Catalunya, España.

Valls-Prieto, J. (2018). *Problemas jurídico penales asociados a las nuevas técnicas de prevención y persecución del crimen mediante inteligencia artificial*. Madrid, España: Dykinson.

Van´t-Wout, E., Valenzuela, E., Asahi, K., Pieringer, C., Torres, D. y Larroulet, P., (2019). Big data para la identificación de comportamiento criminal. En: Centro de Políticas Públicas UC (ed), Propuestas para Chile. Concurso de Políticas Públicas 2018. Santiago: Pontificia Universidad Católica de Chile, (pp. 49-78).

Varol, O., Ferrara, E., Davis, C., Menczer, F., & Flammini, A. (2017, May). Online human-bot interactions: Detection, estimation, and characterization. En *Proceedings of the International AAAI Conference on Web and Social Media, 1(11),* 280-289.

Vilalta-Perdomo, C. J. (2017). *Información para la prevención del delito y la violencia*. Quito: Banco Interamericano de Desarrollo.

Villalobos-Fonseca, H. (2020). El desarrollo tecnológico en materia policial: una receta de éxito para la prevención del delito. *Revista de Relaciones Internacionales, Estrategia y Seguridad*, *15*(1), 79-97.

Wang, H., Kifer, D., Graif, C., y Li, Z. (2016). Crime rate inference with big data. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 635-644).

Weikart, D. P., y Schweinhart, L. J. (1997). *High/Scope Perry Preschool Program.* En G. W. Albee & T. P. Gullotta (Eds.), *Issues in children's and families' lives, Vol. 6. Primary prevention works* (pp. 146–166). Sage Publications, Inc.

Williams, M. L., Burnap, P., y Sloan, L. (2017). Crime sensing with big data: The affordances and limitations of using open-source communications to estimate crime patterns. *The British Journal of Criminology, 57*(2), 320-340.