El analista de inteligencia en la era digital

Coordinadores:
Félix Arteaga, David Ríos Insúa

El analista de inteligencia en la era digital

Coordinadores:
Félix Arteaga, David Ríos Insúa

REAL INSTITUTO
elcano
ROYAL INSTITUTE

ISBN: 978-84-92983-37-7

Índice

Presentación

El libro que el lector tiene en las manos surgió de la colaboración del Real Instituto Elcano (RIE) con el Centro Nacional de Inteligencia (CNI), iniciada a partir del lanzamiento del proyecto de Innovación Metodológica en Análisis de Riesgos para Inteligencia, en el que participaron varios de los autores. Dicho proyecto pretendía plantear algunas mejoras significativas en las técnicas de análisis en inteligencia basándose en los últimos avances de las ciencias matemáticas, de la computación, de la ciencia cognitiva y de las técnicas de juicio experto. Esta colaboración hizo posible un fructífero proceso de aprendizaje mutuo para ambas organizaciones, que comparten el reto de ofrecer información útil, y en el menor tiempo posible, al decisor político.

La obra refleja asimismo el objetivo fundacional del RIE de contribuir en la medida de lo posible a los debates clave para el futuro de la sociedad española. En el ámbito de la transformación digital, por ejemplo, investigadores como Andrés Ortega y Félix Arteaga han ayudado a impulsar una estrategia más coordinada e inclusiva entre los diferentes actores del ámbito empresarial, académico y público. Entre sus contribuciones más señaladas, cabe destacar el documento titulado "Hacia un ecosistema español de Inteligencia Artificial (IA): una propuesta",[1] en el que los autores proponen la construcción de un *hub* de IA en España, que cuente con la participación de actores gubernamentales, económicos y académicos implicados en la digitalización, y que ha servido para orientar algunas de las prioridades de la nueva Estrategia Nacional de IA. También cabe subrayar las aportaciones del RIE a los trabajos de Think20 (T20), la red de *think tanks* de los países que forman parte del G20, tales como el documento de reflexión elaborado para promover la inclusión de la llamada justicia tecnológica en la agenda de dicho organismo.[2]

Este libro es el fruto de la inteligencia colectiva, y no hubiese sido posible sin la contribución de numerosos expertos, investigadores y analistas del ámbito nacional e internacional. Lo han hecho desde diferentes disciplinas (matemáticas, estadística

1 Félix Arteaga y Andrés Ortega (2019), "Hacia un ecosistema español de Inteligencia Artificial: una propuesta", *Elcano Policy Paper*, junio, https://www.realinstitutoelcano.org/policy-paper/hacia-un-ecosistema-espanol-de-inteligencia-artificial-una-propuesta/019.

2 Andrés Ortega, Francisco Andrés y Yarik Turianskyi (2018), "Justicia tecnológica: un objetivo para incluir en la agenda del G20", Discussion Paper, *Economics e-journal*, http://www.economics-ejournal.org/economics/discussionpapers/2018-58.

bayesiana, ciencias de la computación y psicología), sectores (público, privado y académico) y áreas de especialización (inteligencia artificial, analítica de datos, teoría de la decisión, prospectiva, estrategia, seguridad y defensa, análisis de redes y análisis de riesgos).

En el ámbito español destacan las contribuciones de Francisco Herrera, el experto más citado en computación e ingeniería, y uno de los autores de la estrategia española de innovación e investigación en inteligencia artificial. Por su parte, Aníbal Figueiras, uno de los mayores expertos de Aprendizaje Máquina de nuestro país, es quizá uno de nuestros pensadores más originales en lo que a la integración hombre-máquina se refiere. Gracias a la generosidad de sus autores y el trabajo de edición de David Ríos, los lectores en lengua española podrán tener acceso a los trabajos de Eva Chen, científica jefa de *Good Judgement,* la *start-up* surgida del exitoso proyecto en predicción de eventos financiado por el *Intelligence Advanced Research Projects Activity* (IARA) de EEUU; del profesor Ahti Salo, uno de los mayores expertos en el campo del análisis de la decisión; y de Fabrizio Ruggeri, director científico del CNR-IMATI de Milán y uno de los mayores expertos europeos de estadística bayesiana. También ha sido muy valiosa la contribución de uno de nuestros cerebros exiliados en EEUU, Jesús Ríos, investigador del Centro de Investigación de IBM.

La aparición de una comunidad de investigadores en IA no se circunscribe únicamente al ámbito académico. El sector privado español ya es un referente en el sector de Fintech y la ciencia de datos aplicada a la empresa. Tres buenos ejemplos de ello son Juan Murillo, responsable de Divulgación Analítica en el área de *Data Strategy* y *Data Science* del BBVA; Carme Artigas, experta internacional en *Big Data* e IA y la fundadora de *Synergic Partners* (actualmente empresa de Telefónica), que fue pionera en Europa en la analítica de datos para la toma de decisiones empresariales; y Francisco Soler, quien compagina una prometedora labor académica con su actividad en el ámbito empresarial a cargo de proyectos de IA y Analítica de datos.

La obra, sin embargo, hubiese quedado incompleta si no fuera un libro escrito para analistas y por analistas de inteligencia. Los Servicios de Inteligencia gubernamentales, tensionados por la necesidad de desarrollar nuevas y mejores funciones (estratégicas, operativas y ahora preventivas o anticipatorias) pero con menores recursos, se encuentran en un momento crítico, tal y como señalan Antonio Fonfría y Félix Arteaga en su texto. Fonfría es uno de los mayores expertos de nuestro país en economía de la Defensa y conoce como pocos el coste de las malas decisiones en el ámbito de la planificación. En su capítulo, señalan la necesidad de reorganizar

y mejorar las fuerzas internas (talento, innovación y dirección) ante el dilema de cooperar o competir con nuevos actores privados que gozan de mayores (y mejores) recursos. Invertir en el capital humano es la recomendación de Ana Páramo, analista en Seguridad y formadora de analistas durante dos décadas en el CNI. Páramo contribuyó de forma destacada al éxito del proyecto de Innovación Metodológica en Análisis de Riesgos para Inteligencia, y representa la mejor versión de una analista de "la casa" con vocación de servicio al interés general.

El campo de la prospectiva, "la ciencia que estudia el futuro para comprenderlo e influir en él", como ha dejado escrito Gasto Berger, creador de la Escuela francesa de Prospectiva y Análisis, ha tardado en echar raíces en nuestro país, pero comienza a florecer a raíz de la toma de conciencia, a nivel tanto político como empresarial, del coste económico y social de acontecimientos no previstos de gran impacto. En poco tiempo, estos "cisnes negros" se han convertido en un elemento central en la nueva cultura de inteligencia. Así pues, la gran crisis de 2008, la llegada de Trump al poder, la pandemia del COVID-19 y la invasión rusia de Ucrania nos ponen en alerta y modelan nuestras expectativas sobre el resultado del devenir histórico. Sin embargo, y más allá de la epistemología, autores como Taleb sostienen que tales fenómenos no son sólo categorías filosóficas, sino que tienen base empírica, tienen un rastro de "miguitas" estadísticas por descubrir. A seguir ese rastro se dedica precisamente el capítulo II ("La inteligencia prospectiva") y a saber dónde buscar esas "miguitas" se desarrolla en el capítulo III ("Análisis de riesgos").

En las últimas décadas, el análisis de riesgos ha sido el paradigma dominante en el diseño y la operativa de diferentes sectores industriales, financieros, energéticos y, en definitiva, en la sala de máquinas de nuestro sistema económico. La importancia vital que para una empresa tiene a largo plazo la implantación de un sistema de gestión de riesgos es analizada en este libro por Virginia Leal (gerente del Sistema Integrado de Riesgos de Repsol). Desde el sector financiero, que goza de mayor experiencia y madurez en el tratamiento de datos que otros, Ignacio Carnicero (responsable de *BBVA Global Risk Management Analytics*), aun reconociendo la ventaja del uso masivo de datos en la toma de decisiones, nos advierte que perder "el control de una decisión es un gran cambio para muchas personas".

El sector privado ha generado en los últimos años en nuestro país una nueva generación de analistas de inteligencia, ávidos por aprender e incorporar nuevas metodologías. José María Blanco, que inició su carrera en el Centro de Análisis y Prospectiva de la Guardia Civil, es actualmente director de Ciber-inteligencia Estratégica en Prosegur.

Con un enfoque parecido, Jessica Cohen, coordinadora de la Unidad de Análisis en Seguridad Internacional de Prosegur, practica un escepticismo crítico y una autoexigencia que empuja a ambos a explorar nuevas respuestas al difícil arte de la prospectiva. Por su parte, Jordi Serra, director de investigación del *Center for Postnormal Policy & Future Studies* de Londres, nos advierte de la importancia de ser conscientes de nuestra ignorancia ante incertidumbres profundas, de las que no conocemos ni su dirección ni su posible impacto.

Sin esta lección de humildad será muy difícil adaptarse a los cambios que ya inciden sobre nuestras vidas. Este libro pretende ante todo servir de receta para la supervivencia y el éxito. En un mundo cada vez más complejo y volátil, los grandes actores políticos, económicos y sociales necesitan diseñar su propia estrategia y planificar sus decisiones. También necesitamos ser conscientes de que el desarrollo tecnológico de nuestro país definirá en no poca medida nuestra posición relativa futura en el tablero global. Cuando ya se han cumplido 50 años de la llegada del hombre a la Luna en una carrera espacial alentada por la competición entre las dos superpotencias, hoy volvemos a encontrarnos en una competición a escala global por el dominio de tecnología. En las próximas décadas, los países que construyan y consoliden una cultura de innovación en IA serán los mejor situados para preservar su estilo de vida, su bienestar económico y su seguridad nacional. Esperamos, por tanto, que esta obra colectiva sirva al menos para alertar a sus posibles lectores sobre la magnitud de los retos que ello conlleva, así como a esbozar algunas posibles vías de actuación al respecto.

Capítulo I
La inteligencia en el siglo XXI

¿Hacia dónde se dirigen los servicios de inteligencia?

Félix Arteaga y Antonio Fonfría

Introducción

Tradicionalmente los servicios de inteligencia de los países han venido desarrollando su actividad basándose en un capital humano muy conocedor de ciertos riesgos, zonas geográficas, estrategias de países, relaciones empresariales, etc. Hasta ahora este tipo de especialización ha dado sus frutos, pero la variedad de riesgos actual junto con la introducción de nuevas tecnologías impone una revisión.

Los retos a los que se enfrentan los servicios de inteligencia hoy implican la necesidad de importantes cambios de distinta índole cuyo objetivo básico es mejorar el grado de seguridad que ofrecen a sus ciudadanos, a quienes han de proteger, así como mantener los modelos de sociedad vigentes. Yendo más allá, la sociedad digital, el nuevo conocimiento aplicado y las novedades que ello implica en términos de formas de análisis han de adaptarse a la sociedad. Sin embargo, ¿el uso de nuevas tecnologías es realmente una ventaja o, por el contrario, puede convertirse en un arma de doble filo para los servicios de inteligencia? ¿Qué papel desempeñan las sociedades cuyo grado de información es mucho mayor hoy que hace años y son más sensibles a temas como la intromisión a través de las nuevas tecnologías? ¿Hasta qué punto han de transformarse las estructuras de inteligencia para responder a estas nuevas situaciones? ¿Será posible cuantificar la eficacia y eficiencia de los Servicios de Inteligencia en su labor de aportar seguridad nacional? Estas cuestiones se abordan de manera introductoria a lo largo de las siguientes páginas.

Los grandes cambios de contexto: efectos sobre los servicios de inteligencia

La evolución de los Servicios de Inteligencia viene configurada por su actitud ante los cambios del contexto estratégico donde actúan. Cuando consideran que los cambios de contexto van a afectar a su función, adoptan medidas proactivas o preactivas, respectivamente, según traten de adaptarse o anticiparse a los cambios.

En caso contrario, los servicios seguirán en modo reactivo, sin capacidad de prevenir a tiempo a sus usuarios y consumidores de inteligencia.

El contexto estratégico hacia el que avanzan no es otro que el de la sociedad global digital hacia la que se encamina el mundo. No es función de los servicios alertar a las sociedades sobre las consecuencias de los riesgos tecnológicos derivados del progreso,[3] riesgos que exceden a la capacidad protectora de los Estados y con los que deben aprender a convivir sus sociedades. Sin embargo, sí deben advertir del efecto de esos riesgos –y, obviamente de otros muchos–, sobre las funciones tradicionales de los servicios.

La proliferación de riesgos, según muestran los análisis de las distintas estrategias de seguridad nacional, el Foro Económico Mundial y tantos otros análisis públicos y privados de riesgos[4] obliga a concentrar la actuación de los Servicios de Inteligencia en el análisis y prevención de los riesgos prioritarios para la seguridad y prosperidad de cada sociedad. Éstas cuentan con otros servicios dedicados a agilizar la recuperación (resiliencia) tras el impacto de esos riesgos, pero dependen de la anticipación de los servicios de inteligencia para prevenir su materialización.

La misma globalización facilita, por un lado, la proliferación y trasnacionalización de los multiplicadores de riesgos y, por otro, limita la capacidad de actuación de los actores tradicionales para gobernar aquellos. Los instrumentos tradicionales de gobernanza como la regulación, la atribución, la sanción o la rendición de cuentas van perdiendo utilidad por la incapacidad de las instituciones de consensuar y exigir el cumplimiento de las normas (*enforcement*). Al mismo tiempo, la agenda de seguridad se define por actores distintos de los usuarios gubernamentales de los servicios de inteligencia y se "segurizan" cuestiones bajo presión social o electoral de las sociedades, como son la inmigración o la protección de las infraestructuras críticas. Los servicios de inteligencia poseen un papel relevante en identificar el desfase entre necesidades y capacidades de seguridad, para lo que necesitan mejorar sus capacidades de analizar riesgos para establecer prioridades de actuación.[5]

3 Uno de los textos más interesantes sobre este tema es el de Beck, U. (2006) *La sociedad del riesgo global.* Ed. Siglo XXI, Madrid.

4 Un claro ejemplo es la publicación de *World Economic Forum* "The Global Risks Report 2019". Ginebra. En ella se plantean riesgos de diversa índole –económicos, medioambientales, geopolíticos, sociales y tecnológicos-, que suman un total de 31 riesgos a los que habrá de enfrentarse cada sociedad en mayor o menor medida. Igualmente, la *National Intelligence Strategy* de 2019, publicada por EEUU, muestra un amplio conjunto de riesgos adicionales, como el terrorismo o las ciber-amenazas.

5 La mejora de los procedimientos de análisis es imprescindible tanto para establecer prioridades como para asignar recursos o declinar responsabilidades por falta de medios.

La digitalización también genera retos y oportunidades para las sociedades globales. Las tecnologías de la información y comunicación multiplican la conectividad global, un proceso acelerado que erosiona la seguridad. Primero, la conectividad facilita la propagación de riesgos a sectores cuya protección dependía de la separación. Por ejemplo, y en el ámbito industrial, los responsables de seguridad no sólo tienen que preocuparse de la seguridad física, industrial o en un nuevo modelo más comprensivo de seguridad digital. Les corresponde a los servicios de inteligencia ocuparse por defecto de las carencias y limitaciones del sector privado para hacer frente a estos nuevos retos en los servicios esenciales e infraestructuras críticas, entre otros. En el mismo sentido, y debido a la interacción entre riesgos, el perfil de los miembros de los servicios crece, por un lado, hacia la especialización y, por el otro, hacia la combinación de talentos para la innovación (polimatía)[6]. Segundo, la conectividad potencia la posibilidad de riesgos sistémicos en los servicios e infraestructuras globales, independientemente de su origen o de su protección.[7] Tercero, la digitalización reduce y comprime los tiempos de alerta, reacción y decisión, una responsabilidad en la que los servicios de inteligencia deben asesorar a los responsables de la toma de decisiones. Así, la globalización produce perdedores que cuestionan el orden global, internacional, nacional o social establecido. En definitiva, la sociedad global, tal y como exponía Ulrich Beck, genera sus propios riesgos debido al progreso tecnológico.

Por otra parte, la globalización y la digitalización favorecen la aparición de tecnologías disruptivas que representan factores de riesgos para la prosperidad y estabilidad de las sociedades. La inteligencia artificial, la robotización, la informática cuántica o el *big data*, entre muchas otras tecnologías, ponen en riesgo de exclusión a las sociedades, regiones, industrias o individuos que no puedan aprovecharse de ellas, tanto por razones de prosperidad, no participando en sus beneficios, como de seguridad, al ver cómo aumenta su diferencial de inseguridad. La hibridación de actores y amenazas crea nuevas fuentes de riesgo cuya naturaleza y alcance corresponde analizar a los servicios de inteligencia, al igual que se vieron obligados a realizar tras la irrupción del terrorismo yihadista o los ciberataques en el pasado reciente. También crean factores de oportunidad porque permiten a los servicios operar a una velocidad y en una escala que supera las posibilidades humanas.

6 *Polimatía*, Deusto Bussines School y 3M, 2017.

7 Una estimación del impacto económico del ciberataque de WannaCry puede consultarse en Deloitte (2017) *¿Qué impacto ha tenido el ciberincidente de WannaCry en nuestra economía?* Madrid.

Los Servicios de Inteligencia, que han venido monitorizando a los potenciales agresores, deben ahora dedicar atención también a las fuentes de fractura y radicalización interna. La desigualdad, marginación, inmigración o urbanización disminuyen la cohesión social y el sentido de identidad sobre el que se asientan las comunidades nacionales de seguridad. La desafección, el activismo, la emulación y el resentimiento cuestionan la eficacia de los sistemas ordinarios de control social y facilitan el proceso de desestabilización de sociedades que se consideraban seguras (revoluciones de colores, primaveras, revueltas sociales, populismos…). La conjunción de los cambios y, sobre todo, el alto ritmo de su interacción plantea a los servicios de inteligencia el problema del conocimiento variable de los actores, estados de opinión, tendencias y escenarios de evolución, tanto de los potenciales agresores como de sus propios usuarios y beneficiarios. La exposición de las sociedades e individuos frente a la desinformación, las operaciones de influencia, la manipulación de las redes de comunicación social o el carácter viral de sus consecuencias obliga a monitorizar los procesos de cambio si se desea advertir a tiempo de sus consecuencias. La rápida evolución (viralidad) y el elevado impacto (efecto) de los riesgos, obligan a desarrollar las funciones de inteligencia con la anticipación necesaria. La digitalización plantea el reto de procesar una cantidad de información sobreabundante en un tiempo cada vez más escaso (cronointeligencia).

En consecuencia, la inteligencia estratégica debe orientarse a dirigir políticas y medidas en el ámbito de la seguridad (*Intelligence-led)*, dando prioridad a la prevención, al conocimiento situacional y a la previsión estratégica. Por tanto, ya no se trata sólo de que los servicios añadan valor o ventaja al proceso de toma de decisiones.

Como refleja la Estrategia de Inteligencia Nacional de Estados Unidos de 2019, sus Servicios deben ocuparse de la seguridad nacional, la prosperidad y la superioridad tecnológica, un activo este último que se había limitado en los últimos tiempos a funciones de contraespionaje, pero que con la competencia geoeconómica global en curso adquiere un significado estratégico (proteccionismo, activos estratégicos, controles de exportación…). Para ello se proponen integrar, más que coordinar, las actividades de su comunidad, potenciar la innovación internamente, reforzar su colaboración con socios e incrementar la trasparencia, la confianza y la rendición de cuentas.

Los Servicios de Inteligencia deberán seguir atendiendo las funciones tradicionales, ocuparse de las emergentes y continuar con las operaciones. Más funciones, más

urgencia y menos certidumbre, con incrementos limitados de recursos, plantea un problema de reorganización de los recursos en los servicios (gestión, talento, innovación, cooperación, integración). La adaptación es crítica porque los Servicios deberán competir con fuentes de inteligencia privadas, muchas de ellas con mayores recursos y menos restricciones para generar inteligencia estratégica (inteligencia como servicio), lo que reduce la ventaja comparativa de los servicios frente a los usuarios públicos y privados en ese apartado. La situación plantea el dilema de competir o cooperar con las nuevas fuentes de inteligencia.

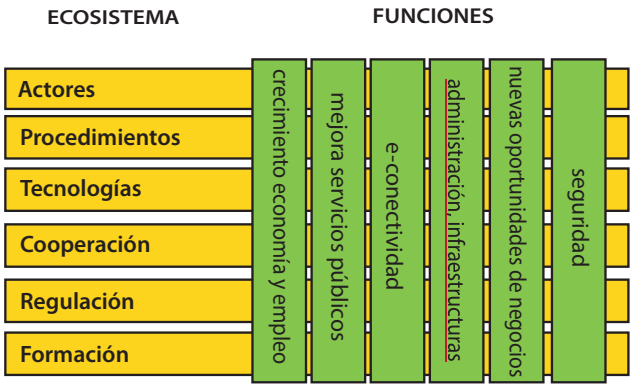
En las funciones de anticipación, frente a los nuevos retos, los Servicios disponen de menor ventaja comparativa de partida, especialmente en la medida que la curva de aprendizaje de las nuevas misiones sea prolongada. La complejidad y el menor conocimiento de los nuevos objetos de inteligencia obliga a desarrollar instrumentos de análisis experimentales, cuantitativos y en entornos de tecnologías disruptivas (inteligencia artificial, informática cuántica, ciberseguridad y *big data*) para recortar el desfase entre las expectativas y las capacidades de inteligencia anticipativa.[8]

El apoyo a las operaciones de inteligencia se mantendrá relativamente continuista mientras los Servicios no asuman nuevos ámbitos de operaciones por defecto. No obstante, las operaciones también se verán influidas por los cambios tecnológicos que afectarán al control de la información y a la cobertura de las mismas.

En la Figura 1 se observa el efecto combinado de los procesos mencionados. En las filas se muestran los elementos principales del nuevo ecosistema de inteligencia. En la fila de actores figuran nuevos y viejos productores y consumidores de inteligencia, públicos y privados que diversifican los ciclos de inteligencia. Junto a los procedimientos tradicionales o a sus innovaciones aparecen instrumentos tecnológicos nuevos o mejorados que multiplican la capacidad de procesamiento instalada. La colaboración con terceros y entre servicios es ahora más importante que antes, así como la regulación de las actividades de los servicios para preservar la confianza social en su función. Finalmente, la formación, reciclaje, captación y retención de talento representan un reto para los servicios.

8 La aplicación experimental de la Inteligencia Artificial a los Servicios de Inteligencia se desarrolla en centros experimentales compartidos por otros servicios de seguridad nacional que disponen de suficientes datos para facilitar el aprendizaje de las máquinas, a la espera de que pueda generalizar su empleo (*waiting for AI*), "Artificial Intelligence and National Security", CSIS Washington, noviembre 2018, pp. 27-34.

Figura 1. Efectos combinados



Fuente: elaboración propia.

En la columna de funciones, aparecen nuevas funciones, públicas, privadas o mixtas de inteligencia (el proceso de segurización podría suprimir funciones antiguas). Los Servicios podrán trabajar para y con múltiples usuarios, clientes y socios, integrando comunidades y culturas de inteligencia de distinta tradición y perfil (públicas y privadas, externas e internas). Corresponde a los Servicios desarrollar nuevos métodos para integrar los distintos ecosistemas (redes, interoperabilidad, centros/tareas de fusión, formación múltiple, compatibilidad tecnológica, nuevas tecnologías…) y nuevos instrumentos de análisis, estructuras organizacionales, procedimientos de trabajo y formación (externalización y tecnologías comerciales) así como nuevas combinaciones de fuentes (inteligencia humana, tecnológica, fuentes cerradas y abiertas).

Como se ha mencionado anteriormente, los servicios de inteligencia no serán los únicos actores en los nuevos ecosistemas de inteligencia y deberán competir con otros productores por los usuarios y productos de especialización. Deberán buscar y defender nuevos nichos de excelencia donde gocen de ventaja comparativa.

Nuevos retos exigen nuevos planteamientos

La situación descrita impone la necesidad de cambios tanto en la forma de hacer inteligencia como en los medios que se utilizan y en el proceso de realización del ciclo de inteligencia. Obviamente esto implica que el resultado de cara al usuario también

habrá de ser distinto al actual, posiblemente con una mayor carga cuantitativa y un efecto superior en el proceso de toma de decisiones. Igualmente, habrán de anticiparse a las necesidades del "mercado" de usuarios de inteligencia y mantener su ventaja comparativa, cual es la de ofrecer productos de inteligencia en exclusiva.

La morfología de la Comunidad de Inteligencia tiende a ser más abierta. Las agencias requieren de conocimientos vinculados a ámbitos como el universitario o el industrial, algo que ya venían haciendo pero que se intensifica con la ampliación y profundización de los riesgos. La magnitud y diversidad de dichos riesgos impone la necesidad de lo multidisciplinar, del trabajo en equipo desde dentro y, hacia afuera, la interconexión de diversos agentes, fuentes y perspectivas. Este aumento de la complejidad requiere de una gestión de los recursos profesionalizada en la cual el objetivo de generación de inteligencia se muestra como una función en la que el número de variables a considerar es cada vez más amplio, yendo desde la gestión de los recursos humanos y el talento, al liderazgo, la asunción de ciertos riesgos tecnológicos y cambios en las ponderaciones de origen cuantitativo y cualitativo con la que se genera la inteligencia.

Junto a ello y, debido a la importancia de la conectividad y de las relaciones –por ejemplo, a través de las redes sociales–, entre diversos actores de la Comunidad de Inteligencia y el conjunto de la sociedad[9], la trasparencia se erige como un arma de doble filo para las agencias al encontrarse en su ADN la necesidad de discreción y un trabajo que no suele trascender socialmente. Sin embargo, la rendición de cuentas de las actividades que realizan es necesaria, lo cual conlleva cierto juicio social. La cuestión en este aspecto se encuentra en un punto entre la tensión generada por la exigencia social de trasparencia y las necesidades de las agencias de minimizar su exposición social. La conciliación de ambos aspectos es compleja ya que las sociedades exigen cada vez más, saber que se hace con sus recursos y qué actividades desarrollan de forma concreta los distintos agentes del sector público.

Así, las agencias de Inteligencia se encuentran obligadas a innovar en sus procesos y a incorporar innovaciones tecnológicas, como la IA, *Machine Learning* o análisis de datos, al objeto de mejorar sus capacidades competitivas para ofrecer un nivel de seguridad nacional más elevado a las sociedades. A su vez el uso de estos instrumentos permite una mayor eficiencia y eficacia en el desarrollo de sus actividades lo cual

9 Un análisis detallado de estos aspectos puede encontrarse en Petersen, K. L. y Ronn, K. V. (2019) "Introducing the special issue: bringing in the public. Intelligence on the frontier between state and civil society" *Intelligence and National Security*, vol. 34 (3).

redunda en una mejora de los niveles de seguridad generales. En este sentido, el paso de la lectura de datos al futuro procesamiento inteligente de la IA, ha de servir de aprendizaje para los servicios de inteligencia.

Sin embargo, la sociedad es cada vez más exigente, como se ha dicho, y toma posiciones con relación a las cuestiones de privacidad, temas éticos, etc., lo cual implica tener un cierto nivel de confianza en las agencias de Inteligencia. Esta confianza no se basa únicamente en lo que la sociedad dictamina, sino en la postura que el estamento político toma con relación a estos aspectos. Por ello se requiere un debate profundo sobre el tema. Dicho debate puede generar como resultado una estrategia de comunicación que fomente la confianza general de la sociedad.

También hay que considerar que las nuevas tecnologías como la IA se utilizan de manera que minimizan el volumen de información necesaria que ha de estudiar un analista, de forma que, aunque el volumen de datos sea muy elevado, la selección de aquellos más relevantes para el análisis posterior que realice un analista será más reducido. Además, el uso de la supercomputación permitirá la inteligencia adaptativa, aspecto que elevará la eficiencia general del proceso de generación de inteligencia.

En definitiva, este debate, que se encuentra abierto en lugares como el Reino Unido y los Países Bajos, ha de ser más global y concluir con una regulación que concilie los diversos aspectos –complejos–, que se han expuesto.

Eficacia y eficiencia en inteligencia

Además de los debates sociales, las crecientes amenazas y los riesgos cada vez más complejos exigen que hayan de tomarse en consideración, con creciente énfasis en el trabajo diario de las agencias de Inteligencia, aspectos que se encuentran vinculados a otro tipo de factores que definen el éxito de sus operaciones. Dos son los fundamentales: la eficacia y la eficiencia.[10]

El primero de ellos se refiere al grado de cumplimiento de un objetivo planteado previamente. En este sentido, el planteamiento del objetivo es el núcleo de la cuestión. Dos son los objetivos básicos en inteligencia. El primero de ellos, de mayor

10 En el caso de las políticas públicas un trabajo que analiza estos dos conceptos es el de K.M. Mokate (2001) "Eficacia, eficiencia, equidad y sostenibilidad: ¿qué queremos decir?", Documentos de Trabajo, I24, *Banco Interamericano de Desarrollo*, Washington.

rango, sería alcanzar el máximo grado de seguridad nacional, siendo este un bien superior y fundamental socialmente. El segundo se refiere a la consecución de diversos objetivos parciales de seguridad, como el control y desmantelamiento de células terroristas o de cualquier riesgo o amenaza concreto encargado a un servicio de inteligencia.

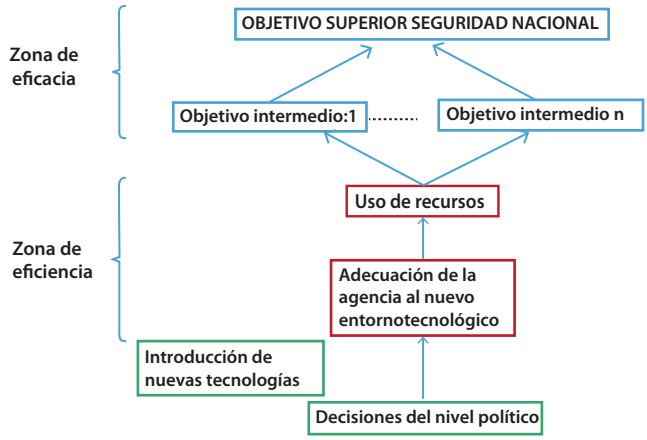
Con relación a la eficiencia, ésta se puede definir como la capacidad para alcanzar el objetivo planteado de forma que el uso de los recursos requeridos sea el menor posible. Esta perspectiva supone que el volumen de los costes de todo tipo –económicos y no económicos–, sean los adecuados. Se podrían incluir los recursos humanos, técnicos, o en general materiales, instalaciones, etc., y otros *a priori* no cuantificables, pero que influyen de forma sustancial en la consecución de los objetivos, como el factor tiempo.

La introducción de las nuevas tecnologías como la IA o el *Machine Learning* en el desarrollo y generación de inteligencia implica importantes cambios internos en las agencias (Figura 2). Estos han de modificar sus rutinas y formas de actuar y adecuarlas a las novedades tecnológicas tratando de alcanzar tanto una mayor eficacia en sus actividades, como una mayor eficiencia en el uso de los recursos. Si bien esto es aplicable a cualquier organización, lo es con más motivo a las agencias de Inteligencia ya que de ellas depende, en parte, lo que hemos denominado un bien social superior como es la seguridad del país y de sus ciudadanos. Bien es cierto, que se puede ser eficaz y no eficiente, pero esta situación es cada vez menos habitual y supone un bajo aprovechamiento de los recursos de todo tipo que las sociedades ponen en manos de las agencias. Adicionalmente, debido a que los servicios de todos los países incorporan de forma paulatina las nuevas tecnologías a sus sistemas de Inteligencia, no hacerlo implica una pérdida de capacidad competitiva y de confianza por parte de los países socios, lo cual va en detrimento de la seguridad.

La adecuación de las agencias al nuevo entorno tecnológico exige una cierta reorganización interna de los recursos. En este sentido, la forma de actuar de los analistas se ve modificada por los nuevos instrumentos que se encuentran a su disposición, lo cual implica que han de desarrollar otras capacidades diferentes, que las necesidades de determinados perfiles profesionales también se verán modificadas, que la relación entre recursos humanos y tecnológicos cambiará, al igual que las destrezas exigidas y la forma de trabajar entre analistas y con la información.

Por último, el nivel político ha de desempeñar un papel fundamental en todo el proceso ya que los estímulos e incentivos a realizar tan profundos cambios han de venir acompañados –incluso impulsados–, desde las instancias superiores y con la financiación suficiente. Sin embargo, en numerosas ocasiones el impulso proviene de abajo hacia arriba siendo más conscientes de las necesidades aquellos que se encuentran más cercanos al día a día.

Figura 2. Efectos de las nuevas tecnologías



Fuente: elaboración propia.

Conclusiones

A lo largo de estas líneas se ha tratado de poner de manifiesto el complejo panorama al que se enfrentan las agencias de Inteligencia, así como que el paradigma tradicional de análisis y de analista se están transformando. La complejidad proviene de muy diversas fuentes, relativas tanto a los cambios geoestratégicos, como a los tecnológicos y sociales. Los retos son de tal magnitud que requieren de enfoques muy diferentes a los que se han dado hasta ahora, e implican importantes transformaciones tanto internas, en el seno de las agencias, como externas, en sus relaciones con el conjunto de la Comunidad de Inteligencia y con otros países.

Las capacidades competitivas de las agencias pasan por una rápida adecuación al nuevo entorno tecnológico. Las exigencias de nuevos perfiles en los recursos humanos, la formación en tecnologías emergentes y la capacidad de absorción de las mismas, la rapidez en las respuestas ante riesgos y amenazas y la capacidad para generar inteligencia de elevado valor añadido que tengan un importante impacto en determinadas políticas nacionales –e incluso internacionales–, son los grandes retos actuales y futuros. Las agencias han de subirse a un tren en marcha cuya velocidad es ya elevada.

Félix Arteaga es investigador principal de Seguridad y Defensa del Real Instituto Elcano y profesor del Instituto Universitario General Gutiérrez Mellado, de la Universidad Nacional de Educación a Distancia (UNED). Oficial de la Escala Superior de las Fuerzas Armadas (en la reserva), ha sido profesor asociado de la Universidad Carlos III de Madrid e impartido clases de postgrado en las universidades Autónoma y Complutense de Madrid. También ha sido investigador principal del Instituto de Cuestiones Internacionales y de los institutos Duque de Ahumada y de Seguridad Interior de la Guardia Civil. Su investigación actual se orienta a evaluar el impacto de la globalización sobre la seguridad nacional e internacional, los cambios en la cultura estratégica de líderes y opiniones públicas, y la adaptación de las políticas de seguridad y defensa ante los nuevos retos para la gobernanza y estabilidad de las sociedades avanzadas.

Antonio Fonfría es profesor de Economía de la Defensa en la Universidad Complutense de Madrid, en el Instituto Complutense de Estudios Internacionales (ICEI) de la misma Universidad, en la Escuela Diplomática, en el Máster en Seguridad y Defensa (UCM-CESEDEN), y en el Máster Oficial en Seguridad y Defensa de la Universidad de Zaragoza entre otros. Ha sido director del Máster en Logística y Economía de la Defensa desarrollado entre la UCM y el Ministerio de Defensa, siendo actualmente director de la Cátedra UCM-ISDEFE. Es o ha sido asesor y colaborador del Instituto de Estudios Fiscales, del Consejo Económico y Social de la Comunidad de Madrid, del CESEDEN, del CESIA y del Ministerio de Defensa (DGAM), del Real Instituto Elcano, de la Foundation pour la Reserche Stratégique (FRS) de Francia, y del Instituto Español de Estudios Estratégicos. Miembro del Consejo Editorial de la Revista del IEEE, es autor de numerosas publicaciones relacionadas con la economía y la industria de la defensa, y ha participado y dirigido más de 20 proyectos de investigación nacionales y europeos. Ha sido ganador del European Award Citizenship, Security and Defence (2011) y del Premio de Defensa de Docencia concedido a la Cátedra Almirante D. Juan de Borbón, CESEDEN-UCM.

El analista de inteligencia de última generación

Ana Páramo

Introducción

La primera cuestión que nos asalta a la hora de abordar este capítulo es la delimitación conceptual del "analista de inteligencia". Existen discusiones doctrinales sobre si es una persona con amplios conocimientos en el ámbito del análisis, con dominio de las técnicas y de los procedimientos, o bien alguien conocedor del tema sobre el que debe analizar. Al margen de que sea posible defender ambas posturas, no cabe duda de que lo ideal es que el analista sea experto en ambos campos, a pesar de que somos conscientes de que no es fácil encontrar profesionales con ese perfil.

A pesar de lo anterior, y sobre todo en el mundo empresarial, es difícil que los analistas puedan permitirse ser expertos en un solo tema, ya sea geográfico (un país o una región) o transversal (tráficos ilícitos, terrorismo, etcétera).

Hoy en día, en el mundo globalizado en el que vivimos, un analista es aquel que sabe producir inteligencia a partir de información sobre cualquier asunto que se le plantee, lo que no significa que no deba documentarse antes de modo riguroso. Esto es, es capaz de aplicar las herramientas analíticas a distintas materias con la finalidad de reducir la incertidumbre y facilitar la toma de decisiones a los destinatarios de esa inteligencia.

Entornos VUCA: qué son y cómo desenvolverse

Además, es preciso dedicar unas líneas para entender un poco mejor los entornos VUCA[11] en los que se desenvuelven los analistas en la actualidad. En ellos existe volatilidad, incertidumbre, complejidad y ambigüedad:

- *Volatilidad*. Lo que es ahora puede dejar de serlo dentro de unos minutos; todo sucede con gran rapidez y, además, se sabe con inmediatez, independientemente del lugar del mundo en que ocurra. De este modo, el analista debe moverse con cautela a la hora de hacer afirmaciones y formular juicios e hipótesis.

11 De las siglas en inglés: *volatility, uncertainty, complexity y ambiguity.*

- Incertidumbre. Se acabó la época en la que una acción llevaba casi irremediablemente a una consecuencia. Ahora, proyectar una situación hacia el futuro supone hacer frente a un alto grado de incertidumbre. Sin embargo, al analista se le exige reducirla cuando difunde su inteligencia.
- Complejidad. Cualquier suceso debe contextualizarse y relacionarse con otros que, en ocasiones, pueden haber sucedido a miles de kilómetros de distancia y que, en apariencia, son independientes e inconexos.
- Ambigüedad. Vivimos rodeados de información, pero no es sinónimo de disponer de datos exactos y fieles a la realidad. El fácil acceso a internet y a las redes sociales permite que cualquier persona desde cualquier lugar pueda transmitir en tiempo real su percepción de un hecho, lo que fomenta la propagación de datos ambiguos, muchas veces contradictorios.

Pues bien, a la incertidumbre intrínseca al análisis de inteligencia, el analista ha de añadir estos atributos propios del mundo actual. Por eso, no basta con ser una analista a la vieja usanza, sino que se impone la necesidad de ser un analista de última generación, capaz de combinar técnica, ciencia, lógica y creatividad. A esta realidad caracterizada por los entornos VUCA se le puede hacer frente con una actitud VUCA[12]: visión, entendimiento, claridad y agilidad.

La visión es precisa para anticiparse a los acontecimientos e, incluso, influir en su evolución; el entendimiento exige empatía, capacidad para conocer y comprender otros puntos de vista; la claridad permite encajar todas las piezas del puzzle con independencia de su forma, color y tamaño; y, por último, la agilidad nos garantiza una perfecta adaptación a los cambios que se produzcan, tanto si constituyen una sorpresa como si se han previsto con antelación.

Actitud del analista en el mundo actual

Además, la inteligencia es el producto de aplicar herramientas de análisis a la información obtenida por diversas fuentes, de manera que no es una mera transmisión de los datos debidamente ordenados, sino que precisa un tratamiento profesional basado en el rigor, la lealtad y la ética.

12 Por las siglas en inglés, VUCA: *Vision, understanding, clarity y agility.*

En la actualidad, asistimos a un escenario en el que proliferan los analistas (financieros, económicos, políticos, sociales…) que, con frecuencia, se limitan a trasladar información con cierta valoración personal, basada en la intuición, en el conocimiento y en la experiencia, y con una tímida proyección hacia el futuro que descansa en la lógica y el pensamiento lineal. En pleno siglo XXI se requieren profesionales de la inteligencia comprometidos con el noble oficio de facilitar la toma de decisiones, muchas veces estratégicas, a la organización o al Estado al que sirven. Es aquí donde entra en acción el analista de inteligencia de última generación.

Esta nueva versión del analista hace referencia a su carácter activo –proactivo diría yo–, inquieto intelectualmente, crítico, dispuesto a someterse a actualizaciones continuas y contrario al anclaje y al "es que siempre se ha hecho así". Debe estar dispuesto a romper moldes e inercias, a enfrentarse, por qué no decirlo, a algunos analistas experimentados que se resisten al cambio, que niegan la profesionalidad y la eficacia de las nuevas generaciones y que, incluso, ven peligrar y cuestionarse sus procedimientos de producción de inteligencia.

Si el mundo ha cambiado, ¿por qué no van a hacerlo los analistas en tanto que profesionales que están obligados a conocer la realidad, comprenderla, interpretarla y explicarla? De nada serviría analizar el hoy y anticiparse al mañana sólo con procedimientos y métodos de ayer.

El pensamiento crítico debe presidir el quehacer diario del analista, guiar su mente y hacer extensivo su empleo a otros colegas. No dar nada por cierto hasta no someterlo a verificación, saber ponerse en el lugar del cliente para poder satisfacer sus necesidades de inteligencia, cuestionarse –eso sí, sin llegar a la paranoia– los juicios propios y los ajenos, tener claro el objetivo para el que realiza el análisis, no dejarse llevar por primeras impresiones ni por intuiciones personales… Todo eso se espera de un analista comprometido, con vocación y leal al fin al que sirve.

Pero no es esta la única cualidad que debe tener el analista de última generación. Ha de actuar con objetividad y rigor intelectual, de acuerdo con una metodología que le permita justificar sus afirmaciones y revisar la línea analítica seguida en busca de posibles errores. La autoconfianza, que no la soberbia, le ayudará a mantener y defender sus tesis sin temor; la empatía le asistirá a la hora de entender otros modos de pensar; el dominio del lenguaje hará que la transmisión de la inteligencia sea eficaz en forma y fondo; la creatividad le hará sencilla la difícil tarea de imaginar nuevas vías de solución a problemas diferentes y complejos; ya no basta con esperar

a que sucedan los acontecimientos, hay que ir en su busca, incluso provocar que ocurran; trabajar en solitario es producto de otra época, la capacidad para trabajar en equipo ha de ser una cualidad primordial del analista de inteligencia, y aún podrían añadirse más aptitudes que serían exigibles al analista de última generación…

Salvo excepciones puntuales, el analista no nace, sino que se hace, y esto se traduce irremediablemente en la necesidad de formación continua a la que está sometido. Incluso podríamos afirmar que nunca llega a estar hecho del todo. En este punto, me gustaría mencionar la idea de un colega según la cual hoy ya no es posible que existan analistas expertos puesto que la rapidez de los cambios impide estar a la última en todo momento. La formación debe presidir el antes y el durante el ejercicio de la profesión. Hay que obtener conocimientos que le permitan llegar a ser un analista; hay que practicar para adquirir destreza en el arte del análisis, y hay que seguir participando en actividades formativas variadas para mantenerse actualizado en un mundo en el que la ausencia de la formación lleva a la obsolescencia temprana.

Formación del analista

Desde hace unos años, en España está aumentando la oferta de formación en materia de inteligencia, si bien no hay que cerrarse a la proveniente de otros países. Es preciso practicar también el autoaprendizaje, para lo cual internet es sin duda un buen recurso, pero sin que el idioma suponga una traba.

En relación con lo anterior, el analista de última generación debe estar formado, al menos, en las siguientes materias:

Conceptos básicos de inteligencia

Se tiende a impartir y recibir cursos en materia de análisis de inteligencia en los que se abordan desde el principio técnicas de análisis; lo mismo ocurre con los monográficos sobre obtención, etc. Desde mi punto de vista esto es un error, puesto que la formación del analista ha de comenzar por la adquisición de conocimientos generales que le permitan manejarse en el ámbito de la inteligencia.

De poco sirve conocer herramientas analíticas si no se domina la terminología propia de esta disciplina. Podemos terminar hablando idiomas diferentes si no están perfectamente entendidos los conceptos más elementales y si no hay unidad doctrinal sobre sus acepciones.

En este sentido, no sirve empezar la casa por el tejado: hay que cimentar bien para que la construcción sea sólida y resistente. Lo mismo sucede con la instrucción del analista de última generación: por muy ansioso que esté por aprender metodología de análisis y aplicarla, necesita familiarizarse con el lenguaje propio de la inteligencia. La especialización no está reñida con el conocimiento de todo cuanto tenga relación con el oficio al que vamos a dedicarnos.

Ciclo de inteligencia, seguridad, inteligencia económica, inteligencia competitiva o empresarial, ciber-inteligencia, contrainteligencia, comunidad de inteligencia, inteligencia básica, actual y estimativa, inteligencia táctica, operacional y estratégica, prospectiva, cultura de inteligencia, riesgos transnacionales, Estrategia de Seguridad Nacional, error de análisis, fuentes de información… Estos son algunos de los conceptos con los que el analista debe estar más que familiarizado para poder iniciarse en el oficio de analista.

Pensamiento crítico

Como ya hemos apuntado, el pensamiento crítico constituye el eje fundamental en torno al cual gira la tarea analítica. Debe constituir una asignatura obligatoria, pues no es posible aplicar de modo riguroso ninguna técnica de análisis si no empleamos este tipo de pensamiento.

Esto no excluye el empleo de otros modos de pensar, más bien al contrario: exige la aplicación del pensamiento creativo, el lógico, el lateral…, porque de eso va el dominio del pensamiento crítico como actitud y como procedimiento, de estar abierto a todo lo que implique introducir nuevas perspectivas para poder acercarnos a la realidad con el menor margen de error posible.

Evitar el etnocentrismo es más sencillo si aplicamos el pensamiento crítico al análisis y si practicamos la difícil tarea de ponernos en el lugar del otro. Cuestionarnos nuestro propio pensamiento es la mejor manera de evitar los sinsabores de un análisis *post mortem*.

Pero el pensamiento crítico hay que entrenarlo y ejercitarlo continuamente, con casos reales y con supuestos de escuela, para convertirlo en una manera natural de afrontar los problemas.

Planificación y estrategia

La inteligencia no se produce porque sí, sino que tiene como finalidad facilitar la toma de decisiones con vistas al cumplimiento de un plan. Por este motivo, tanto en los servicios de inteligencia como en las organizaciones donde existen unidades de inteligencia es preciso que los departamentos de análisis estén próximos a la dirección, para que aquellos trabajen de acuerdo con los objetivos marcados por la institución, y en paralelo y al servicio de la estrategia diseñada para alcanzarlos.

De acuerdo con esos fines que quieren alcanzarse, los analistas deben establecer su propio plan, de forma que los requerimientos de información y las necesidades de inteligencia vayan encaminados a la elaboración de inteligencia de utilidad. También se hace preciso establecer medidas de seguridad que permitan proteger lo más valioso que tienen las organizaciones, esto es, la información.

Procedimientos de adquisición de información

Hasta hace unos años el analista mantenía una postura pasiva en tanto que permanecía a la espera de la llegada de la información para convertirla en inteligencia. El analista de última generación, inquieto, comprometido y proactivo, está obligado a mantenerse informado y contar con los datos actualizados y contrastados, a pesar de que es consciente de que nunca dispondrá de toda la información disponible debido a las características de los entornos VUCA en los que vivimos.

El problema al que se enfrentan hoy día los analistas no es de falta de información, sino de exceso (*infoxicación*). Resulta complejo determinar cuáles son los datos relevantes entre tanto ruido informativo, así como obtener aquellos que sean reales y que arrojen una imagen completa de la situación objeto de estudio.

Disponer de una guía actualizada de fuentes es una herramienta de trabajo útil y necesaria. Dicho listado no debe ceñirse a las fuentes tradicionales para obtener información, sino que se impone la creatividad para imaginar otras posibles vías para conseguir los mejores ingredientes para elaborar la mejor inteligencia.

Ceñirse a las fuentes cibernéticas puede limitar las posibilidades de obtención, si bien es preciso dominar los procedimientos OSINT, SOCMINT y todo lo que tenga que ver con el *ciberhumint*.[13] La obtención HUMINT de siempre no puede dejarse de lado; los metadatos no pueden desplazar a los datos más aparentes o próximos.

13 Gestión de fuentes humanas en la red.

Vivimos en la era de los *big data*, pero, como afirma el profesor Carlos Pobre[14], no debemos obviar los *small data*, porque es en ellos donde muchas veces reside la información de mayor valor.

Metodología de análisis

De un analista se espera que sepa analizar, por eso es vital que se forme en técnicas de análisis de inteligencia. Debe conocerlas todas o casi todas para determinar sus capacidades y sus limitaciones; solo así sabrá cuál es la mejor para cada caso concreto, con independencia de que se sienta más cómodo con unas o con otras. No debe reducirse al aprendizaje de las técnicas estructuradas porque limita las posibilidades analíticas.

La formación en este campo ha de ser progresiva e ir de las más sencillas a las más complejas hasta llegar a las propias de la prospectiva. Nociones en estadística y en el análisis bayesiano pueden ser de gran ayuda para el profesional de la inteligencia.

Tecnología e informática aplicada al análisis

La práctica totalidad de las técnicas de análisis de Inteligencia tienen asociadas herramientas informáticas que permiten ponerla en práctica de una manera más sencilla y automatizada. Ello no justifica que sea suficiente manejar la aplicación para poder emplear la técnica; primero hay que conocer los procedimientos de análisis, sus pasos –con sus dificultades y su rutina–, su potencial, sus restricciones…

El analista no precisa conocer los algoritmos que maneja el ordenador, pero sí por qué los hace y qué finalidad persiguen. En este punto, es oportuno señalar la estrecha relación que existe entre el análisis de Inteligencia y las matemáticas, sobre todo la estadística, fundamentalmente si consideramos las técnicas cuantitativas.

Comunicación

Planificar la producción de inteligencia, obtener información oportuna y pertinente y elaborar inteligencia útil para el cliente es vital, pero no suficiente. Recordemos que la misión del analista es poner a disposición del destinatario la inteligencia necesaria para que este pueda tomar decisiones con la menor incertidumbre posible.

De ahí es de donde se deriva la trascendencia de la comunicación de la inteligencia. Saber comunicar, ya sea de forma escrita, ya sea de forma oral, es una cualidad innata o adquirida que el analista ha de tener. Pero, como todo lo demás, puede aprenderse. Para ello existen cursos y talleres donde se practican la expresión verbal

14 Carlos Pobre, sesión formativa en el Posgrado de Inteligencia Económica y Seguridad de MasConsulting ICADE. Madrid, mayo de 2018.

y no verbal –sean o no específicos para transmitir inteligencia–, en los que el analista puede participar para mejorar, si lo precisa, en este aspecto.

Metodologías ágiles / gestión del tiempo / trabajo en equipo y gestión de equipos
Los procesos de digitalización en las organizaciones exigen un cambio de mentalidades y de actitudes que deben gestionarse mediante las metodologías ágiles, según las cuales la adaptación a los rápidos cambios y a las particularidades del producto está garantizada.

Íntimamente relacionado con lo anterior está la necesaria buena gestión del tiempo y de sus ladrones para optimizarlo y poder destinarlo a lo que de verdad importa: la producción de inteligencia. Atender lo urgente no debe impedir afrontar lo importante.

Si trabajar en equipo para elaborar inteligencia es vital en la actualidad, el analista de última generación está obligado a formarse en esta materia. Saber que está sujeto a los principios de la comunicación, la complementariedad, la confianza, la cooperación y el compromiso (las cinco "C" del trabajo en equipo) le servirá para desenvolverse mejor en un mundo donde se impone el desempeño colaborativo.

Deontología profesional
Todos los conocimientos teóricos y prácticos adquiridos por el analista serán inútiles si no los complementa con una buena formación en los principios éticos que deben inspirar su quehacer diario. Obtener información con sometimiento a la ley y analizarla con rigor moral imprimirán al trabajo final una cualidad extra, intangible pero presente, y aportará al analista la tranquilidad del trabajo hecho con todas las garantías que la deontología del profesional de la inteligencia exige.

De este aspecto es cierto que no se ocupan mucho los programas de formación, pero existe literatura al respecto que puede ayudarnos a comprender que es posible desempeñar la tarea con respeto a los derechos de los demás.

A esta lista abierta de materias habría que añadir cuántas se le ocurran al analista (creatividad) y puedan facilitarle su labor o contribuir a mejorar su producto final.

En cualquier caso, tanto los analistas como las organizaciones han de mentalizarse que la formación de aquellos no es un mero gasto, sino una inversión cuyo retorno es tangible a corto, medio y largo plazo: conduce a la consecución de los objetivos

estratégicos y, en su caso, a una mejora en la cuenta de resultados, todo ello como efecto de la oportuna toma de decisiones sobre la base de una inteligencia útil. Se trata de cambiar el paradigma de las unidades de inteligencia como "centros de coste" por el de "centros de beneficio" que aportan valor añadido y que contribuyen a un mejor desarrollo de las empresas, instituciones, corporaciones, *think tanks*…

Conclusión

En pleno siglo XXI es preciso contar con analistas preparados para afrontar los retos que los nuevos tiempos presentan. No sirve con "saber analizar"; es preciso ser capaz de obtener con buenas prácticas información útil entre todos los datos disponibles, elaborar inteligencia y ponerla a disposición del cliente casi con inmediatez.

Para desempeñar esta profesión y avanzar, la formación constituye un pilar fundamental para el analista. Este ha de estar al día en multitud de materias y disciplinas, y exigirse a sí mismo rigor metodológico y ético.

En resumen, el analista de inteligencia debe ser un todoterreno de última generación, con todos los extras posibles, con actitud para adaptarse a los cambios, con capacidad para asumir las actualizaciones disponibles y con una proactividad que le permita anticiparse a los acontecimientos en todos los sentidos.

Bibliografía

D. Pérez, *Nuevos líderes VUCA ¿conoces el futuro que nos espera?* www.pagepersonnel.es

P. Canal (2015), *Educación en entornos VUCA*, www.iebschool.com

Diversos autores (2018), *Postgrado de Inteligencia Económica y Seguridad*, 7ª edición, Universidad Pontificia Comillas / MasConsulting

M. Jiménez (2018), *Metodologías Ágiles*, Instituto Hune

J. Goldman (2005), *Ethics of spying: a reader for the intelligence professional*

C. Sánchez Capdequí (2017), *El ethos creativo: debates y diagnósticos sobre el nuevo imperativo moderno*, Política y Seguridad, 54 (3), págs. 621-640.

Ana Páramo es licenciada en Derecho por la Universidad Pontificia Comillas ICAI-ICADE (E-1). Ha participado en el Curso de Experto en Análisis de Inteligencia por el Instituto de Ciencias Forenses y de la Seguridad de la Universidad Autónoma de Madrid, así como en el Postgrado en Inteligencia Económica y Seguridad por Masconsulting/ ICADE. Tiene experiencia en el sector de seguros de crédito a la exportación y en el ámbito del análisis de inteligencia orientado a la seguridad, tanto como analista como formadora de analistas.

Capítulo II: La inteligencia prospectiva

Perspectivas para el análisis de tecnología orientado al futuro

Ahti Salo

Introducción

Los avances tecnológicos y las innovaciones basadas en la tecnología son fuentes duraderas de ventaja competitiva en el sector privado, civil y militar. También requieren decisiones en torno a la planificación y la gestión de la investigación y el desarrollo tecnológico (IDT), que pueden necesitar recursos financieros importantes y el desarrollo de competencias especializadas, y llevan mucho tiempo. Algunos ejemplos de estas decisiones serían la formulación e implementación de las prioridades de los programas IDT, la elección entre tecnologías en competición directa, la explotación de los derechos de propiedad intelectual y la gestión de los ciclos de vida de plataformas consecutivas de productos. Muchas de estas decisiones son sumamente estratégicas y por lo tanto deben estar bien fundamentadas.

El ámbito del análisis de tecnología orientado al futuro (ATOF), como término general, abarca una serie de métodos sistemáticos que pueden utilizarse para producir análisis fundamentados sobre la evolución de futuras tecnologías, o sobre las consecuencias de su utilización en ciertos productos y servicios. Dependiendo de sus requisitos específicos, estos análisis pueden concentrarse en un ámbito reducido o, al contrario, abordar ampliamente una serie de impactos sociales, tecnológicos, económicos, medioambientales, y políticos (STEMP), concentrándose tanto en efectos deseados como no-deseados (donde la "deseabilidad" se mide según los intereses de las partes implicadas). También se puede distinguir entre impactos directos e indirectos, dependiendo de si estos efectos pueden atribuirse directamente al uso de la tecnología (por ejemplo, la emisión de sustancias al medio ambiente) o si, por el contrario, son el resultado de una cadena de acontecimientos más compleja que pueda incluir comportamientos humanos impredecibles (por ejemplo, la falta de precaución conduciendo debido a una confianza desmesurada en los dispositivos de seguridad automáticos).

Más específicamente, en el ATOF existen varios enfoques complementarios para justificar decisiones y políticas:

- Los *pronósticos de tecnología* buscan generar predicciones sobre las capacidades tecnológicas específicas que existirán en un momento dado del futuro para proporcionar máquinas, técnicas y servicios útiles.
- Las *evaluaciones de tecnología* (ET) se centran en la descripción y evaluación de los impactos del uso de la tecnología. Las ET se utilizan frecuentemente en la toma de decisiones regulatorias, por ejemplo, el Parlamento Europeo tiene un panel de Evaluación de Opciones Científicas y Tecnológicas[15] (STOA por sus siglas en inglés) que se dedica a analizar aquellos desarrollos científicos y tecnológicos que puedan requerir nuevas medidas regulatorias o que merezcan la atención de los responsables políticos.
- La planificación tecnológica[16] (PT) se suele utilizar en un contexto de gestión tecnológica o de ingeniería de sistemas. Puede utilizarse para apoyar las primeras fases del desarrollo de un producto, ya que explica cómo distintas combinaciones de opciones individuales tecnológicas pueden ser utilizadas para construir nuevas capacidades. Puede ser útil, por ejemplo, a la hora de planificar las distintas generaciones de un mismo producto, a través de las cuales se reemplazarían antiguas soluciones tecnológicas por otras más actualizadas.
- La *inteligencia tecnológica* se refiere a las actividades que ayudan a las empresas a identificar las oportunidades y amenazas tecnológicas que podrían afectar su crecimiento y su supervivencia. Esta disciplina se centra en recoger, diseminar e interpretar información tecnológica que luego se utiliza en la planificación estratégica y la toma de decisiones. De este modo, la inteligencia tecnológica podría considerarse una variante de la inteligencia estratégica y competitiva.
- La *previsión*, un enfoque general del ATOF, consiste en una serie de actividades que se llevan a cabo para, por ejemplo, entender el pasado reciente; brindar conocimientos a través del escrutinio sistemático de horizontes aplicado a tendencias y eventos potencialmente relevantes; generar visiones de futuro preferidas; y desarrollar estrategias a través de las cuales se puedan alcanzar estas visiones. De esta forma, la previsión, el pronóstico, la planificación estratégica y la gestión del cambio tienen mucho en común.

15 Ver http://www.europarl.europa.eu/stoa/en/home/highlights. Para más información sobre las actividades de planificación tecnológica de los miembros de la Red de Evaluación de Tecnología Parlamentaria Europea, ver https://eptanetwork.org/.

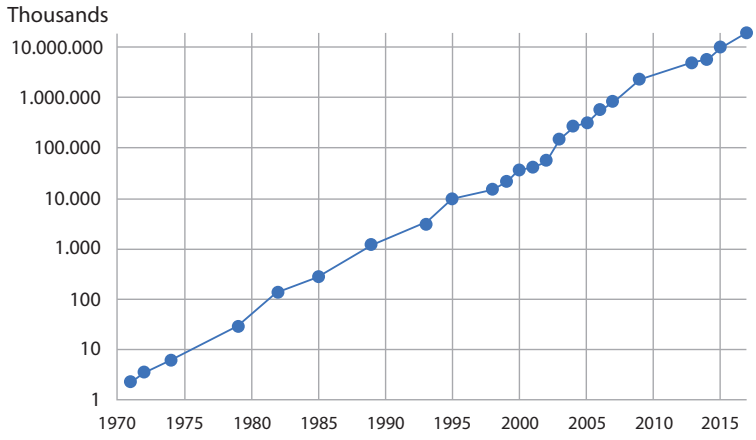
16 Zhang *et al*. (2016).

Metodológicamente, estos enfoques no son completamente diferentes; más bien, muchos de ellos se basan en métodos parecidos, cuyas principales características resumimos más adelante.[17] Desde una perspectiva histórica, el pronóstico tecnológico es probablemente el más antiguo, ya que se le reconoció como una disciplina propia en los años 50, cuando la Unión Soviética y EEUU competían por la supremacía militar y la conquista del espacio. En esa época, el desarrollo de sistemas intercontinentales y espaciales requería una cantidad enorme de recursos y tardó muchos años en completarse. Esto creó la necesidad de desarrollar el pronóstico tecnológico de forma sistemática, basándose en el uso de técnicas estadísticas y en consultas con expertos.

Métodos estadísticos

El uso de métodos estadísticos en el ATOF se basa en la idea de que la evolución de las características del rendimiento tecnológico y la divulgación de nuevas innovaciones suelen exhibir patrones sistemáticos que pueden utilizarse para generar predicciones. Un ejemplo muy celebrado de dicho patrón sería la Ley de Moore, que expresa que el número de transistores en circuitos integrados se duplica cada dos años, demostrando así su crecimiento exponencial. Como se demuestra en la Figura 1, esta Ley lleva demostrando su validez más de 45 años.

Figura 1. Transistores por microprocesador



Fuente: https://ourworldindata.org/technological-progress

17 Para un resumen más exhaustivo, ver el artículo TFAWG (2004).

Las curvas de crecimiento –por ejemplo, aquellas que encontramos en los modelos de Bass, Gompertz y Fischer-Pry– pueden ser útiles para predecir la difusión de nuevas tecnologías o la penetración de mercado de ciertas innovaciones. Estos modelos son similares al basarse los tres en funciones de logística parametrizadas; sin embargo, también existen diferencias, por ejemplo en la medida en la que los distintos modelos suponen que: (i) las tecnologías son completamente nuevas y no sustitutos mejorados de tecnologías anteriores (por ejemplo, las cámaras digitales sustituyendo a las que usan celuloide); (ii) la adopción de tecnologías se debe a los efectos de red (por ejemplo, el nivel de utilidad de las redes sociales para un usuario individual depende del número total de usuarios que las utilicen); y que (iii) existen niveles de saturación que impondrán límites en la difusión de estas tecnologías (por ejemplo, debido a su dependencia de infraestructuras relacionadas con ellas). Dado que las curvas de crecimiento sólo utilizan unos pocos parámetros numéricos, la calidad de sus predicciones depende mucho de la precisión de estas estimaciones, particularmente de aquellas relacionadas con el nivel de saturación y los distintos efectos de red que puedan impulsar su difusión, que pueden ser críticas. En la práctica, uno de los desafíos de las curvas de crecimiento es que sus predicciones sobre la difusión suelen ser de mayor interés cuando todavía existen pocos datos sobre las nuevas tecnologías e innovaciones que investigan, lo cual dificulta la elaboración de estimaciones precisas de los parámetros del modelo e incluso la elección entre modelos de crecimiento alternativos.

Hasta cierto punto, los métodos de análisis de correlación abordan este desafío ya que, en vez de intentar analizar directamente la breve historia de la innovación y de las nuevas tecnologías, buscan generar nuevos conocimientos a base de identificar y explorar correlaciones entre desarrollos tecnológicos similares. Por ejemplo, en el caso de los dispositivos electrónicos de consumo, innovaciones tales como el microondas o el televisor de alta definición se han popularizado a través patrones de difusión comparables entre sí. De este modo, cuando emerge una nueva innovación, estos patrones pueden utilizarse para: (i) identificar cuál de las innovaciones anteriores se parece más a la innovación en cuestión; y (ii) formular predicciones sobre la difusión de esta nueva innovación basándose en datos sobre los patrones de difusión de las anteriores innovaciones.

La minería de tecnologías (conocida como *tech mining*[18]) se parece a los métodos estadísticos en la importancia que otorga al análisis de datos empíricos y se basa notablemente en aplicar métodos de minería de textos (y, más generalmente, de

18 Porter & Cunningham (2004).

aprendizaje automático y de ciencia de datos) a documentos técnicos. Una de sus variantes es la minería de patentes, que utiliza la minería de textos para analizar documentos de patentes. Los resultados de los análisis basados en la minería de tecnologías se pueden sintetizar de varias maneras, sea preservando la riqueza cualitativa de las fuentes originales en distintos contextos, o generando información estadística cuantitativa que subraye las tendencias observadas en el desarrollo de nuevas tecnologías (lo cual puede ser útil para determinar políticas de desarrollo de la investigación y la tecnología). En los últimos años han crecido tanto la usabilidad como la utilidad de la minería tecnológica, gracias a la mayor disponibilidad de documentos técnicos en internet y en distintas bases de datos, así como a los avances en el aprendizaje automático y la disponibilidad de instrumentos mejorados de minería de textos. La minería tecnológica puede ser especialmente útil para obtener inteligencia tecnológica.

En principio, los métodos de simulación y otros métodos cuantitativos pueden ayudar a generar previsiones sobre la emergencia de nuevas capacidades tecnológicas y la adopción de nuevas innovaciones, particularmente en aquellas situaciones en las que a estos procesos dinámicos los rigen interrelaciones entre variables tecnológicas, económicas y sociales que se puedan capturar en un modelo plausible (por ejemplo, la difusión de los coches eléctricos depende de la disponibilidad de estaciones de carga, inversiones en el DTI, posibles subvenciones y el precio de modos de transporte alternativos). Sin embargo, suele ser complicado utilizar estos modelos para hacer pronósticos fiables, dado que el desarrollo de nuevas tecnologías suele venir acompañado de mucha incertidumbre y que, por definición, las innovaciones son nuevas. Sin embargo, siguen siendo útiles para determinar nexos causales entre variables, subrayando así las combinaciones de variables que más afectarán a los futuros desarrollos tecnológicos.

Juicio de expertos

Los métodos de juicio de expertos siempre han desempeñado un papel central en el ATOF debido a la incertidumbre que existe alrededor del desarrollo tecnológico y sus impactos. El más conocido es Delphi[19], un método sistemático para generar pronósticos a través de consultas estructuradas a paneles de expertos. En cada ronda, se invita a los miembros de estos paneles a proporcionar: (i) previsiones sobre sucesos concretos que puedan ocurrir en el futuro (por ejemplo, "una nave espacial

19 Dalkey & Helmer (1963). Para un resumen reciente de distintos métodos de juicio de expertos, ver Dias *et al*. (2018).

con pasajeros humanos aterrizará con éxito en Marte"); y (ii) los argumentos en los cuales basan sus respuestas. Estas respuestas se sintetizan y se envían de vuelta a los miembros del panel, acompañadas de información sobre la dispersión estadística de sus respuestas (por ejemplo, la media de los pronósticos, así como los cuantiles superiores e inferiores) y un resumen de los argumentos presentados para defender estos pronósticos. Los miembros del panel pueden, si lo desean, revisar sus pronósticos iniciales basándose en esta información. El número de rondas de consultas que se llevan a cabo depende de lo rápido que se estabilicen estos pronósticos y de cuánto aprendan los expertos en cada ronda. Los resultados del Delphi consisten en los pronósticos, sus propiedades estadísticas, y los argumentos utilizados por los expertos para fundamentar sus pronósticos.

Delphi puede generar una gran riqueza de información basada en declaraciones explícitas sobre eventos futuros y los plazos en los que se espera que ocurran estos eventos. Hoy en día, con el apoyo de herramientas modernas, estos procesos pueden llevarse a cabo de forma eficaz a través de Internet. Por otro lado, elegir los temas adecuados para analizar puede ser complicado. En particular, si no se identifican sucesos importantes en las primeras rondas, es difícil corregir estas omisiones en rondas posteriores. Otra limitación de Delphi es que las preguntas se estructuran alrededor de un tema en particular, complicando así la identificación de interdependencias entre los distintos temas.

Previsión

En términos de su alcance, la previsión tecnológica constituye el enfoque más amplio del ATOF, como demuestra el hecho de que se pueda definir como "un proceso que intenta, de forma sistemática, observar el futuro a largo plazo de la ciencia, la tecnología, la economía y la sociedad, con el objetivo de identificar los ámbitos clave de investigación estratégica y las tecnologías emergentes que más probablemente brindarán los mayores beneficios económicos y sociales a la humanidad".[20] De acuerdo con este objetivo tan amplio, la previsión tecnológica como método ha buscado ayudar a definir políticas de DIT, lo cual constituía uno de los principales objetivos de los estudios de previsión a gran escala que se llevaron a cabo en los años 90 y a principios de los 2000.[21]

20 Irvine y Martin (1984).
21 Salo & Cuhls (2003).

A lo largo de los años, la previsión como disciplina se ha desarrollado para abarcar también otros objetivos, por ejemplo, la mejora de programas de aprendizaje organizacional, el desarrollo de redes de colaboración, y la mejora de capacidades para el "escrutinio de horizontes" y el *sense-making*. Al abarcar estos objetivos, la previsión se acerca bastante a la planificación estratégica, especialmente con respecto al análisis de escenarios[22], un método muy popular que se utiliza para apoyar ciertas decisiones estratégicas en un contexto de gran incertidumbre. Estas decisiones se pueden fundamentar en un análisis sistemático de fuentes clave de incertidumbre; una descripción de cómo podrían materializase estos factores; y la elaboración de distintos escenarios basados en distintas combinaciones de los mismos, en caso de que se materializasen, para construir una visión equilibrada de posibles fuentes de incertidumbre relevantes y generar así conocimientos para los gestores de la organización.

Tendencias y desarrollos

La previsión como método suele criticarse utilizando ejemplos relevantes e incluso entretenidos de ocasiones en las cuales ha funcionado mal.[23] Pero estos ejemplos no reducen el nivel de responsabilidad y de presión que supone tener que tomar ciertas decisiones inmediatas o a corto plazo. De este modo, sigue siendo importante reflexionar sobre cuál es la mejor manera de fundamentarlas, si a través de acceder a mejores fuentes de datos o de sintetizar información obtenida de distintas fuentes utilizando metodologías adecuadas.

Dentro de este contexto, se pueden observar ciertas tendencias que afectan el uso del ATOF.[24]

- Las innovaciones suelen ser cada vez más sistémicas, ya que cada vez existen más y mayores interfaces entre distintos sistemas habilitadores. Esto significa que no se pueden analizar por sí solas, sino en su contexto sistémico.
- Dado que desarrollar un alto nivel de especialización y conocimiento suele llevar mucho tiempo, y que este conocimiento suele estar limitado por el

22 Ver, por ejemplo, Bunn y Salo (1993).

23 Por ejemplo, Ken Olsen, fundador del *Digital Equipment Corporation*, 1977: "No existe razón alguna por la cual alguien querría tener un ordenador en su casa". Ver https://www.pcworld.com/article/155984/worst_tech_predictions.html

24 Ver, por ejemplo, las presentaciones que se hicieron en la conferencia *FTA2018 - Future in the Making*, https://ec.europa.eu/jrc/en/fta2018 , o las actividades presentadas por la *European Foresight Platform*, http://www.foresight-platform.eu/ .

contexto en el cual se desarrolla, el ATOF debería realizarse a través de actividades colaborativas que incluyan a expertos de orígenes muy diversos, mitigando así los peligros del pensamiento de grupo.[25]

- Los beneficios que aporta el proceso –tales como el desarrollo de redes profesionales, el que los participantes tomen mayor conciencia de las distintas fuentes de información que existen, e incluso la humildad que aporta entender lo poco que sabemos sobre futuros desarrollos tecnológicos– pueden ser tan importantes (si no más) como los usos más instrumentales del AFTO en apoyar la toma de decisiones.

Estas tendencias han afectado de muchas maneras los procesos de ATOF participativos en los que ha tomado parte el autor de este artículo. Estos incluyen la gestión del primer estudio de AT llevado a cabo para el parlamento finlandés y el diseño y ejecución del primer ejercicio de previsión nacional,[26] así como actividades más especializadas para evaluar ciertos desarrollos tecnológicos y sociales que guíen el ejercicio de priorización en programas de DIT.[27] Una lección clave de estos procesos ha sido que es importante aclarar qué expectativas acerca del ATOF tienen las distintas organizaciones e instituciones, para así tomar las decisiones metodológicas adecuadas que permitan que estas expectativas se cumplan de forma eficaz y económica.

Bibliografía

Bunn, D. y Salo, A. (1993). Forecasting with Scenarios. *European Journal of Operational Research 68*, 291–303.

Dalkey, N. y Helmer, O. (1963). An Experimental Application of the Delphi Method to the use of experts. *Management Science 9/3*, 458–467.

Dias, L.C., Morton, A. y Quigley, J. (eds.) (2018). *Elicitation: The Science and Art of Structuring*, International Series in Operations Research & Management Science Vol. 261, Springer.

Irvine, J. y Martin, B. (1984). *Foresight in Science: Picking the Winners*. Pinter.

Könnölä, T., Brummer, V. y Salo, A. (2007). Diversity in Foresight: Insights from the Fostering of Innovation Ideas. *Technological Forecasting and Social Change* 74/5, 608–626.

25 Ver Könnölä *et al.* (2007).
26 Salo *et al.* (2009).
27 Vilkkumaa *et al.* (2014).

Porter, A.L. y Cunningham, S.W. (2004). *Tech Mining: Exploiting New Technologies for Competitive Advantage.* Wiley.

Salo, A. y Kuusi, O. (2001). Developments in Parliamentary Technology Assessment in Finland. *Science and Public Policy* 28/6, 453–464.

Salo, A., Brummer, V. y Könnölä, T. (2009). Axes of Balance in Foresight–Reflections from FinnSight 2015. *Technology Analysis & Strategic Management* 21/8, 989–1003.

Salo, A. y Cuhls, C. (2003). Technology Foresight—Past and Future. *Journal of Forecasting* 22/2–3, 79–82.

Technological Futures Analysis Methods Working Group (2004).Technology Futures Analysis: Toward Integration of the Field and New Methods. *Technological Forecasting & Social Change*, 71/3, 287–303.

Vilkkumaa, E., Salo, A. y Liesiö, J. (2014). Multicriteria Portfolio Modelling for the Development of Shared Action Agendas. *Group Decision and Negotiation* 23/1, 49–70.

Zhang, Y., Robinson, D.K.R., Porter, A.L., Donghua, Z., Guangqua, Z. y Lu, J. (2016). Technology roadmapping for competitive technical intelligence. *Technological Forecasting & Social Change* 110, 175–186.

Ahti Salo es profesor en el Departamento de Matemáticas y Análisis de Sistemas de la Universidad de Aalto. Ha trabajado extensamente en el desarrollo de métodos de toma de decisiones analíticos, la gestión de la innovación, la gestión de riesgos, la previsión tecnológica y el análisis de eficiencia. Ha publicado extensamente en revistas académicas internacionales de referencia (incluyendo Management Science y Operations Research) y ha recibido premios de investigación de la Decision Analysis Society del Institute for Operations Research and the Management Sciences (INFORMS). Es miembro del consejo editorial de varias revistas académicas de referencia. Ha dirigido una serie de proyectos de investigación básicos y aplicados, financiados por empresas líderes del sector industrial, federaciones industriales, y organismos de financiación. Ha sido catedrático invitado en la London Business School, la Université Paris-Dauphine y la Universidad de Viena. Ha sido presidente de la Sociedad Finlandesa de Investigación Operativa (FORS) durante dos trimestres bienales. En 2010-2011, ejerció como Representante Europeo y del Medio Oriente en el Comité de Actividades Internacionales de INFORMS. En 2010-2016, fue miembro del jurado del Premio de Tesis Doctoral EDDA de la Asociación Europea de Sociedades de Investigación Operativa (EURO), y ejerció como presidente del jurado en 2016. Ha sido miembro del Consejo de la Asociación de Miembros e Investigadores Parlamentarios (Tutkas) desde 1999.

Inteligencia prospectiva: construcción de escenarios

Jessica Cohen y José María Blanco

Inteligencia prospectiva y construcción de escenarios

El tardío desarrollo de la prospectiva en el ámbito español, tanto en el espacio público como el privado, ha influido en su escasa visibilidad, aun estando este concepto estrechamente ligado a la estrategia. Así, pese a la reciente aparición de documentos de estrategia nacional (de seguridad nacional, internacionalización de la empresa o acción exterior, entre otros) la prospectiva es aún un aspecto residual.

Desde nuestra visión y experiencia, hemos optado por la definición de prospectiva propuesta por Gastón Berger (1964, 1967, 1991), según el cual, se trata de "*la ciencia que estudia el futuro para comprenderlo y poder influir en él*". En este punto es donde pueden surgir reticencias respecto a la imposibilidad de estudiar algo aún no acontecido, en parte por la confusión de esta con la predicción, la proyección o la adivinación. Nada más lejos de la realidad, aunque el futuro aún no se haya materializado, en el presente tenemos un amplio conocimiento sobre tendencias y sucesos que acontecerán: el envejecimiento de la población es visible; disponemos de agendas que contemplan cuando tendrán lugar las próximas elecciones; las estimaciones sobre cantidades de recursos disponibles cada vez son más certeras; etc.

Ejemplos como los descritos guardan relación con la visión que Peter Drucker, experto en gestión empresarial, tiene sobre la planificación a largo plazo, al asegurar que esta "*no es pensar en decisiones futuras, sino en el futuro de las decisiones presentes*". Por esta facultad la prospectiva es también conocida como "*inteligencia prospectiva*" o "*prospectiva estratégica*", estando ambos conceptos ligados íntimamente con la dirección estratégica.

La utilización de modelos prospectivos permite diseñar escenarios de futuro. El término "*escenario*" como herramienta de apoyo a la investigación de futuros y el análisis de políticas fue acuñado por Herman Kahn en la década de los 50, definiéndolo en 1967 como "*descripción narrativa del futuro que centra la atención en los procesos causales y los puntos de decisión*".

Es especialmente reseñable el valor estratégico de la prospectiva y, en particular, de la elaboración de escenarios. La dirección estratégica (misión, visión y valores) debe contar con una gestión del conocimiento (datos e información estructurados de interés para el organismo en cuestión –público o privado–), un análisis de inteligencia (actores, variables, riesgos y amenazas; es decir, datos e información evaluados, integrados, puestos en contexto y orientados a la toma de decisiones) y prospectiva (visión a largo plazo basada en el estudio del futuro).

Figura 1. Prospectiva y estrategia

Fuente: elaboración propia.

El estudio sobre el futuro puede ser un punto de partida para actualizar la visión de las organizaciones, entendida como la forma en que éstas pretenden lograr su misión. Misión, que debe ser adaptable al entorno y éste, a su vez, únicamente es susceptible de análisis a través de modelos holísticos. De esta forma la visión de las organizaciones debe evolucionar de acuerdo con la sociedad en la que éstas prestan servicio y sus valores. En este sentido, la prospectiva puede contribuir a mejorar el servicio al ciudadano o al cliente (Blanco, 2016).

En opinión de los presentes autores, el principal valor de la prospectiva deriva de su uso a modo de sistema de seguimiento y evaluación continua, más que de simple aporte estático de información (Blanco y Jaime, 2014). En este proceso, la detección de tendencias y la identificación de indicadores propician el hallazgo de señales débiles que pueden utilizarse a modo de alertas tempranas, advirtiendo así a los usuarios sobre la existencia o evolución de tendencias, riesgos o eventos, en principio no detectados. Por este motivo no se debe considerar la prospectiva como un producto (informe), sino como un sistema ajustable y evaluable continuamente. Un sistema de gestión estratégico (Blanco y Cohen, 2018).

Como añadido, huelga decir que el simple hecho de pensar en futuro procura la adopción de un patrón cognitivo que facilita pensar en problemas no actuales, potenciando la creatividad y el pensamiento crítico de los analistas (Blanco, 2016). El pensamiento y reflexión sobre el futuro es susceptible de introducir, por sí mismo, causalidad en el sistema. El conocimiento sobre la probabilidad de un futuro más o menos negativo contribuye a adaptar las estrategias de los actores del sistema, pudiendo dar lugar al fenómeno de las "profecías que se autoanulan". De manera similar, se pueden dar las "profecías que se autocumplen", escenarios futuros que, al explicitarse, presentan los elementos necesarios para su consecución (Blanco y Cohen, 2018).

La prospectiva permite romper la concepción de linealidad del tiempo. Esta característica que *a priori* puede resultar difícil de entender se basa en que el pasado y el presente condicionan el futuro, pero no lo determinan en su totalidad. Las expectativas y deseos sobre el futuro afectan también a las decisiones que se adoptan en el presente. De este modo, es posible entender la existencia de una multiplicidad de futuros en cada momento, cuya caducidad o perpetuidad deviene influida por los hechos y las decisiones que se van adoptando en cada momento.

Por este mismo motivo, el uso de escenarios muestra una mayor utilidad cuando se tienen en consideración todos en su conjunto. Sin duda, es posible centrarse en el más favorable y, desde el presente, iniciar la toma de decisiones que lo procuren; sin embargo, considerar el conjunto de escenarios también permite advertir eventualidades no deseadas ante las que estar preparados, aumentado así la resiliencia frente a las mismas.

Dentro de sus ventajas se encuentra también la reducción de incertidumbre que su uso aporta, al exigir la asignación de probabilidades a cada suceso esperado. Por

supuesto, el margen de error de este ejercicio será mayor cuanto a más largo plazo se haga el análisis.

Partiendo de las utilidades por las cuales se diseñan escenarios, destacadas en el *Millennium Project*, junto a las reconocidas por los presentes autores en su experiencia, se destaca su beneficio en:

- Identificar, ante una toma de decisiones, aquellas cuestiones que se conocen que se desconocen; las que se conocen que se conocen; y las que no se conocen que se conocen (aunque en este último caso el grado de posible éxito es muy limitado).
- Priorizar las estrategias que pueden ser de mayor relevancia a partir del abanico de escenarios analizado.
- Detectar riesgos y oportunidades, no sólo generados por la evolución del entorno, sino también derivados de eventuales decisiones.
- Acercar el proceso analítico al decisor facilitando la comprensión de lo planteado, mejorando así la toma de decisiones.
- Comprender y reducir la incertidumbre.
- Contribuir al fomento de un pensamiento más estructurado y cercano al futuro.
- Potenciar la creatividad en los analistas, conectando las diferentes informaciones de una forma, inicialmente, no intuitiva o lógica.
- Facilitar el ulterior y continuo seguimiento, en la actualidad potenciado por la posibilidad de hacer consolas digitales, al permitir una rigurosa identificación y clasificación las diferentes tendencias, indicadores, actores y sucesos desde el inicio.

Principales técnicas en prospectiva para la construcción de escenarios

La compañía Shell International Petroleum fue una de las primeras corporaciones privadas en hacer uso de los escenarios para estudiar el futuro, aunque fue General Electric la primera empresa privada en aplicar el método de escenarios. Sin ánimo predictivo, trataban de identificar las tendencias de cambio en la economía y la sociedad de EEUU a finales de la década de los años 60 y comienzos de la década de los 70.

En general, los escenarios se establecen a partir de las variables clave de futuro y la combinación de posibles configuraciones de esas variables. Esta es la aproximación utilizada para el Consejo Nacional de Inteligencia de EEUU o el *World Economic Forum*. "Paradox of Progress" (2017), del Consejo Nacional de Inteligencia de EEUU, propone tras identificar las claves de futuro los siguientes escenarios:

- Islas. El crecimiento estancado de muchas economías incentiva una vuelta al proteccionismo frente a la globalización.
- Órbitas. Grandes centros de poder compiten, tratando de influir en el exterior mientras intentan consolidar la estabilidad interna en sus estados.
- Comunidades. Las mayores expectativas públicas, unidas a la incapacidad de los Estados por cumplir, llevan a que se generen núcleos de poder de carácter local y no estatales.

El *World Economic Forum*, analizando la seguridad en 2030, identifica siete fuerzas de cambio clave:

- La innovación tecnológica.
- El cambio climático y el acceso a los recursos.
- La buena gobernanza, transparencia e imperio de la ley.
- La competencia geoestratégica.
- El cambio demográfico.
- La cohesión social y la confianza.
- Las amenazas híbridas y asimétricas en un mundo hiperconectado.

La combinación de éstas produce tres escenarios, a los que se dota de una denominación (se utiliza habitualmente títulos de efecto) y se ofrece un relato completo sobre las características de cada uno de ellos.

- Ciudades amuralladas. Islas de orden (ciudades) en un mar de desorden (el mundo).
- Regiones fuertes. Un mundo competitivo con diferentes centros de poder.
- Guerra y paz. Un mundo sumido en un gran conflicto que lleva a un replanteamiento del sistema.

Peter Schwarz (1991), propone el siguiente modelo para construir escenarios:

Fase 1. Alcance de la cuestión a estudiar. Se trata de fijar claramente la pregunta a responder, y su alcance geográfico y temporal.

Fase 2. Identificación de los principales *drivers*. Estas variables clave se pueden identificar a través de un *brainstorming* o paneles de expertos. Posteriormente se evalúan según su importancia y grado de incertidumbre. *Environmental scanning* o PESTEL, serían técnicas utilizables a tal fin.

Fase 3. Identificar las incertidumbres críticas. Son aquellos agentes o factores con mayor impacto en el marco temporal marcado, o aquellos que generan mayores incertidumbres. Son aplicables el método MICMAC de análisis de variables de Godet o las técnicas de consulta a expertos.

Fase 4. Matriz de escenarios. Se suelen tomar las incertidumbres críticas por pares, configurando distintos escenarios (una variable en "x" y otra en "y", dando lugar a un cuadrante).

Fase 5. Identificación de indicadores. Se identifican aquellos indicadores que explican cómo se llega a cada escenario. Se elabora una matriz de doble entrada, con indicadores en filas y posibles escenarios en columnas, que permita visualizar la configuración de cada escenario en cada momento, y su evolución cuando se modifican los indicadores.

La denominada "Caja de Godet", nacida en Francia en el año 1993 y acuñada así por su creador es otra de las aproximaciones tradicionales, cuya metodología surgió en el Laboratorio de Investigación de Estudios Prospectivos de París. El modelo desarrollado por Godet (2011) ofrece un conjunto de metodologías que permiten llegar a la construcción de escenarios futuros. Para ello, partiendo de un análisis previo sobre la materia u organización a estudiar, identifica las variables clave (para lo que utiliza la metodología MICMAC), analiza el juego de actores (metodología MACTOR), explora el campo de los futuros posibles (a cuyos efectos utiliza el análisis morfológico, las encuestas Delphi, el ábaco de Régnier o el método de impactos cruzados) y evalúa las opciones estratégicas (árboles de pertinencia y el método Multipol). La limitación de estas aproximaciones reside en el sesgo de arrastrar la experiencia y los conocimientos previos hacia un futuro en el que no necesariamente han de existir.

The Millennium Project fue fundado en el año 1996 como un *think tank* independiente formado por futurólogos, investigadores, decisores que trabajan en organizaciones internacionales, gobiernos, ONG y universidades. Sus trabajos prospectivos son publicados bajo la denominación de *State of the Future*. Destaca, a nivel metodológico su colección *Futures Research Methodology*.

Con el objetivo de sortear alguna de las limitaciones que los citados modelos planteaban, desde el año 2013, los presentes autores trabajaron en el Centro de Análisis y Prospectiva (CAP) de la Guardia Civil, en una aproximación al futuro en dos direcciones: del presente al futuro y del futuro al presente.

- Del presente al futuro

  Se consigue a través del análisis de tendencias. Un ejercicio que resulta sencillo de comprender para el decisor, al percibir elementos del presente en el futuro. Este análisis facilita el seguimiento de los impulsores e indicadores que influyen en cada tendencia y permite la evaluación de sus efectos en seguridad, o ámbito bajo estudio. De este modo, la evaluación y la actualización es continua.

  En este paso, el CAP hace uso de metodologías como el análisis de tendencias, su evaluación, DAFO, PESTEL, revisión bibliográfica, indicadores, análisis del entorno, técnicas de creatividad, paneles de expertos, etc.

- Del futuro al presente

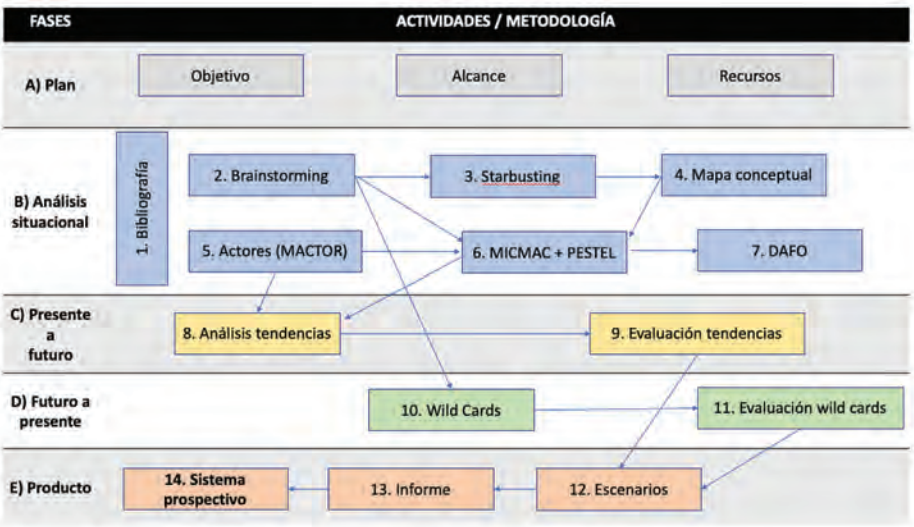
  Se ha pensado con objeto de limitar la proyección de conocimientos y experiencias previas que se permiten en el primer enfoque. El conocimiento y las experiencias previas son un sesgo cognitivo, y el futuro no siempre es una proyección del pasado y del presente.

En este caso, se parte de un futuro imaginario para, bajo un proceso de *backcasting*, llegar a determinar las variables que han conducido al mismo. El seguimiento de las variables procura la detección de señales débiles, pudiendo generar, como se ha indicado anteriormente, alertas tempranas sobre la materialización de ciertas tendencias o eventos.

Son útiles técnicas como *what if?*, *wild Cards* o la creación de distopías o utopías. En esta última propuesta, la que corresponde a la creación de escenarios bajo el criterio de confección de una distopía, o utopía en su caso, es preciso un mayor esfuerzo cognitivo, al forzar a nuestro pensamiento a pensar en situaciones lo más alejadas posibles del mundo conocido.

1. En la Figura 1 se muestra el paso a paso del presente modelo:
   - Planificación del trabajo prospectivo. Se trata de delimitar los parámetros clave del análisis prospectivo a realizar:
   - Entendimiento del requerimiento a satisfacer.
   - Delimitación del objeto de estudio y su alcance
   - Identificación del marco geográfico.
   - Fijación del horizonte temporal.
2. Análisis estructural. El objetivo es disponer de datos e información sobre el fenómeno, las variables involucradas, así como sobre los contextos histórico y geográfico. Se utilizan diversas técnicas entre las que se incluyen a su vez un análisis de variables, actores, facilitadores e inhibidores, así como de posibles riesgos. Todo análisis prospectivo para la construcción de escenarios precisa de análisis de inteligencia.
3. Análisis de presente a futuro. Apoyado en el análisis estructural, se desarrolla un análisis y evaluación de tendencias. Un ejemplo de aplicación sería el informe de *RAND Europe*: "Europe's Societal Challenges. An analysis of global societal trends to 2030 and their impact on the EU".
4. Análisis de futuro a presente. El objetivo es salir de los marcos cognitivos marcados por los conocimientos y experiencias de los analistas, y tratar de anticipar posibles hechos disruptivos. Como técnicas básicas se usan: *backcasting*, *what if*, *wild cards* y distopías/utopías.
5. Construcción de escenarios. Determinación de las claves de futuro del fenómeno a analizar, construyendo escenarios a través del análisis morfológico, y todo ello basado en los *inputs* analíticos de las fases previas.

Figura 2. Elaboración de escenarios



Fuente: elaboración propia.

Juicio crítico sobre las aproximaciones desarrolladas en prospectiva

Las principales críticas que se le puede hacer a la prospectiva enfocada a la seguridad nacional son las siguientes:

(a) Baja confianza y apoyo institucional, con continuidad en el tiempo, al estudio del futuro

El principal motivo de la desconfianza hacia la prospectiva deriva de su consideración, absolutamente errónea, como técnicas para la adivinación y predicción del futuro. La crítica más fundamentada puede derivar, frente a lo señalado por Berger, de la imposibilidad de su consideración como una ciencia, ni siquiera como una disciplina científica. Si se sigue el clásico método científico, la amplitud de los plazos de los informes prospectivos impide una adecuada evaluación de los mismos.

Asimismo, desde la perspectiva de los técnicos resulta necesaria una clarificación metodológica con la finalidad de limitar las críticas al subjetivismo de estas aportaciones, contrastando en algunas de ellas la ambición y los medios disponibles, con los parcos comentarios metodológicos incorporados (Jaime y Blanco, 2014). Así, realizar estudios prospectivos rigurosos requiere dedicación, recursos amplios y unos impulsores políticos y técnicos decididos y conscientes de su utilidad.

(b) El recurso a la prospectiva como producto, en lugar de como sistema

Las técnicas prospectivas configuran un marco de cohesión de conocimientos diversos y fraccionados. Aportan una visión holística, al precisar una concepción amplia, desde distintas disciplinas, del fenómeno a analizar. La prospectiva no tiene gran valor si se considera como una información de futuro y estática, sino como un sistema continuo de seguimiento y evaluación, que permita realizar correcciones sobre los escenarios que se plantean (Blanco y Jaime, 2014).

El International Centre for Counter-Terrorism (Bakker, 2012), desarrolló una investigación sobre referencias al futuro del terrorismo en 60 informes realizados por instituciones y expertos de prestigio. La conclusión fue que la mayoría de ellos carecían de unas bases metodológicas, y que prácticamente en ningún caso se mencionaban posibles dinámicas de cambio que permitieran determinar indicadores que permitieran monitorizar la evolución del fenómeno. En el mejor de los casos, estos trabajos eran proyecciones del presente u opiniones personales más o menos argumentadas, basadas en la intuición o la experiencia.

(c) Incertidumbre, presente y futuro

La prospectiva tiene sus pilares en las dificultades para gestionar la incertidumbre, la complejidad, la volatilidad y la ambigüedad, los elementos de lo que se ha venido a conocer como entornos VUCA (acrónimo en inglés de los cuatro elementos). De esta manera, el esfuerzo metodológico debe tratar de combinar el conocimiento sobre el futuro existente (por ejemplo: cuándo se celebran unas elecciones, las proyecciones de población o la introducción en el mercado de un desarrollo tecnológico) con otras variables o cuestiones inciertas (sobre su producción, evolución, plazo, impactos, alcance, etc.).

En este sentido, es de interés la aportación de Espósito (2011) que sigue la teoría sociológica de Niklas Luhman: el presente y el futuro incluyen una multiplicidad de futuros pasados y presentes. Aun en el desconocimiento, es posible disponer de una aceptable orientación sobre el futuro. De nuevo, se pone de manifiesto de la consideración de la prospectiva como un sistema continuo. Un trabajo que exige imaginación y creatividad (Blanco y Cohen, 2018).

(d) Aproximaciones cualitativas y escasez de estudios cuantitativos o mixtos

Muchos estudios de prospectiva, en línea con lo señalado por Bakker, se basan en aproximaciones, más o menos metódicas, de carácter cualitativo. Se muestra, especialmente en el ámbito de la seguridad, una ausencia de estudios cuantitativos

o que combinen ambas aproximaciones. Siguiendo a Bas (1999), es preciso, con objeto de tener una visión holística del futuro del fenómeno a analizar, combinar las diferentes aproximaciones, tal y como se ha señalado en la categorización de técnicas en el epígrafe previo.

(e) Multiplicidad e interdependencia entre variables

Por otra parte, los estudios de futuro existentes mayoritariamente obvian las interrelaciones entre variables múltiples, optando por atajos para determinar las variables clave. No incluyen análisis de correlación y de regresión, fundamentales para tratar de predecir evoluciones futuras.

Tratamientos que exigen conocimientos especializados y el apoyo tecnológico, puesto que el ser humano es incapaz de procesar la elevada cantidad de variables interrelacionadas en los fenómenos de seguridad, de determinar el grado de influencia (ponderación) de cada una de ellas, o de detectar posibles patrones.

(f) Escaso tratamiento de la probabilidad

Los escenarios son representaciones del futuro. Estos escenarios, en el ámbito de la seguridad, pueden llevar a sugerir futuros utópicos (más raramente) o distópicos (futuros indeseados). La prospectiva enraíza claramente con la evaluación de riesgos, siendo necesario realizar un adecuado análisis de probabilidad e impactos.

En general, en los estudios prospectivos hay una ausencia de tratamiento probabilístico adecuado, posiblemente debido a la falta de formación en dicha materia de gran parte de los analistas de seguridad y defensa.

(g) Ausencia de consideración de hechos disruptivos

Podríamos llegar a definir el futuro con esta fórmula:

F = (a x T) + (b x E) + (c x P), donde:

T = Tendencia histórica

E = eventos o sucesos inesperados

P = propósitos u objetivos individuales o colectivos

a, b, c = parámetros, siendo a + b + c = 1

El informe de investigación de los atentados del 11S incluyó un capítulo dedicado a la prospectiva, destacando que la falta de imaginación fue el mayor error a la hora de evitar el ataque. La creatividad no es habitual en los sistemas burocratizados y

en el pensamiento de decisores muy orientados a corto plazo. El informe señala la necesidad de "burocratizar" la imaginación, incorporar la misma en los procesos de análisis. Los *unknowns unknowns*, señalados por Donald Rumsfeld, son, muy probablemente, imposibles de gestionar. Taleb, por la imposibilidad de detectar los que denomina cisnes negros, hechos de baja probabilidad pero alto impacto, desarrolla el concepto de antifragilidad, una forma de gestionar la incertidumbre, no a través de la detección previa de cisnes negros, sino a través del desarrollo de las capacidades necesarias para enfrentarse al cambio, la disrupción, saliendo fortalecido de la adversidad.

(h) Tratamiento de los actores

Siguiendo la fórmula metafórica anterior, el futuro será en gran parte un efecto de las acciones que diferentes actores puedan acometer. En el presente, a título de ejemplo, se observa que frente a los desarrollos de la Inteligencia Artificial surgen una serie de actores, de influencia, que tratan de modelar los posibles impactos negativos de estas tecnologías, configurando de esta manera su utilización futura. En esta línea de acción es importante incorporar a los estudios prospectivos técnicas como el análisis de riesgos adversarios (Banks, Ríos y Ríos, 2015).

(i) Método Delphi para contrastar

Muchos estudios de futuro se basan en la utilización del método Delphi. Desarrollado por la *RAND Corporation* se trata de una aproximación que trata de lograr un consenso. Está alejado de algunos de los objetivos del análisis de inteligencia y de la prospectiva, donde se busca ampliar el número de alternativas y de hipótesis. En muchas ocasiones, el desarrollo de un Delphi presenta serias brechas metodológicas, tanto en la construcción de los cuestionarios como en la selección de los expertos. A su vez, surgen aproximaciones que ponen en duda el valor predictivo de los expertos, como el proyecto *The Good Judgement*, y las consideraciones de Tetlock y Gardner (2015).

(j) Escasez de formación en la materia

Prácticamente no existe una oferta formativa en la materia, siendo abordada de manera residual en cursos sobre análisis de inteligencia.

Bibliografía

Banks, L.; Ríos, J. M. y Ríos, D. (2015), *Adversarial Risk Analysis*, Chapman and Hall/CRC

Bas, Enric (1999), *Prospectiva. Cómo usar el pensamiento sobre el futuro*, Ariel social.

Berger, Gas *Phénoménologie du temps et prospective.* Paris: Presses Universitaires de France, 1964.

Berger, Gaston: *Étapes de la prospective.* París: PUF, 1967.

Blanco, J. M. (2016), Inteligencia prospectiva para la Seguridad y la Defensa, Revista Ejército nº 902, mayo 2016. http://www.ejercito.mde.es/Galerias/multimedia/revista-ejercito/2016/902/accesible/revista_ejercito_mayo_902.pdf

Blanco, J.M. y Oscar Jaime (2014). Toma de decisiones y visión de futuro para la seguridad nacional. En De la Corte y Blanco (2014) "Seguridad nacional, amenazas y respuestas". Editorial LID.

Blanco, J. M. y Cohen, J. (2018). La inteligencia en el mundo de la prospectiva, en Del Toro Jímenez (2018) La inteligencia empresarial en España (2018). Editorial Borrmart.

Godet M. y Coates, J. F. (2007). Prospectiva Estratégica: problemas y métodos. Cuadernos LIPSOR nº 20. http://www.prospektiker.es/prospectiva/caja-herramientas-2007.pdf

Kahn, H. y Wiener, A. J. (1967). The Year 2000: A Framework for Speculation on the Next Thirty-Three Years. New York: The Hudson Institute.

Lia, Brynjar (2005). *Globalisation and the Future of Terrorism. Patterns and Predictions.* Contemporary Security Studies. Routledge.

Makridakis, S., Weelwright, S. y McGee, V. (1983). Forecasting. Methods and Aplications. Nueva York: John Wiley and Sons.

National Commission on Terrorist Attacks: *The 9/11 Commission Report*, New York: Norton, 2004.

National Intelligence Council (2017). "Global Trends. Paradox of Progress". de https://www.dni.gov/files/documents/nic/GT-Full-Report.pdf

RAND Europe: "Europe´s Societal Challenges. An analysis of global societal trends to 2030 and their impact on the EU". RAND Europe, 2013. Recuperado de http://www.rand.org/pubs/research_reports/RR479.html[Consultado el 28/11/2017].

Schwartz, Peter: *The Art of the Long View.* New York: Currency Double day, 1996.

Tetlok, P. y Gardner, D. (2015). Superforecasting: The Art and Science of Prediction. Cornerstone Digital

Jessica Cohen es analista de inteligencia en materia de seguridad internacional; responsable de Unidad de Análisis de Inteligencia en el sector privado. Licenciada en Criminología, Máster en Análisis de Inteligencia. Profesora en la Universidad Autónoma de Madrid en materias como metodología, fuentes abiertas, estrategias de búsqueda de información, red teaming, apoyo a la toma de decisiones, insiders y gestión de crisis. Profesora en seminarios y títulos específicos en la Universidad San Jorge y en la Universidad Pontificia de Comillas. Investigadora y formadora en proyectos europeos sobre terrorismo, violencia política, extremismo y radicalización. Autora de múltiples artículos y capítulos de libros en materia de metodología de análisis, prospectiva, extremismo, radicalización y polarización social, tanto a nivel nacional como internacional.

José María Blanco es director de Ciberinteligencia Estratégica en Prosegur, siendo con anterioridad jefe del Centro de Análisis y Prospectiva de la Guardia Civil. También es jefe del Área de Estudios Estratégicos y de Inteligencia en el Instituto de Ciencias Forenses y de Seguridad (ICFS, Universidad Autónoma de Madrid), profesor de Análisis en Inteligencia en varias universidades, asesor en el Instituto Universitario de Investigación en Seguridad Nacional (IUISI-UNED) e investigador en varios proyectos europeos del Séptimo Programa Marco y del H2020 tales como VIRTUOSO (de inteligencia de código abierto), EPOOLICE (alerta temprana y crimen organizado), TRANSVERSAL BRIDGES (amenaza emergente de alianzas terroristas transversales y la radicalización del clima social de la UE) y MINDb4ACT (sobre radicalización terrorista). Tiene varias publicaciones sobre estrategias de seguridad, extremismo y radicalización, terrorismo, crimen organizado, conflictos internacionales e inteligencia.

Anticipando en tiempos postnormales

Jordi Serra del Pino

En 1993 Jerome Ravetz y Silvio Funtowicz pusieron de relieve la perplejidad y la incertidumbre que se siente ante situaciones en las que "los datos son inciertos, los valores en disputa, las apuestas altas y las decisiones urgentes" (1993: 744). Su respuesta ante esta situación fue definir un nuevo tipo de ciencia, la Ciencia Postnormal. Posteriormente, y ante la constatación de que este tipo de situaciones se estaban haciendo más y más recurrentes, Ziauddin Sardar teorizó que la Postnormalidad no se aplicaba solamente al ámbito científico, sino que era un fenómeno transversal. De hecho, postuló que la Postnormalidad era lo que definía la época que nos ha tocado vivir. En su artículo "Bienvenidos a los tiempos postnormales" (2010), Sardar describe este período según sus tres factores constituyentes: la complejidad, el caos y las contradicciones. Ciertamente, resulta muy complicado refutar que nuestra realidad es hoy significativamente más compleja que en el pasado. Igualmente es difícil obviar que asistimos a una mayor reiteración de sucesos caóticos.[28] Y es precisamente la mayor presencia de fenómenos complejos y/o caóticos, lo que ha provocado que la presencia de contradicciones se haya hecho mucho más prevalente. Por lo tanto, de alguna manera la realidad ha cambiado respecto a la de tiempos precedentes, pero no sólo eso, el propio cambio también ha mutado. En un mundo híper-conectado vemos cómo los sucesos se despliegan de forma acelerada, expansiva, incremental y simultánea. Así, el suceso más nimio puede convertirse en viral, llegar a cualquier rincón del planeta casi instantáneamente y, lo que resulta más inquietante, tener un impacto muy por encima de lo que pudiera haberse previsto anteriormente mientras se retroalimenta con otros sucesos. Por esta razón ya no sorprende que un *tuit* pueda arruinar una carrera política o profesional, que un don nadie condicione la geopolítica internacional, que una mentira pese más que un dato constatable o que un lego tenga más credibilidad que un académico. Efectivamente, el mundo se ha vuelto postnormal.

Pero, si aceptamos que la realidad ha cambiado, entonces, también tenemos que cuestionarnos cómo la percibimos. Es decir, si somos plenamente conscientes de esta transformación. Dicho de otro modo, hay que cerciorarse de que los mecanismos de percepción, procesado y análisis de esta nueva realidad son idóneos. Y, sin embargo,

28 Aquí conviene señalar que caótico no debe entenderse como lo opuesto a orden o a ordenado. Si no como aquella propiedad de algunos sistemas que, ante mínimas variaciones en sus estados iniciales, pueden experimentar modificaciones de gran impacto. Como postula, por ejemplo, el Efecto Mariposa.

en lo que a anticipación se refiere, seguimos utilizando las mismas herramientas y, dado que están dando muestras de agotamiento, lo razonable sería revisarlas y ponerlas al día. En este sentido, la Teoría de los Tiempos Postnormales (TTPN) propone una actualización de la prospectiva clásica que se ha basado, fundamentalmente, en el análisis del cambio pasado y presente. La TTPN en cambio, propugna un modelo de anticipación en el que, aunque también se arranca de ese mismo planteamiento clásico, se configura un enfoque que permite ampliar y profundizar el análisis del futuro. Y es justo en este punto donde hay que volver a examinar nuestros procesos cognitivos para ver si realmente son idóneos para este nuevo tipo de realidad y cambio en el que estamos inmersos.

Esto no es una cuestión menor en absoluto. Una de las capacidades más singulares de las personas es la de adaptarse bastante rápido a cualquier situación. Esta capacidad ha sido conceptualizada por Venkatesh Rao (2012) como el Campo de Normalidad Manufacturada (CNM). Este campo es un mecanismo extremadamente útil cuando nos enfrentamos a sucesos que nos superan; fundamentalmente, porque cuando algo nos genera angustia y estrés, conseguir normalizarlo nos calma. Pero el CNM también es una estrategia del cerebro para ahorrar energía. De hecho, la investigación en neurobiología confirma que nuestro cerebro está continuamente buscando formas de economizar energía (nuestra capacidad para olvidar sería un buen ejemplo de ello). En cualquier caso, el CNM trabaja incesantemente para convencernos de que lo que sea que experimentemos es normal y como tal hemos de tratarlo; sea lo que sea que percibimos, no hace falta inquietarse porque lo que vivimos es… normal y, en tanto que normal, no justifica un dispendio de energía extraordinario para procesarlo o preocuparse.

Pero hete aquí que nos hallamos en un momento en que nuestra vida se hace compleja, caótica y contradictoria; una época de cambio acelerado, expansivo, incremental y simultáneo que, como acertaron a detectar Ravetz y Funtowicz, hace que estemos, casi permanentemente, sumidos en una incertidumbre perpleja. Es como si nuestro CNM tuviera que estar haciendo horas extra para mantener una mínima apariencia de normalidad. Y es aquí donde distintos sesgos cognitivos, que los humanos hemos ido desarrollando con el paso del tiempo, intervienen para gestionar la creciente complejidad de nuestra vida social. De entre estos mecanismos, el pensamiento lineal ha sido uno de los más exitosos. Algo que, probablemente, empezó como la identificación de una secuencia en la ocurrencia de sucesos concretos y que pronto derivó en una inferencia de causalidad. En algunos casos, esta inferencia supuso una ventaja competitiva para aquellos que poseían ese conocimiento, como descubrir que ciertos cambios en el comportamiento animal podían anticipar una tormenta.

La necesidad de encontrar la razón o causa de los distintos fenómenos se convirtió en un gran motor del progreso humano y, a fecha de hoy, nuestra primera reacción ante cualquier problema es tratar de identificar la variable clave que nos permita resolverlo. Sin embargo, en la actualidad esta estrategia se enfrenta a un problema: en un entorno progresivamente más complejo el pensamiento lineal resulta insuficiente. El vínculo causal ya no puede representarse con una simple línea, sino que, la mayoría de las veces es como una red enmarañada con múltiples variables e interacciones. Por consiguiente, A ya no lleva a B y resulta iluso pretender que después llegaremos a C. Aquí el verdadero problema reside en que hemos interiorizado tanto la linealidad que cuando no la podemos identificar recurrimos a enfoques simplistas. Reducir una situación a un modelo simplista con el que se pretende entender lo que esté sucediendo puede convertirse en un ejercicio análogo a hacerse trampas jugando al solitario. Y así, aunque a menudo constatamos que el uso de modelos o enfoques simplistas falla, seguimos usándolos. Y sí, en algunos casos esto puede deberse al hecho de carecer de una alternativa mejor, pero en otros no es más que nuestra incapacidad (o arrogancia) para darnos cuenta de que hemos caído en un sesgo cognitivo. De hecho, pasa lo mismo con los planteamientos dicotómicos o el uso de la inducción. Sabemos que un enfoque dicotómico no siempre es fiable en un contexto cada vez más cuántico; pero seguimos creyendo que si algo es cierto lo opuesto será falso a pesar de que la profusión actual de contradicciones nos muestra que muchas veces las cosas son y no son a la vez. Y la inducción supone un caso similar, si bien puede ser una manera válida de deducir principios generales a partir de una observación limitada, también puede ser el origen de graves errores si olvidamos que todo se origina de una observación que es limitada. Por tanto, y ante cualquier intento de anticipar lo que pueda acontecer, un primer ejercicio de prudencia sería intentar descubrir y corregir estos sesgos cognitivos.

Pero centrémonos ahora en esa sensación de incertidumbre perpleja que describían Ravetz y Funtowicz. Efectivamente, la incertidumbre parece ser el *zeitgeist* de nuestra época y permea los distintos ámbitos de las sociedades contemporáneas. Con todo, resulta bastante obvio que no todo nos genera un nivel de incertidumbre equivalente. Por ejemplo, anticipar la nueva generación de teléfonos inteligentes, puede ser algo relativamente sencillo usando solamente la información de la que disponemos en la actualidad; en cambio, pronosticar el efecto a largo plazo de la inteligencia artificial en nuestros modelos económicos puede ser algo bastante más complicado. Por tanto, seguramente podemos aceptar que existen distintos niveles de incertidumbre; y a pesar de ello, tendemos a responder a todas estas incertidumbres de un modo bastante uniforme: recopilación de datos y búsqueda de información

para elaborar hipótesis explicativas. La TTPN propugna discriminar los distintos tipos de incertidumbre y buscar el enfoque más idóneo para cada uno. Y el concepto que la TTPN ha desarrollado para encapsular esta respuesta a la incertidumbre es el de ignorancia. Por lo tanto, y desde esta perspectiva teórica, la ignorancia no es sólo el déficit de conocimiento respecto a un tipo concreto de incertidumbre, sino también el proceso cognitivo con el que se intenta corregir ese déficit. Y a cada tipo de incertidumbre le corresponde una forma concreta de ignorancia.

Dentro de la TTPN se distinguen tres tipos de incertidumbre. La primera es la Incertidumbre Superficial que respondería a aquellas situaciones en las que ya poseemos una buena indicación de en qué dirección puede ir el cambio y, también, qué impacto pueda provocar este cambio. Un ejemplo de este tipo de incertidumbre puede ser el que genera la evolución del automóvil: si combinamos lo que sabemos sobre los desarrollos en la motorización (con modelos eléctricos, híbridos o tradicionales) con todo el campo de conducción autónoma y cuestiones socioeconómicas y jurídicas, seguramente podremos hacer una composición de lugar bastante ajustada. Probablemente, recurriremos a la investigación para obtener datos, procesaremos la información disponible y llegaremos a conclusiones mediante inferencias o analogías. Este tipo de respuesta define el tipo de ignorancia que la TTPN asocia a la Incertidumbre Superficial, la Ignorancia Plana. Este tipo de ignorancia es la más básica y, como tal, supone el método estándar de aprendizaje. En efecto, la respuesta tipo a cualquier déficit cognitivo consiste en aprender mediante la búsqueda de información, a menudo dando preferencia a los datos cuantitativos, para poder formular hipótesis de resolución y, cuando alguna se valida (o simplemente no se refuta), confirmarla como respuesta idónea; pero, y esto es lo más relevante aquí, es un tipo de aprendizaje que aplica abiertamente los sesgos que se han expuesto anteriormente: el pensamiento lineal, la dicotomía y la inducción. Es más, el binomio Incertidumbre Superficial e Ignorancia Plana vendría ser el modo por defecto del CNM, su *business as usual*. Por un lado, tenemos un tipo de incertidumbre que no es particularmente estresante y, por otro, podemos tratarla de un modo en el que podemos recurrir a toda la batería de mecanismos que nuestra mente ha desarrollado para el cerebro pueda ahorrar energía. Sería lo más parecido a la situación ideal para el CNM.

Sin embargo, los tiempos postnormales provocan que a menudo la incertidumbre sea más honda: situaciones en las que ya no estamos seguros de en qué dirección irá el cambio o, mucho menos, de qué impacto pueda tener. Estaríamos hablando de cuestiones como cuál puede ser el impacto a medio plazo de la política aislacionista

de Trump o del auge de los populismos autoritarios; pero también del consumo de alimentos transgénicos o de los efectos secundarios de la cirugía ocular con láser. Esta es la Incertidumbre Somera, una incertidumbre que no puede ser resuelta con la Ignorancia Plana: aquí hay que recurrir a la Ignorancia Vencible. Esta supone un salto cualitativo respecto a la plana, ya que, de entrada, requiere un paso previo para poder accionarla: ser consciente de lo que se desconoce. Si bien la Ignorancia Plana puede ser accionada de una manera directa, a piñón fijo por así decirlo; en el caso de la Vencible no funcionará si no se ha dado este paso anterior de ser consciente de ese déficit de conocimiento. Esta constatación puede llevarnos a dos tipos de conclusiones: o bien que comporte el descubrimiento de que nuestro ámbito de conocimiento no es suficiente y necesitamos incorporar saber de otros ámbitos o campos, si se quiere, que tenemos que abrir nuestra perspectiva; o, en el segundo caso, puede resultar en tener que aceptar que simplemente hay cuestiones que no se pueden dilucidar en el momento presente y que requieren amasar un cierto grado de experiencia para poder comprender su funcionamiento e implicaciones. Por tanto, la Ignorancia Vencible implica disponer de un cierto grado de humildad y reconocer que no todo es accesible desde una única perspectiva o momento.

Pero los tiempos postnormales también se caracterizan por ser caóticos y eso significa que, en algunas ocasiones, nuestra incertidumbre será mucho mayor y virtualmente insondable. A modo de ejemplo, aunque podamos especular, lo cierto es que no tenemos manera de saber qué efecto puede tener un cambio climático exacerbado y acelerado en nuestras sociedades, de igual modo que no podemos abarcar los efectos de un uso descontrolado de la ingeniería genética en nuestra especie. Aquí estamos delante de la Incertidumbre Profunda, seguramente la incertidumbre más genuina e inquietante porque es la que más nos pone delante del espejo de nuestra vulnerabilidad y nuestras carencias. Y ante este tipo de incertidumbre recurrir a la Ignorancia Plana (o aprendizaje habitual) no servirá. Por un lado, difícilmente encontraremos datos o información que nos sean de utilidad y, por el otro, y aquí es donde más se pone de relieve el riesgo de nuestros sesgos cognitivos, si el fundamento implícito de nuestro enfoque tiende a minusvalorar la dimensión e impacto del cambio, incluso ante fenómenos de alto poder disruptivo, dicho sesgo nos conducirá a conclusiones erróneas. Aquí necesitamos un nuevo enfoque, el que la TTPN ha conceptualizado en el tercer tipo de ignorancia, la Ignorancia Invencible.

Lo que postula la Ignorancia Invencible, especialmente en lo que a anticipación respecta, es que a menudo lo que más se interpone en la comprensión del futuro es justamente lo que creemos que sabemos. Si aceptamos que nuestra mente ha

generado un mecanismo para normalizar el cambio para reducir la generación de estrés y el consumo de energía, entonces resulta legítimo preguntarse qué partes de ese conocimiento están condicionando nuestra visión del futuro. En este punto hay que fijarse en la cosmovisión del individuo. El ser humano hace milenios que ha ralentizado su evolución biológica a la par que ha ido acrecentando su evolución cultural. Actualmente, la vida en sociedad supone incorporar un conjunto de normas y pautas que nos permiten conducirnos en una cultura dada. Pero cada cosmovisión comprende un conjunto de axiomas y suposiciones, generalmente implícitas, que justamente nos permiten "aculturarnos" en esa tradición, pero que también provocan que determinados aspectos de esa cultura queden fuera del análisis crítico. Pues bien, ante situaciones de Incertidumbre Profunda hay que cuestionarse hasta qué punto nuestra cosmovisión, nuestro campo teórico o nuestro enfoque nos están impidiendo comprender aquello que investigamos, aquí no nos quedará más remedio que desaprender. Es decir, tendremos que decidir qué partes de nuestro sistema cognitivo hay que desarmar para poder construir otro más idóneo o mejor adaptado al tema que nos ocupe.

Ahora bien, y aunque aceptemos todo lo que se ha expuesto hasta ahora, hay que tener en cuenta que el CNM siempre intentará aplicar su modo por defecto, Incertidumbre Superficial con Ignorancia Plana. Es decir, hará que nuestra primera respuesta ante cualquier tipo de crisis, situación o reto incierto sea intentar aprender sobre ella aplicando la batería habitual de recursos (pensamiento lineal, dicotómico e inductivo). El problema radicará si esta crisis se enmarca en un tipo de incertidumbre vencible o invencible, en estos casos la Ignorancia Plana no sirve, se queda corta o, peor aún, agrava la situación al promover respuestas inadecuadas. Y, sin embargo, el efecto más negativo de aplicar una respuesta de Ignorancia Plana es que genera la sensación de que se está haciendo algo al respecto, quizás incluso, que se está solventando la cuestión, cuando, en el mejor de los casos, sólo se estén tratando sus síntomas. En estos supuestos se produce una fractura entre el tipo de incertidumbre en el que se contextualiza la crisis, que es más profunda, y nuestra respuesta, que es más superficial. Este desfase recibe el nombre de Retraso Postnormal en la TTPN y uno de los ejemplos más flagrantes que podemos encontrar sería la negación del cambio climático. Pero es precisamente este ejemplo el que nos pone sobre la pista de otro aspecto importante en lo que a anticipación respecta: ¿Hasta qué punto nuestro juicio racional se ve condicionado por nuestras preferencias irracionales? Es decir, cuando negamos el cambio climático ¿lo hacemos basándonos en un análisis o reflexión controvertidos o porque nos desagrada aceptarlo? Todo lo expuesto aquí hasta ahora se centraría en tratar de corregir los sesgos que enturbian el análisis

racional, pero lo cierto es que resulta muy ingenuo creer que es posible obviar completamente nuestras preferencias irracionales cuando se explora el futuro. Por lo tanto, resulta imprescindible dotarnos de instrumentos para detectar, y corregir, este tipo de sesgos. Y precisamente por esta razón la TTPN ha concebido un instrumento para poder aumentar nuestra capacidad de reconocer y combatir todo tipo de sesgos: la fauna postnormal.[29]

El primero es el "Elefante Negro" y se trataría de aquellos sucesos muy probables con un alto impacto potencial asociado y, a menudo, ampliamente anunciados, que ignoramos, o preferimos ignorar, individual o colectivamente. Claramente, es el animal que nos puede ayudar a cuestionar si nuestras conclusiones son el resultado de un análisis riguroso o si hemos caído en un cálculo teleológico. Con todo, los elefantes son el miembro de la fauna comparativamente más fácil de descubrir siempre que se actúe con un mínimo de honradez profesional. El segundo animal, es el "Cisne Negro" (concepto acuñado por Nicolas Nassim Taleb), quizá es más conocido, pero resulta mucho más arduo de localizar, ya que se refiere a fenómenos atípicos, que generalmente no hemos percibido ni anticipado. No es infrecuente que se hayan desestimado previamente como fallos del sistema, ruido de fondo o excepciones a la regla que no merecen más reflexión. Aquí tratamos con el sesgo cognitivo más genuino ya que a menudo es el resultado de no haber querido profundizar en contradicciones o aspectos que cuestionaban el resultado de un enfoque lineal, dicotómico o inductivo. Y, además, resulta más difícil de ver porque el CNM trabaja para hacerlo invisible (ya que contradice su modo estándar de percepción y comprensión de la realidad). A lo sumo un Cisne Negro será percibido como un vacío o una ausencia, como una discontinuidad o una discordancia. Finalmente, llegamos al tercer animal, creado específicamente para la TTPN, las "Medusas Negras"; aquí hablaríamos de fenómenos "normales" que creemos conocer y controlar pero que, bajo determinadas circunstancias y de forma inesperada, pueden exhibir un comportamiento imprevisto con un alto efecto asociado. Las medusas nos ponen en guardia ante falsas sensaciones (o expectativas) de control, aquellos casos en que se minimiza la relevancia de algo porque se considera nimio y/o fácil de controlar. De alguna manera, las Medusas Negras tratan de ayudarnos a identificar qué elementos de nuestro trabajo pueden tener un comportamiento caótico y salirse del guion previsto.

29 Es de justicia reconocer aquí que no todos los componentes de la fauna postnormal han sido creados desde la TTPN. En dos casos hemos aprovechado conceptos preexistentes que encajan perfectamente con nuestra intención en esta cuestión.

Llegado a este punto, me pregunto si alguien ha empezado la lectura de este artículo esperando alguna indicación sobre cómo poder anticipar de una manera más fiable y, quizás, más sencilla. Me temo que ese caso estará decepcionado. Cuanto más entendemos la lógica evolutiva de lo que analizamos, más nos damos cuenta de la necesidad de tener una actitud muy humilde y autocrítica para desarrollar nuestro análisis. No obstante, esto no disminuye un ápice la fascinación y la gratificación que se puede obtener cuando se consigue desentrañar algún aspecto de lo que estamos analizando.

Bibliografía

Funtowicz, Silvio O. y Ravetz, Jerome R. (1993), "Science for the Post-Normal Age", *Futures* 25 (7) 739-755.

Rao, Venkatesh (2012), "Welcome to the Future Nauseous", *Ribbonfarm*. https://www.ribbonfarm.com/2012/05/09/welcome-to-the-future-nauseous/

Sardar, Ziauddin (2010), "Welcome to postnormal times", *Futures* 42 (5) 435-444.

Jordi Serra del Pino es consultor especializado en prospectiva, estrategia e inteligencia. Actualmente es el director de investigación del Center for Postnormal Policy & Futures Studies; director académico del Observatori de l'economia social OES21; profesor asociado en Blanquerna - Universitat Ramon Llull; Fellow y vicepresidente del Capítulo Iberoamericano de la World Futures Studies Federation. También es miembro del comité editorial de Futures, World Future Review (revistas especializadas en prospectiva) y de la Revista IAPEM (Instituto de Administración Pública del Estado de México). Igualmente ha desarrollado una intensa actividad como escritor, contribuyendo en las principales revistas especializadas y otros medios generalistas, como docente en distintas universidades españolas y como ponente en varios tipos de actividades en los cinco continentes.

El sesgo de la historia: reflexiones sobre la causalidad

Eva Moya y Eugenia Hernández

Famoso es el lema de los historiadores, *conocer el pasado, para comprender el presente y determinar cuál será el futuro.* La base de este lema está formada por un conglomerado de estudios sobre "qué provocó lo que pasó". Es decir, un estudio exhaustivo de las causas que llevaron a que tuviera lugar un cambio histórico, un acontecimiento considerado relevante por los historiadores o simplemente una razón de ser del porqué de un comportamiento sociocultural.

Cabría pensar que la proyección hacia el pasado resulta más sencilla que la proyección hacia el futuro. De esta manera, se podría considerar al historiador un analista que juega con la ventaja de analizar acontecimientos que ya han sucedido y que, por tanto, dispone de más información para determinar cuáles fueron sus causas. Sin embargo, no es así. Cuanto más nos alejamos en el tiempo, menos información tenemos y más obligado está el historiador a interpretar lo que sucedió. Mientras, cuanto más nos acercamos al presente, más información y más expuestos estamos a las consecuencias del término de moda "infoxicación". Ante tanta información dispar, que en ocasiones incluso es falsa, el historiador vuelve a tener que interpretar la realidad.

Para elaborar el discurso, uno de los focos de atención de los historiadores es la investigación de las causas que han participado en la evolución de la historia. Sin embargo, cabe destacar que la mirada del historiador se ha visto influenciada por el momento cultural en el que vivía. Por tanto, los historiadores tenían a bien prestar atención a una o varias tipologías muy concretas de causas. Sin embargo, como se verá en este capítulo, conforme el estudio de la historia se ha ido profesionalizando a lo largo de los siglos, la selección de causas ha sido cada vez más rica.

Así pues, quisiéramos comenzar realizando un breve paseo por el tiempo historiográfico para comprender cómo el momento cultural ha ido influyendo en la interpretación de la realidad y, por tanto, en la selección de las causas e interpretación de las consecuencias que la conformaron.

La primera de las reflexiones tiene que ver con la esencia del tiempo en sí misma. ¿Considera el lector que el tiempo es lineal? ¿Es circular?

A lo largo de los diferentes momentos culturales el tiempo ha variado en su concepción. Aparentemente, puede parecer irrelevante para el analista, sin embargo, dicha concepción temporal determinará la naturaleza de la selección e influencia de las causas, por extraño que pueda parecer.

Si realizamos una retrospección historiográfica será interesante reflexionar sobre la concepción del tiempo en época premoderna y, así, podremos retrotraernos a los orígenes de la cuestión.

En la antigua Grecia surge el concepto del "eterno retorno", es decir, una visión cíclica del tiempo. Esta corriente era aplicada a los mitos, la narrativa histórico-mítica, la agricultura y la vida cotidiana. Inspirados en los ciclos astronómicos y en las estaciones, los antiguos griegos consideraban que, tarde o temprano, "todo vuelve". Y bajo esta concepción circular, organizaban su día a día. También en la Edad Media existía la creencia de que el tiempo, desde el punto de vista de la vida cotidiana, era cíclico o circular. En este caso, inspirados por una visión regeneradora y glorificadora del universo, que emanaba de la intervención divina. Así pues, a cada vuelta, el ser humano podía perfeccionarse a sí mismo. Con esta mirada, y desde el punto de vista del análisis de Inteligencia, si el analista considera el tiempo en su concepción cíclica puede asumir una alta probabilidad de que las mismas causas pasadas tendrán las mismas consecuencias futuras.

Sin embargo, desde el punto de vista contrario, puede tenerse en cuenta la concepción lineal del tiempo introducida y desarrollada también en época premoderna. Así pues, si atendemos al padre de la historiografía, Heródoto, en la antigua Grecia, descubriremos en sus textos el reconocimiento de que nada se repite en la historia. Según sus palabras en la *Historia de las Guerras Médicas*, "no hay día igual". Así pues, los griegos convivían con la idea de un tiempo lineal para explicar el devenir de las civilizaciones desde el punto de vista político y militar. Respecto a la Edad Media, sucede algo similar a raíz de la visión del "tiempo escatológico" cristiano. Esta visión considera la evolución y desarrollo histórico como algo lineal. De esta manera, la historia resulta ser el desarrollo de una serie de causas y consecuencias que dirigen la evolución hasta la llegada del Mesías, como fin último de la historia de la humanidad. Desde el punto de vista del análisis de Inteligencia, si el analista considera que el tiempo es lineal, será más difícil valorar la probabilidad de que una misma causa tenga las mismas consecuencias y, por tanto, dificulta su valoración para las proyecciones hacia el futuro.

Como se observa, en las culturas premodernas convivieron ya las dos concepciones más reconocidas del tiempo. Una circular asociada a los ritmos de la naturaleza y una lineal como consecuencia de los ritmos marcados por el ser humano.

Si bien es cierto, hoy día se reflexiona sobre un punto intermedio entre ambas. Historiadores como Noah Harari plantean la posibilidad de que el tiempo sea espiral. Esto significaría que las grandes tendencias presentes en la historia de la humanidad, como por ejemplo las luchas de poder y los conflictos entre pueblos, se repitan a lo largo de la historia con diferencias. Si el analista de inteligencia considera este planteamiento, podría seleccionar posibles causas y consecuencias ya confirmadas en el pasado para la proyección de escenarios, pero teniendo en cuenta que debería ir un paso más allá para ser capaz de adelantar cuáles serían las variaciones.

Este planteamiento nos conduce a la siguiente reflexión. ¿Cuál es la naturaleza de las causas que se deberían analizar relacionadas con las consecuencias observadas?

Hasta la época premoderna las causas culturales tenían un origen divino o sagrado, mientras que las causas políticas y militares pertenecían al mundo de los humanos. Así, la historia como tal, se elabora según una serie de criterios observados en el momento de la narrativa. Estos criterios contemplaban, por ejemplo, todo tipo de causas relacionadas con los gobernantes y el poder. Si ascendía al poder o al trono una personalidad más beligerante o deseosa de unificar territorios, era probable que se produjeran varios conflictos con la intención de expandir el territorio. En esta misma línea, los matrimonios reales se convertían en una causa de adhesión política y cultural, y generaban nuevas afinidades y descontentos que transformaban el contexto histórico.

El ámbito militar estaba muy presente. Las probabilidades se medían según el número de recursos humanos y tecnológicos, y los servicios de inteligencia se centraban en obtener información de qué movimientos y cuándo iban a realizar los contrarios. Por tanto, desde este punto de vista, los analistas observaban sólo aquellas causas y movimientos relacionados con el poder, obviando todo lo demás, por considerar que no influía en el cambio histórico como tal.

A partir del siglo XIX, con la llegada de la industrialización, el capitalismo y sucesos históricos de gran trascendencia como la Revolución Francesa o las Guerras Mundiales, los historiadores comienzan a reflexionar sobre otros ámbitos propios de la sociedad en la que viven. El estudio de la historia se convierte en una disciplina

científica y los historiadores realizan un profundo análisis desde diferentes puntos de vista e ideologías propias del momento.

Para el desarrollo de esta reflexión no es necesario acudir a todas ellas. Simplemente se mencionarán aquellas consideradas más relevantes para el objetivo de este capítulo.

Así, por ejemplo, una de las más relevantes es la escuela marxista. Esta escuela interpreta las causas y consecuencias bajo la lupa del llamado "materialismo histórico". Bajo esta óptica, adquiere mucha fuerza la interpretación de causas socioeconómicas derivadas de la industrialización, así como las migraciones y la presión demográfica del campo a la ciudad. Del nacimiento del concepto de "clase" aparecen nuevas causas de estudio derivadas del choque y luchas de poder entre las mismas.

Por otro lado, en el siglo XX surge otra escuela llamada "Annales". Los Annales transforman por completo el objeto de estudio histórico. El grupo de historiadores que sigue esta corriente se obsesiona con la idea de volver a investigar el pasado en busca de todo aquello que pasó desapercibido y pudo tener un efecto real en la transformación histórica, más allá de lo político y militar. Por ejemplo, aparece el concepto de "historia total", que plantea que no se puede entender ningún hecho histórico sin analizar las relaciones sociales o realidades colectivas. En definitiva, promovían el saber enciclopédico, integrando múltiples perspectivas y múltiples causas de naturalezas muy diferentes para interpretar la realidad.

Otro fenómeno muy interesante tiene lugar también en el siglo XX. Se produce el desarrollo de una nueva perspectiva en la disciplina historiográfica denominada "microhistoria". La microhistoria va un paso más allá de la interpretación de clases como estudio social para centrarse en el análisis del individuo y su vida cotidiana. Los ejemplos más conocidos de este tipo de estudio se reflejan en dar respuesta, por ejemplo, a cómo es su alimentación, cómo viste, cómo es la vida de las mujeres y los niños, etc. Una reflexión que, de nuevo, multiplica exponencialmente las posibilidades de estudio del pasado y permite una reinterpretación de los hechos históricos. La microhistoria consolida el contexto en el que tuvieron lugar, originando un racimo de nuevas causas propias del entorno fértil necesario para que se dé una consecuencia.

Así pues, la realidad se vuelve más compleja en época contemporánea, dado el estudio de causas de naturaleza política, militar, económica y social. El entramado de interacción se hace más grande y la asignación de probabilidades más complicada. Por otro lado, el reto ya no sólo era identificar las causas como tales, sino analizar su interacción y si alguna causa había tenido más influencia que otra sobre las consecuencias que transformaron el ciclo histórico.

A estas alturas, el lector ya se habrá hecho consciente de la relevancia de la perspectiva humana en la determinación de las causas que provocan determinadas consecuencias. A cada paso histórico aparecen nuevas perspectivas que lejanamente habría podido contemplar un analista de una época anterior. Cabe pensar que, si esto ha sucedido a lo largo de la historia, sea una tendencia que continúe sucediendo en el futuro. Nadie sabe pues qué naturalezas causales adicionales podrán llegar a identificarse o plantearse para poder proyectar escenarios más probables.

De hecho, hoy día ya vivimos otra gran transformación gracias al desarrollo acelerado de las nuevas tecnologías que están cambiando por completo la visión actual del mundo. Si bien es cierto que a lo largo de la historia se ha reflejado la evolución tecnológica, especialmente para explicar la evolución del hombre prehistórico, en el futuro cabe pensar que se estudiará Internet y las nuevas formas de comunicación como se estudió en su momento la aparición de la rueda.

Es muy probable que también se tengan en cuenta fenómenos medioambientales, como causas de la transformación histórica. Algo a lo que se le había prestado atención en el estudio de los periodos de glaciación y algunos momentos del pasado que podrían explicar las malas cosechas y el surgimiento de enfermedades que mermaron la capacidad demográfica de las civilizaciones.

Hemos reflexionado ya sobre la concepción del tiempo o la influencia del momento cultural en la selección de las causas y consecuencias. Pero queremos ir un poco más allá en nuestras reflexiones y terminar adentrándonos en otro punto que consideramos relevante: el conocido debate sobre los efectos derivados de la interacción entre el observador (historiador, analista, investigador, etc.) y lo observado (por ejemplo, la relación entre causas y consecuencias). Un debate siempre candente en el ámbito de las ciencias sociales, gracias, entre otros a Bronislaw Malinowski quien en *Los Argonautas del Pacífico Occidental* plantea la gran paradoja de la investigación: el observador observado.

Así pues, ¿hasta qué punto el observador es transformado por lo que observa? Gombrich, en su *Historia del Arte*, por ejemplo, realiza una reflexión profunda sobre los sentimientos y conceptos que se transmiten a través de diferentes canales y formas artísticas como la pintura, la escultura, la música, etc. Demuestra que la mera observación del objeto artístico transforma al observador, que pasa a integrar las nuevas ideas y sensaciones, afectándole en subsiguientes observaciones. Esta analogía puede servirnos para entender cómo el objeto de estudio condiciona al propio observador en otras disciplinas, como por ejemplo en la investigación histórica o en el análisis de inteligencia. Las condiciones de contacto con los hechos causales pueden transformar la mente del analista para sucesivas interpretaciones y proyecciones de los escenarios.

Por otro lado, sabemos que el observador realiza transferencias propias sobre el objeto de estudio. Si bien es cierto que existe mucha literatura respecto al conocido como "sesgo del analista", nos vemos en la necesidad de recordar su relevancia para cerrar este capítulo. Las transferencias se producen como elemento común en las áreas del conocimiento en el que el investigador interactúa con el objeto de estudio, entre ellas, el estudio de las causas, consecuencias e interacciones, más aún con la intención de proyectar escenarios futuros.

Malinowski, en su planteamiento de la "observación participante", señala que el investigador es el instrumento final de recogida del dato cualitativo, por lo que este debe saber reconocer las disimilitudes propias frente al objeto de estudio y reducirlas. Es decir, la predisposición del investigador a realizar una "traducción" del objeto observado (datos, información, escenarios, etc.) le obliga a conocer sus sesgos cognitivos y sus marcos de referencia si desea llegar a la verdad del objeto de estudio.

Como ya sabemos, el marco inicial del propio analista viene configurado por un perfil determinado de personalidad, una formación determinada, unos valores y maneras de leer el mundo que lo rodea que se convierten en los pilares del producto intelectual. En este sentido, se debe tener en cuenta también aquellas situaciones en las que el observador puede llegar a ser múltiple, por ejemplo, cuando intervienen las instituciones en el proceso de análisis e interpretación de los hechos.

Como es natural y lógico, las instituciones se inspiran en políticas públicas que marcan los intereses de estudio de cada sociedad, en cada momento cultural. A través de nuevas leyes y regulaciones, mediante la financiación, o incluso apoyando

a través de premios y menciones, las instituciones participan en un posible sesgo en la selección de lo que se consideran causas relevantes frente a otras en la interpretación de escenarios futuros.

Llegados a este punto, cabe reseñar los estudios realizados por Philip Tetlock en los que plantea e identifica la existencia necesaria de unos "superpronosticadores". Dedicado a estudiar un grupo de control al que se le plantearon preguntas de futuros probables, Tetlock identificó fallos en la CIA (inteligencia estadounidense) al no haber sido capaz de predecir la caída del muro de Berlín. En su estudio, descubrió que un grupo de su investigación sí evaluó esta posibilidad. Paradójicamente el grupo estaba formado por agentes recién llegados que afloraron la insostenibilidad de la situación, pero fueron ignorados por el sistema.

Los denominados "superpronosticadores"[30] son librepensadores, grupos de individuos que analizan las probabilidades de futuro, al margen de los informes de expertos. No se aferran a grandes teorías, al momento cultural en el que viven, ni les generan rechazo los conceptos nuevos. Incluso son capaces de adoptar ideas o propuestas contrarias a las plateadas por ellos previamente si consideran que existe una mayor conexión con la realidad.

Quisiéramos recordar las palabras de San Agustin *Nullus hostis metuatur extrinsecus: te vince et mundus est vinctus* o "no temáis a ningún enemigo exterior: véncete a ti mismo y el mundo será vencido". Estas palabras nos retrotraen a un momento singular de la historia de la humanidad, en la que no se buscaba la objetividad. San Agustín, sin embargo, como buscador de la libertad personal, nos recuerda la necesidad del desprendimiento personal. Superarse a uno mismo y ser capaz de superar el marco que nos constriñe (el enemigo externo) se hace imprescindible para poder vencer los elementos marco que nos rodean. Así, el desprendimiento personal se convierte en el perfecto censor.

A modo de conclusión, en lo que se refiere a la selección, análisis y evaluación de las causas y consecuencias con la intención de proyectar escenarios futuros el historiador, investigador o analista debe vencer el condicionamiento del propio pensamiento y de las restricciones conceptuales propias del momento cultural en el que vive.

30 Para más información, véase el capítulo acerca de "The Good Judgment Project".

Precisamente, en nuestro momento cultural actual, cabe destacar la relevancia de no confundir causalidad con correlación. Especialmente ahora que el aprendizaje automático se está incorporando a todas las áreas de la sociedad, incluyendo al área de Inteligencia.

Los analistas, investigadores y los equipos de desarrolladores de herramientas para Inteligencia, ante la necesidad de trabajar grandes volúmenes de información y datos, podrían sentir la tentación de señalar que ciertas correlaciones se comportan como relaciones causales, favoreciendo así la aparición del sesgo y contaminando todo el análisis que de ellos se pudieran derivar.

Bibliografía

Ginzburg, Carlo. (1994) *Microhistoria: dos o tres cosas que sé de ella*, Manuscrits, 1994.

Gombrich, E.H. (2008) *La historia del arte*, Phaidon Press Limits.

Iggers, Georg. (2014) *La historiografía del siglo XX. Desde la objetividad científica al desafío posmoderno*, Fondo de cultura económica.

Kawulich, B. (2005) *La observación participante como método de recolección de datos*, Forum: qualitative social research.

Ortega Cervigón, José Ignacio. (1999) *La medida del tiempo en la Edad Media. El ejemplo de las crónicas cristianas*, Universidad Complutense de Madrid.

Plácido Suarez, Domingo. (2004) *El tiempo, la ciudad y la historia en la Grecia clásica*, Universidad Complutense de Madrid.

Tetlock, P. (2016) *Superforecasting: The Art and Science of Prediction*, Broadway Books.

Eva Moya es licenciada en Historia por la Universidad de Vigo. Máster en Análisis de Inteligencia (UC3M y URJC) y Máster en Dirección y Gestión de la Identidad Corporativa (Madrid School of Marketing). Especializada en Análisis de Inteligencia en Fuentes Abiertas y Análisis de Redes Sociales. En los últimos 10 años se ha dedicado al Diseño, Implantación, Gestión y Dirección de Unidades de Inteligencia con orientación a ciberseguridad, seguridad corporativa y reputación; así como a la consultoría en compañías del IBEX35. Actualmente es Cyber Threat Intelligence en Mnemo. Profesora de OSINT y gestión del conocimiento en programas Universitarios en Inteligencia y de las Fuerzas y Cuerpos de Seguridad del Estado. Entre sus publicaciones destacan "Las Redes Sociales en el Devenir de los Estados" (Cuadernos de Estrategia 197 IEEE), "Open Goverment en Seguridad Pública e Inteligencia" (Ministerio de Interior) y "Redes Sociales y seguridad ciudadana. Casos de Éxito" (GIGAAP).

Eugenia Hernández es licenciada en la Facultad de CC Políticas de la UCM, especializada en RRII y Análisis Político. Máster en acción Política en el Estado de Derecho por la Universidad Francisco de Vitoria. Experta en Análisis de Inteligencia, en el ámbito de las Reservas de Inteligencia y Comunicación, así como en Análisis del Discurso. Fundadora de la Consultora Prime Project, trabaja desde hace más de 10 años como directora de Desarrollo de Proyectos, Relaciones Institucionales y Asuntos Públicos en proyectos multinacionales de sectores estratégicos por todo el mundo. Colabora con diferentes Instituciones para el desarrollo de la Cultura de Seguridad y Defensa en España. Investigadora en el Centro de Seguridad Internacional del IPI (UFV). Personal investigador de la Fundación Ortega y Gasset, realiza en la actualidad su tesis doctoral sobre "Retos de la Inteligencia Española ante la Cuarta Revolución Industrial-Revolución Tecnológica". Es miembro de Spanish Women in Security.

Capítulo III: Análisis de riesgos

Métodos bayesianos para el análisis de Inteligencia

Fabrizio Ruggeri

Resumen

Para ser coherentes, la evaluación y análisis de problemas complejos de toma de decisiones requieren la ponderación de múltiples fuentes de incertidumbre, de objetivos contradictorios y de preferencias multifacéticas que cambian con el tiempo, así como la integración de las opiniones y deseos de diversos grupos de partes interesadas. Nuestro objetivo general es presentar las metodologías bayesianas, instrumentos gráficos y computacionales para fundamentar esta toma de decisiones de forma racional, inclusiva y analítica. Nos focalizaremos especialmente en los dos aspectos principales del proceso de toma de decisiones, que son las más adecuadas para nuestro contexto: la estructuración de problemas y la elicitación de creencias (*beliefs elicitation*). A través de ejemplos muy básicos, ilustraremos cómo los procesos de toma de decisiones pueden ser descritos utilizando diagramas de influencia, y cómo ciertos problemas complejos que contienen estocasticidad pueden ser modelados utilizando una estructura de dependencia simple proporcionada por las redes bayesianas. Acabamos el artículo con un breve ejemplo del método estadístico bayesiano, que es capaz de combinar el conocimiento de los expertos y la información que nos proporcionan los datos, y también mencionamos cómo abordar un aspecto crítico de este método: la sensibilidad con respecto a las distintas creencias.

Introducción

El paradigma bayesiano del análisis de decisiones nos proporciona lo que podría ser el único marco justificado y unificado para fundamentar la toma de decisiones. Sin embargo, este método estaba confinado hasta muy recientemente a contextos muy sencillos, limitado a fundamentar las decisiones de pequeños grupos coherentes de tomadores de decisiones, ya que no se podía utilizar más ampliamente debido a obstáculos técnicos y computacionales. En estos últimos años, este ámbito ha experimentado una revolución técnica que ha abierto el camino a que los modelos y el análisis bayesianos se utilicen a mayor escala. Nuevos desarrollos en áreas como la modelización no lineal, la estructuración de problemas, la selección de modelos y

la computación bayesiana, así como el abaratamiento de los equipos informáticos, ahora permiten llevar a cabo una modelización más compleja y realista que puede ser explorada y presentada utilizando interfaces gráficos interactivos e intuitivos.

Los sistemas basados en métodos bayesianos han sido criticados en el pasado porque a los usuarios les parecían opacos, porque no opinaban que los razonamientos en los que se basaban fuesen intuitivos, y porque la cognición no guiada y el comportamiento en situaciones de incertidumbre no suelen ajustarse al método bayesiano. Sin embargo, los avances que han experimentado los sistemas de explicación basados en la inteligencia artificial, así como el acceso generalizado a internet, nos ofrecen muchas nuevas maneras mediante las cuales estas técnicas pueden utilizarse para fundamentar el proceso de toma de decisiones.

Las fases típicas de un proceso (bayesiano) de toma de decisiones son las siguientes: la estructuración de problemas, la modelización de la incertidumbre, la modelización de las preferencias, la maximización de la utilidad esperada, y el análisis de sensibilidad. En este artículo nos centraremos sobre las dos primeras y la última, ya que la modelización de las preferencias y la maximización de la utilidad esperada son fases posteriores a aquellas que realizan los analistas en los cuales se focaliza este artículo. Sin embargo, los métodos expuestos en este artículo también son útiles para los tomadores de decisiones, que podrán analizar y comparar distintos escenarios. Los analistas y los tomadores de decisiones podrán estructurar los problemas que tienen que abordar, así como explorar la dependencia estocástica de las entidades en cuestión, a través del uso de instrumentos gráficos tales como los diagramas de influencia y las redes bayesianas, que describiremos en las próximas secciones.

La estructuración de problemas

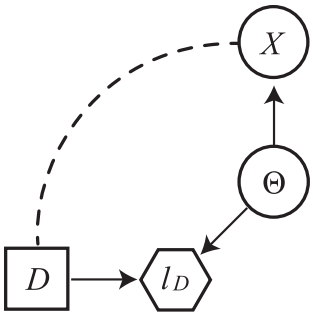
Uno de los problemas más difíciles de resolver en este tema es el de cómo estructurar el problema en cuestión en una toma de decisiones de forma adecuada. Hasta cierto punto, suele ser más bien un arte que requiere mucha creatividad y comprensión de distintas perspectivas cognitivas por parte de los analistas para llegar a la esencia de un problema. Recientemente, se han propuesto muchos métodos para capturar la estructura de un problema en términos de incertidumbres, decisiones y evaluaciones, que incluyen los diagramas de influencia y las redes bayesianas que tratamos en las próximas secciones.

Diagramas de influencia

Un diagrama de influencia (DI) es un método que se utiliza para representar problemas de toma de decisiones de forma gráfica. Un DI es un grafo acíclico dirigido (GAD) que contiene varios tipos de nodos: nodos de azar, representados con un círculo, nodos de decisión, representados con un cuadrado, pérdidas, representadas con un hexágono, y nodos deterministas, representados con cuadrados con bordes redondeados. Si un DI sólo contiene nodos de azar, se le llama una red (bayesiana) de creencias. El nodo determinista deriva su nombre del hecho de que el valor que toma es una función del valor que toman sus nodos padres. Las flechas representan relaciones condicionales, excepto en el caso de las flechas punteadas, que representan la información disponible a la hora de tomar la decisión correspondiente. Pearl (2005) es un artículo interesante que aborda los DIs, también desde un punto de vista histórico. El DI visualiza las dependencias entre variables y la información que existe cuando se toman las decisiones, pero no exhibe los posibles valores ligados a cada variable de decisión o de azar. De este modo, es frecuente que un problema se represente primero como un DI y luego se convierta en un árbol de decisiones para computar la solución, lo cual significa determinar la utilidad esperada de cada posible decisión para ser maximizada en el espacio de la toma de decisiones.

En la Figura 1 mostramos un ejemplo de un DI, que representa a un tomador de decisiones D que debe tomar una decisión d basándose en una observación x que depende del estado θ y toma valores de un conjunto Θ. D sufre una pérdida lD(d; θ) que depende de lo que decida y de lo que esté ocurriendo.

Figura 1. Ejemplo de un diagrama de influencia



Redes bayesianas

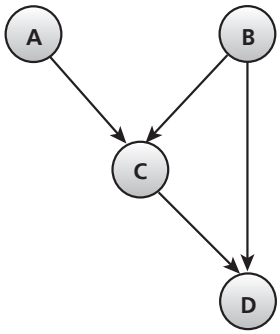
Una red bayesiana (RB) es una representación gráfica de la distribución de probabilidad conjunta de un conjunto de variables, y se utiliza para razonar en una situación de incertidumbre. Una RB representa las variables de interés como nodos, y las dependencias entre variables como arcos. Las variables subyacentes a los nodos de la RB pueden ser continuas, ordenables o categóricas o, alternativamente, las variables continuas pueden discretizarse para facilitar la elicitación y la computación. Los valores booleanos, ordenables, enteros y en rango serían ejemplos de nodos discretos comunes. El número de categorías se suele elegir según el contexto, las inferencias que se desean hacer, la información disponible y la complejidad computacional.

La Figura 2 representa una RB con nodos binarios. Es obvio que esta representación equivale a un grafo acíclico dirigido (DAG) compuesto de nodos que representan las variables de interés, arcos que muestran las influencias directas que existen entre estas variables, tablas de probabilidad a priori para aquellos nodos que no tengan padres, y tablas de probabilidad conjunta (TBCs) para los demás nodos,

Para ser más precisos, en un grafo dirigido acíclico que se obtiene de G = (V; E), donde V es el conjunto de nodos y E es el conjunto de vínculos dirigidos entre nodos, una distribución de probabilidad conjunta P(XV) sobre el conjunto de variables XV se puede factorizar como P(XV) =∏vєV P(XV|Xpa(v)) donde Xpa(v) es el conjunto de las variables padres de la variable XV para cada nodo v en V. Esto nos proporciona la propiedad que define a una RB, eso es, que la distribución conjunta de un nodo solo está condicionada por los padres de ese nodo.

La representación explícita de la estructura de independencia condicional impone ciertas limitaciones al modelo que reducen enormemente el número de parámetros a calcular. En la red representada en la Figura 2(a) con los estados listados en la Figura 2(b), la tabla de probabilidad conjunta que se obtiene de P(A,B,C,D) tendría 24 = 16 entradas. Sin embargo, las limitaciones que derivan de la estructura de independencia condicional de la Figura 2 (a) conllevan a P(A,B,C,D) = P(A)P(B) P(C|A,B)P(D|B|C), lo cual solo requiere 1+1+4+4=10 probabilidades.

Figura 2. Un ejemplo de una red bayesiana, con 4 nodos de interés, (b) los estados de cada variable y (c) las tablas de probabilidad condicional subyacentes al nodo C, debido a los nodos A y B



(a)

Node	States
A	true, false
B	true, false
C	true, false
D	true, false

(b)

	A = true		A = false	
	B = true	B = false	B = true	B = false
C = true	0.45	0.2	0.6	0.77
C = false	0.55	0.8	0.4	0.23

(c)

Nodo Estados

A = verdadero. A = falso

La terminología utilizada en la literatura de RBs e ilustrada en la Figura 2(a) es la siguiente. En primer lugar, un nodo es el padre de un hijo si un arco va desde el primero al segundo; por ejemplo, los nodos A y B son los padres de C y los nodos C y D son los hijos de B. En segundo lugar, donde existe una cadena dirigida de nodos, un nodo es el ancestro de otro si está posicionado anteriormente en la cadena, y es el descendiente de otro nodo si está posicionado posteriormente en la cadena; por ejemplo, el nodo D desciende de A. En tercer lugar, un nodo sin padres es un nodo raíz, por ejemplo, el nodo A. En último lugar, un nodo sin hijos es un nodo hoja, por ejemplo, el nodo D.

La Figura 2(C) es un ejemplo típico de una tabla de probabilidad condicional, que un experto ha de especificar para cada nodo y para cada combinación de sus valores y aquellos de sus padres. En la Figura 2 se analiza al nodo C y a sus padres A y B. Para simplificar, un nodo puede tomar dos valores posibles, a los cuales llamamos

"True" (Verdadero) y "False" (Falso). Para cada combinación de valores a y b de A y B respectivamente, la probabilidad condicional P(C=T| A=a, B=b) debería especificarse. A modo de ejemplo en la Figura 2(c), tenemos P(C=T| A=T, B=T)= 0.45, y, consecuentemente, P(C=F| A=T, B=T)= 0.55.

La visualización de una RB también permite reducir el tiempo que lleva computar las probabilidades marginales, la operación más común que se realiza en una RB. Cuando se obtiene nueva información, se pueden actualizar las creencias de forma sencilla. Una introducción útil a las RBs sería Jensen y Nielsen (2007), mientras que Hugin, GeNie y Netica son los *softwares* más utilizados para modelar RBs.

Como ejemplo relacionado a la Figura 2, uno podría interesarse en la probabilidad (marginal) P(C = T), donde T significa "True" y F "False". Obtenemos

P(C=T)=P(C=T|A=T,B=T)P(A=T)P(B=T)+P(C=T|A=T,B=F)P(A=T)P(B=F)+P(C=T|A=F,B=T)P(A=F)P(B=T)+
+ P(C=T|A=F,B=F)P(A=F)P(B=F).

Supongamos que P(A=T)=0.3 (and P(A=F)=0.7) y P(B=T)=0.6 (and P(B=F)=0.4), entonces:

P(C=T)=0.45*0.3*0.6+0.2*0.3*0.4+0.6*0.7*0.6+0.77*0.7*0.4=0.5726.

Es evidente que las computaciones de las probabilidades marginales de los nodos (que podrían representar sucesos en un escenario futuro) son muy sencillas una vez que el problema ha sido estructurado en forma de una cadena de dependencia en la cual se especifican las probabilidades condicionales de cada nodo dependiendo de sus padres, como se hace en las RBs.

La modelización de la incertidumbre

La incertidumbre aparece en cuanto empezamos a recolectar información. Por eso mismo, es esencial intervenir en la forma en la que se recolecta la información, teniendo en cuenta los protocolos de elicitación de creencias. La clave de representar la incertidumbre en un modelo de probabilidad son las estructuras de dependencia/independencia que se modelan. Si todas las variables se representan como variables independientes, no existe la posibilidad de aprender de los datos y, lo que es más importante, es poco probable que el modelo represente las complejas interacciones

del mundo. Asimismo, si el modelo permite dependencias entre todas las variables, es probable que su complejidad conlleve a que sea computacionalmente intratable. Por estas razones, es importante encontrar un equilibrio. Muchos formalismos, tales como los diagramas de influencia y las redes bayesianas, han sido desarrollados para representar dependencias clave de forma computacionalmente tratable, y que permita que la estructura de dependencia sea eficazmente elicitada de los expertos. Estas técnicas de modelización de dependencias tienen la ventaja de ser gráficas, resultando en interfaces gráficos intuitivos en los *softwares* de modelización. De esta manera, esperamos que estos formalismos sean esenciales a la hora de explicar a todas las partes interesadas los modelos en los cuales se basan los análisis en cuestión.

La importancia acordada al uso de las opiniones de expertos es una de las características de la estadística bayesiana. El nivel de interés en el método bayesiano está creciendo, no sólo entre matemáticos de estadística, sino también entre científicos y expertos de distintos ámbitos. Una de las razones que explican el nivel de interés que atrae este método es la posibilidad de hacer inferencias o pronósticos que no se basan puramente en datos sacados de experimentos estadísticos, sino también en los conocimientos de expertos. Ambas fuentes de información se combinan formalmente mediante el Teorema de Bayes. El típico método bayesiano combina distribuciones a priori y funciones de pérdidas para estimar parámetros, testar hipótesis y hacer pronósticos sobre futuras observaciones. Si tomamos un modelo descrito mediante la variable aleatoria X, con una densidad $f(x \mid \theta)$, el experto proporciona información que se convierte en una distribución a priori $\pi(\theta)$ del parámetro θ que, combinado con una muestra $X=(X1, \ldots, Xn)$ de X, conlleva a una distribución a posteriori $\pi(\theta \mid X)$ a través del Teorema de Bayes

$$\pi(\theta \mid X) = f(X \mid \theta)\, \pi(\theta) / f(X), \text{ donde } f(X)= \int f(X \mid \theta)\, \pi(\theta)\, d\, \theta.$$

Las inferencias y los pronósticos se basan en la distribución a posteriori; el parámetro θ, en particular, se estima a base de especificar una función de pérdidas $L(\theta, a)$ y elegir un valor de a que minimice las pérdidas posteriores esperadas.

Como ejemplo, podemos evaluar el caso de la durabilidad de un objeto, modelado con una distribución exponencial; es decir, la densidad de la durabilidad de un objeto (por ejemplo, una bombilla) la proporciona $f(x \mid \theta) = \theta \exp(-\theta x)$. Supongamos que se observan n objetos que se suponen iguales y utilizados en iguales condiciones; este contexto nos llevaría a la noción estadística de las variables aleatorias independientes

e idénticamente distribuidas. Este conjunto de observaciones lo mostramos como X=(X1, …, Xn). Se elige una distribución gamma $\pi(\theta) = \beta\alpha\ \theta\alpha\text{-}1\ \exp(-\beta\ \theta)/\ \Gamma(\alpha)$ como distribución a priori, de forma que el Teorema de Bayes nos lleve a una distribución a posteriori que siga siendo una distribución gamma $\pi(\theta) = (\beta+\sum Xi)\alpha+n\ \ \theta\alpha+n\text{-}1\ \exp(-(\beta+\sum Xi)\ \theta)/\ \Gamma(\alpha+n)$, mientras que el estimador del parametro θ nos lo proporciona $(\alpha+n)/(\beta+\sum Xi)$ (en jerga estadística, es el estimador óptimo bajo una función de pérdidas al cuadrado). Se puede observar cómo se obtiene el estimador combinando la opinión del experto (cuantificada utilizando los parámetros α y β) con los datos (n a y $+\sum Xi$).

Antes de seleccionar una distribución del parámetro θ, es necesario llevar a cabo la elicitación de opiniones expertas. Para ello, se han propuesto distintos métodos, desde la especificación directa de las distribuciones a priori, hasta las atribuciones por azar y los juicios cualitativos, que se transforman en valores cuantitativos utilizando métodos controvertidos tales como el Proceso de Análisis Jerárquico. Las opiniones de expertos conllevan, sea de forma directa o indirecta, a la especificación de algunas características de la distribución *a priori*, por ejemplo, los momentos (en general, la media y la varianza) o los cuantiles. Basándose en estas características, el estadista selecciona una forma funcional adecuada para la distribución *a priori* y sus hiperparámetros, que se ajuste mejor a los valores evaluados por el experto. Le elección de la distribución *a priori* solía basarse en razones matemáticas, tales como la necesidad de obtener las distribuciones *a posteriori* de formas funcionales ya conocidas; un ejemplo típico de esto serían las distribuciones a priori conjugadas, cuyas distribuciones *a priori* y *a posteriori* tienen la misma forma (como las distribuciones gamma cuando el modelo es exponencial o de Poisson). Hoy en día, la existencia de ordenadores de alto rendimiento y, sobre todo, de técnicas de simulación (las más populares serían MCMC, las cadenas de Markov, y Montecarlo) permiten la elección de casi cualquier distribución *a priori.* Hay muchos libros que explican el método bayesiano. Una introducción recomendable al tema sería Bolstad y Curran (2016).

Análisis de sensibilidad

El análisis de sensibilidad (AS), en su sentido más amplio, es el estudio de la variación en el resultado de un modelo que puede distribuirse, cualitativa o cuantitativamente, en distintas fuentes de variación, y de cómo un modelo en concreto depende de la información en la cual se basa. La modelización bien practicada requiere que el responsable de la misma proporcione su evaluación de confianza en el modelo,

posiblemente a base de evaluar las incertidumbres del proceso de modelización y del resultado del propio modelo. El análisis de sensibilidad bayesiano busca superar la principal objeción al análisis bayesiano, que es su dependencia de aportaciones subjetivas, sobre todo en lo que se refiere a la distribución a priori y la pérdida. Su propósito es determinar el impacto que tienen las aportaciones a un análisis bayesiano (eso es, la distribución *a priori*, la pérdida y el modelo) en su resultado cuando estas aportaciones se encuentran en ciertas clases. Si el impacto es considerable, esto significa que el modelo es sensible y que habría que refinar la información disponible. Una forma de hacer esto sería añadir límites adicionales a las clases y/o obtener datos adicionales. Si el impacto no es considerable, significa que el modelo es robusto y que no se requiere más información o un análisis más detallado. La sensibilidad bayesiana se considera un elemento clave del análisis bayesiano.

También surgen problemas de sensibilidad cuando lidiamos con diagramas de influencia y redes bayesianas. Una forma práctica de realizar un análisis de sensibilidad a las redes bayesianas es cambiar las probabilidades en cada nodo y ver cómo esto afecta el valor de las probabilidades marginales en los nodos que nos interesan, especialmente los nodos hoja (como el nodo D en la Figura 2).

Cogiendo el ejemplo anterior de la computación de la probabilidad marginal P(C=T) en la RB de la Figura 2, supongamos que hay incertidumbre en la especificación de P(A=T) y que P(A=T)=0.4 podría ser una alternativa a 0.3. En ese caso, obtendríamos P(C=T)= 0.5408, que no se distingue en sensibilidad al resultado anterior (0.5726). Si hubiésemos considerado la posibilidad de P(A=T)=0.8 habríamos obtenido P(C=T)=0.4136. El análisis de sensibilidad daría un resultado de robustez en el primer caso, donde la diferencia en creencias no afecta al nodo C, mientras que la falta de robustez conllevaría a evaluaciones completamente distintas de la ocurrencia del nodo C. Este último resultado no denota un error en el procedimiento matemático, sino que formaliza un problema relacionado a las distintas opiniones de los expertos, que podría resolverse buscando un consenso o simplemente dejarse sin resolver, dejando espacio para distintas evaluaciones. Para un resumen detallado del análisis de sensibilidad bayesiano, pueden consultar Ríos Insúa y Ruggeri (2000).

Conclusiones

En este artículo nos hemos concentrado en el papel de los analistas, es decir, en cómo pueden estructurar un problema complejo utilizando herramientas gráficas tales como los diagramas de influencia para modelar los flujos del proceso de toma

de decisiones, así como las redes bayesianas para modelar la dependencia entre distintos factores y su relación probabilística. También hemos explicado cómo la estadística bayesiana nos permite formalizar el uso de dos fuentes relevantes de información: los datos sacados de observaciones y el conocimiento de los expertos. También hemos ilustrado cómo abordar los problemas de sensibilidad creados por la diversidad de creencias. Nos gustaría añadir unas pocas palabras sobre la forma en la que los tomadores de decisiones (TDs) podrían utilizar los instrumentos brevemente expuestos en este artículo. Los TDs podrían utilizar y, posiblemente, agradecer las representaciones gráficas del proceso de toma de decisiones que les proporcionan los diagramas de influencia: estos presentan, de forma simplificada, los distintos aspectos del proceso, incluyendo las aportaciones cuyo valor se conoce, aquellas que son aleatorias, la relación entre entidades (cuál influye a cuál), y las consecuencias de las decisiones tomadas por los TDs. Los TDs pueden utilizar las redes bayesianas para explorar distintos escenarios futuros. En el ejemplo que nos proporciona la Figura 2, diversos escenarios futuros podrían conllevar a distintos valores de, por ejemplo, P(A=T) con consecuencias para, por ejemplo, P(C) y P(D), similares a aquellas que discutimos en la sección sobre el análisis de sensibilidades. Por último, la estadística bayesiana podría permitir que los tomadores de decisiones expresen sus creencias y (en el contexto del análisis de decisiones bayesiano) sus preferencias, aunque este proceso requeriría la cooperación de un técnico estadístico, ya que el proceso de elicitar la distribución a priori y las pérdidas (es decir, los instrumentos que se utilizan para expresar las creencias y las preferencias, respectivamente) pueden ser bastante complejos.

Bibliografía

Bolstad, W.M. and Curran, J.M. (2016), Introduction to Bayesian Statistics, Third Edition, Wiley, Chichester.

Jensen, F.V. and Nielsen, T.D. (2007), Bayesian Networks and Decision Graphs, Springer-Verlag, New York.

Pearl, J. (2005) Influence Diagrams—Historical and Personal Perspectives, Decision Analysis, v. 2, pp. 183-244.

Rios Insua, D. and Ruggeri, F. Eds. (2000), Robust Bayesian Analysis, Springer-Verlag, New York.

Fabrizio Ruggeri es director de Investigación en el CNR-IMATI, Milán, Italia, Profesor Adjunto en QUT, Brisbane, Australia, y miembro de la facultad en varios programas de doctorado (Matemáticas, Pavia-Milano Bicocca, Italia, y Estadística, Valparaíso, Chile). Sus principales intereses son la estadística bayesiana y la estadística industrial. Es vicepresidente del International Statistica Institute, antiguo presidente de la European Network for Business and Industrial Statistics y de la International Society for Bayesian Analysis, y presidente electo de la International Society for Business and Industrial Statistics. Es autor de más de cien artículos y de cinco libros. Es redactor jefe de Applied Stochastic Models in Business and Industry y de Wiley StatsRef. Es miembro de la ISBA y de la American Statistical Association, y receptor de la Medalla Zellner.

Teoría de juegos y análisis de riesgos adversarios para predicción de acciones de adversarios

Jesús Ríos y David Ríos Insúa

Resumen

Numerosas cuestiones relativas a inteligencia geopolítica se refieren a predecir las decisiones de otras organizaciones que las toman para su propio beneficio con propósito bien definido y carácter adaptativo. La Teoría de Juegos y el Análisis de Riesgos Adversarios aportan conceptos y métodos para tratar situaciones de este tipo. En este capítulo se hace una breve revisión comparativa de ambas perspectivas y se presenta cómo pueden emplearse en problemas de inteligencia competitiva.

Introducción

Numerosas cuestiones relativas a inteligencia geopolítica se refieren a predecir las decisiones de otras organizaciones que las toman para su propio beneficio con un propósito bien definido y carácter adaptativo. Por ejemplo, en el estudio presentado en Chen et al (2016) sobre extracción de información geopolítica usando combinación de opiniones cerca del 30% de las preguntas planteadas involucraban a adversarios de alguna manera, siendo un ejemplo del tipo de preguntas allí planteadas: *¿Empleará Siria armas químicas o biológicas antes de enero de 2013?* (pregunta planteada en 2011 y 2012).

Las técnicas estructuradas de juicios de expertos, véanse los artículos de Chen, Salo y Ruggeri en este libro o la introducción de Cooke (1991) y los ejemplos de uso en Goossens *et al* (1998), pueden utilizarse con éxito en estudios de inteligencia geopolítica y, más generalmente, en la elaboración de análisis de riesgos y de decisiones (Bedford y Cooke, 2001). Un rasgo esencial de estas disciplinas es su énfasis en descomponer problemas complejos en piezas más pequeñas de más fácil comprensión que luego se recombinan para resolver el problema inicial. Se simplifican así las tareas cognitivas complejas y se mitiga el uso de heurísticas de razonamiento que introducen sesgos sistemáticos en juicios de valor y creencias. La descomposición típicamente conlleva más juicios, aunque éstos tienden a ser más sencillos y con mayor contenido, por lo que conducen a mejores predicciones y decisiones una vez combinados para producir una respuesta conjunta al problema de predicción o de decisión.

Sin embargo, tales metodologías estándar no tienen en cuenta los rasgos específicos de carácter estratégico de los problemas en que se necesita la predicción de acciones de adversarios inteligentes. Como se suele decir, no es lo mismo predecir la evolución de un huracán que las acciones de un terrorista. Aquí describimos una perspectiva que permite centrarse en este tipo de predicciones basada en ideas de la Teoría de Juegos y del Análisis de Riesgos Adversarios (ARA). El ARA se origina a partir de la observación de que las condiciones de conocimiento común de la Teoría de Juegos (Hargreaves-Heap y Varoufakis, 1995) no se dan en aplicaciones como la lucha contra el terrorismo y la ciberseguridad, pues los competidores tienden a ocultar información. El ARA propone la descomposición del problema de decisión con adversarios creando no sólo un modelo de decisión para el defensor, sino también modelos explícitos de cómo los adversarios razonan para decidir su estrategia de ataque. Hasta aquí el ARA y la Teoría de Juegos proceden de la misma manera. Sin embargo, en esta última disciplina se presupone que todos los modelos de decisión, tanto del defensor como de los atacantes, son conocimiento común para los agentes participantes. Es decir, no sólo todos los decisores saben qué modelos emplea cada decisor para elegir su estrategia, sino también todos saben que los otros decisores saben que cada decisor dispone de tal conocimiento. El ARA debilita esta hipótesis considerando que los modelos de cada decisor son subjetivos, desde el punto de vista del defensor, y la falta de conocimiento total o incertidumbre sobre los modelos de decisión de los adversarios se traduce en distribuciones de probabilidad predictivas sobre la acción que cada adversario elegirá.

A continuación, mostramos la diferencia entre el ARA y la Teoría de Juegos a través de un sencillo ejemplo numérico.

Teoría de Juegos

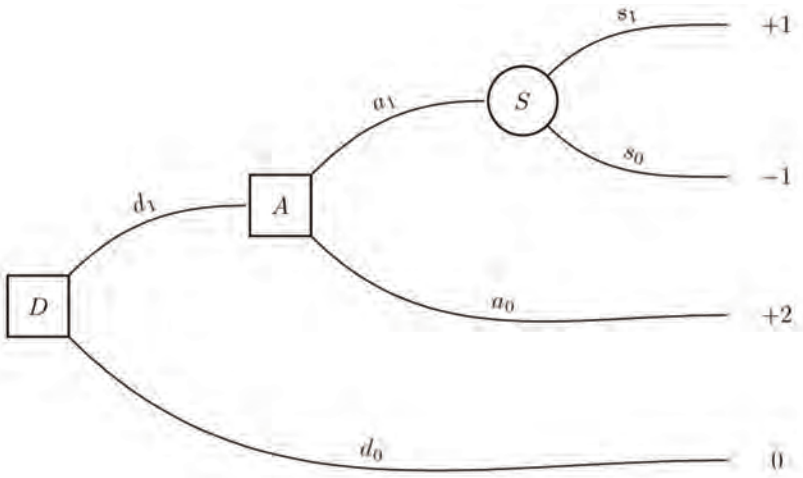
Para simplificar la discusión, nos centraremos en situaciones secuenciales de Defensa-Ataque: un defensor (la organización a la que aconsejamos) D toma primero su decisión *d*; un atacante (la organización de la que nos protegemos) A observa *d* antes de tomar su decisión y responde seleccionando su mejor acción de ataque a contra *d* (incluyendo la posibilidad de no atacar). El resultado será generalmente no determinista y dependerá de *d* y *a*.

Como ejemplo, consideremos un país D, que tiene la opción de intervenir en otro país A con el objetivo de cambiar el actual régimen de gobierno, lo que, a la larga, beneficiaría al defensor. Un análisis de los costes y beneficios esperados asociados

con dicha opción ayudará al defensor a decidir si dicha intervención merece la pena. Así el defensor dispone en este problema de decisión de dos opciones: *d1* (intervenir) y *d0* (no intervenir). El país A puede responder a dicha intervención con una de dos posibles estrategias: someterse a la intervención del país D sin ofrecer resistencia (*a0*); o resistirse con fuerza (*a1*), en cuyo caso existirá la incertidumbre de si la resistencia será suficiente para frenar la intervención de D. Para simplificar, supongamos que la única incertidumbre S en este problema de decisión lleva asociada sólo dos posibles resultados: el éxito (*s0*) o no (*s1*) de la resistencia por parte de A, si D decide intervenir. El problema se describe en la Figura 1 como un árbol de juego.

Cada rama del árbol representa un escenario al que se llega dependiendo de las decisiones tomadas por los países D y A, así como la posible realización de la incertidumbre S. El defensor evaluará las consecuencias relevantes asociadas con cada una de dichas ramas usando, por ejemplo, un análisis de costes y beneficios o, más en general, a través de la asignación de su función de utilidad multi-atributo (Keeney y Raiffa, 1993). Así, el escenario de *statu quo*, consecuencia de no intervenir (*d0*) es evaluado con 0 unidades de utilidad por el defensor. El escenario más deseado para el defensor es aquel en que A no opone resistencia (*a0*) a una intervención (*d1*) y se evalúa con +2 unidades de utilidad. En el caso en que A se resistiese (d1, a1) son posibles dos escenarios: la intervención resulta en un cambio de régimen de gobierno tras imponer medidas de fuerza (*s1*), valorado con +1 unidad de utilidad dado el esfuerzo extra que el país D tiene que realizar; o el país A es capaz de resistirse al intento de imponer un cambio de régimen de gobierno (*s0*), que es el resultado menos deseado por el defensor con -1 unidades de utilidad.

Figura 1. Árbol de juego



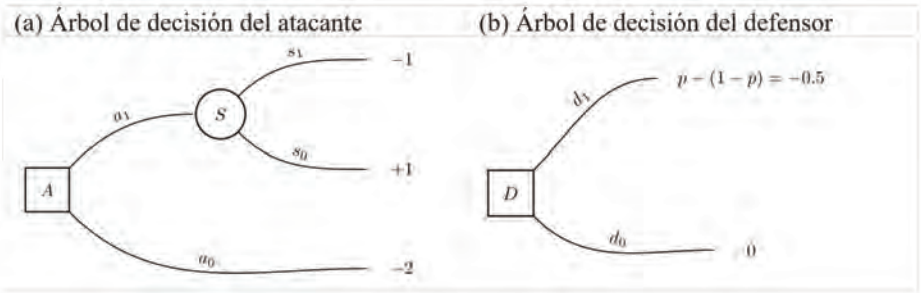
La Teoría de Juegos resuelve a la vez este problema de decisión para el defensor y el atacante. Para ello, supone que ambos agentes buscan maximizar simultáneamente su utilidad esperada y que los problemas de decisión de cada agente son *conocimiento común*. En particular se supone que ambos agentes conocen las opciones disponibles y las utilidades del otro para cada una de sus decisiones y consecuencias. Una forma típica[31] en Teoría de Juegos de llegar a esto es suponer que el juego es de *suma-cero*, es decir, que las utilidades del defensor y atacante suman cero. Esto es equivalente a suponer que las utilidades del atacante son opuestas a las del defensor en signo y su maximización por parte del atacante es equivalente a la minimización de las utilidades del defensor. Con respecto a la única incertidumbre S de este problema, se supone que hay una probabilidad objetiva[32] asociada a la misma. Supongamos que p=0.25 es la probabilidad de que la intervención no falle, pese a la resistencia de A: escenario (*d1, a1, s1*). Esta probabilidad podría justificarse si su estimación se basase, por ejemplo, en la tasa de éxito de intervenciones similares.

31 Los juegos no han de ser de suma-cero. Por sencillez adoptamos esa hipótesis, pero la discusión que damos se extiende al caso general.

32 En el sentido de que es común y conocida por ambos. En el caso de información incompleta, esta idea se aplica a la denominada distribución de tipos.

La solución se calcula resolviendo hacia atrás el árbol de juego (French y Ríos Insua, 2000). Este procedimiento implícitamente resuelve primero el problema de decisión del atacante mostrado en la Figura 2a. La acción de máxima utilidad esperada para el atacante es a1 pues éste obtendría una utilidad esperada de (-1) *p* + (1) (1-*p*) que es mayor que la utilidad de -2 asociada con *a0*, cualquiera que sea la probabilidad Pr(*S*=*s1*| *d1, a1*) = p. Bajo las hipótesis de la Teoría de Juegos, el defensor podría anticipar con absoluta certidumbre la respuesta del atacante a una posible intervención (*d1*). Así, sabiendo esto, el defensor resolvería el problema de decisión de la figura 2b, eligiendo no intervenir (*d0*) que le proporciona 0 unidades de utilidad; en comparación con *2p* -1 = -0.5, la utilidad esperada que obtendría si eligiese *d1*.

Figura 2. Árboles de decisión del defensor y atacante considerados en la solución de Teoría de Juegos



Análisis de Riesgos Adversarios

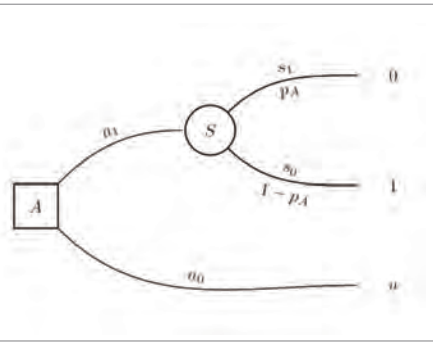
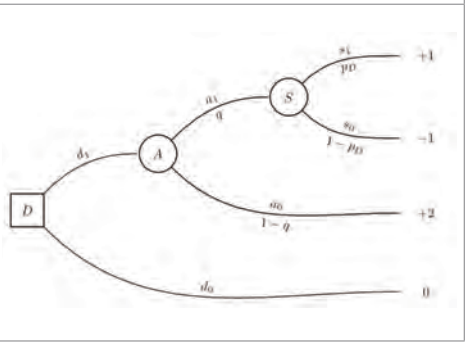
En ARA, el objetivo es ayudar a uno de los agentes (el defensor) a resolver su problema de toma de decisiones. Para ello, se necesitará predecir las acciones de los otros agentes intervinientes. Este problema se puede ver como un problema de decisión bayesiano, donde las acciones de los otros agentes llevarían asociada cierta probabilidad de ocurrencia, con lo que el problema del decisor se centra en la búsqueda de la acción de máxima utilidad esperada con respecto las distribuciones de probabilidad sobre las acciones de los otros agentes. Esta perspectiva fue propuesta, de manera no constructiva, en Kadane y Larkey (1982) y Raiffa (1968, 1982). Dada su falta de sentido práctico, la perspectiva bayesiana a los juegos fue criticada por Harsanyi (1982) y Myerson (1991), entre otros. Por ejemplo, las probabilidades asociadas con las acciones de los otros agentes representan el grado de creencia del decisor (defensor) y no está claro desde un punto de vista práctico cómo incorporar elementos estratégicos de teoría de juegos en la estimación de dichas probabilidades. Así, el principal obstáculo en la implementación de la perspectiva bayesiana en estos problemas es la falta de mecanismos explícitos que permitan codificar dichas probabilidades.

ARA proporciona una metodología para resolver dicha dificultad mediante la construcción explícita de un modelo para la toma de decisiones de su oponente. Dicho modelo incorporará incertidumbre sobre las creencias y valores del adversario a través de distribuciones de probabilidad, induciendo una distribución de probabilidad sobre las acciones del mismo. El ARA, entonces, no necesita recurrir a la hipótesis de *conocimiento común* para crear un modelo predictivo sobre las acciones del adversario. Esto es esencial en casos donde no es realista suponer que los agentes estén dispuestos a compartir información privada sobre sus problemas de decisión. En nuestro ejemplo, el defensor (país D) no tiene por qué tener acceso a las auténticas utilidades del atacante (país A) ni a sus creencias sobre la probabilidad p de que la resistencia a una posible intervención falle, aunque puede tener acceso a fuentes de información más o menos fiables.

El objetivo del ARA es generar modelos más realistas que los de la Teoría de Juegos para la predicción de acciones de adversarios no basándonos en conocimiento común. Para ello, el defensor intentará ponerse en la piel del atacante y resolver su problema de decisión. Por ejemplo, el problema del atacante podría ser visto a través del árbol de decisión de la Figura 3a. Como el defensor no sabe las utilidades y probabilidad que el atacante usará para resolver su problema, tendremos que estimar éstas a través de distribuciones de probabilidad que representarán la falta

de conocimiento total sobre ellas por parte del defensor y modelizan la posible información disponible sobre el atacante por medios directos e indirectos. Así, el defensor utilizará estas estimaciones para resolver el problema del atacante desde su punto de vista, obteniendo una distribución de probabilidad sobre los posibles ataques.

Figura 3. Modelos de decisión para predecir al adversario y ayudar a decidir al defensor en la solución de ARA

(a) Árbol de decisión del atacante desde el punto de vista del defensor	(b) Árbol de decisión del defensor
	

En concreto, son posibles tres escenarios para el atacante (resistirse sin éxito, hacerlo con éxito y no resistirse) como se muestra en la Figura 3a. Supongamos, por ejemplo, que el defensor estima que el orden de deseabilidad desde el punto de vista del país A (atacante) entre las consecuencias de estos escenarios es: (*d1,a1,s0*) > (*d1,a0*) > (*d1,a1,s1*). Es decir, para el país A resistirse a una intervención (*a1*) es una acción más costosa que *a0* pero potencialmente podría producir el resultado más favorable si es exitosa (*s0*); sin embargo, una resistencia que resulta en fracaso (*s1*) sería lo menos deseable pues el coste extra de resistirse no produciría beneficios para el país mejores que *a0*. Sin pérdida de generalidad, normalizamos la utilidad del atacante (país A) asignando un valor de utilidad de 1 a sus mejores consecuencias y *0* a sus peores, como se muestra en la Figura 3a. Dada esta normalización, el defensor sólo tendrá que estimar la utilidad u que el atacante asignaría a las consecuencias resultado del escenario (*d1,a0*).

Como el defensor no tiene acceso directo a la utilidad *u* del atacante y, por lo tanto, no podrá estimar su valor exacto con total certidumbre, éste tendrá que

codificar sus creencias (O'Hagan et al, 2006) acerca de u mediante una distribución de probabilidad subjetiva *U* con soporte en [0,1], pues 0<*u*<1 está normalizada. Esto define la función de utilidad aleatoria del atacante. Claramente, estamos relajando la utilización de la hipótesis de suma-cero que es representativa de cómo la Teoría de Juegos ingenuamente estima las utilidades del adversario, lo que nos permite generar modelos de predicción de acciones adversarias más realistas. Obsérvese como bajo esta hipótesis de suma-cero el orden de preferencia entre escenarios asignado al país A supone que (*d1,a0*) < (*d1,a1,s1*), es decir, las consecuencias de resistirse sin éxito serían más deseables para el país A que las de no resistirse, pese al coste extra, lo cual no parece muy razonable.

Más aún, el paradigma bayesiano permite distinguir el caso en que defensor y atacante perciben con diferentes probabilidades el resultado *S* de resistencia a una intervención. Estas probabilidades representan creencias de los agentes y son, por tanto, subjetivas. En comparación, la Teoría de Juegos se limita esencialmente al uso de probabilidades objetivas. Así, ahora usaremos *pA* y *pD* para representar respectivamente las probabilidades del atacante y del defensor de que *S*=*s1* (la resistencia fracase) basadas en sus respectivas evaluaciones. Como el defensor no sabrá con exactitud el valor exacto de la probabilidad *pA* del atacante, asignará una distribución de probabilidad *PA* sobre dicho valor, que define la probabilidad aleatoria del atacante (desde la perspectiva del defensor).

Si el defensor cree que el atacante busca maximizar su utilidad esperada, éste usará su estimación de la utilidad y probabilidad aleatoria del atacante, *U* y *PA* respectivamente, para predecir la elección del atacante mediante la resolución del árbol de decisión de la Figura 3a desde su perspectiva. La utilidad esperada aleatoria del atacante, asociada con la acción *a1* (resistirse) es *0 PA + 1 (1-PA) = 1-PA*. Mientras que la utilidad aleatoria asociada con *a0* es *U*. Así, el atacante (país *A*) elegirá la acción a1 (resistir) si y sólo si 1-PA > U. Por ejemplo, si el defensor estima que *pA ~ PA* sigue una distribución *beta*(4, 1), codificando su creencia de que el país *A* es muy pesimista en cuando a su posibilidad de resistirse con éxito, y que *u ~ U* sigue una distribución uniforme en (0, 0.6), entonces la probabilidad con que el país *D* cree que el país *A* se resistirá en caso de una intervención es aproximadamente *q = Pr(A = a1) = Pr(1-PA > U)* = 0.33. Obsérvese cómo, en ARA, el defensor en lugar de usar los valores de utilidad y probabilidad del atacante, que no conoce con exactitud, para calcular la acción del atacante que maximiza su utilidad esperada, usará probabilidades y utilidades aleatorias que representan sus creencias sobre tales valores para obtener una distribución de probabilidad predictiva sobre la decisión del atacante.

Finalmente, el defensor usará esta distribución predictiva sobre las acciones del atacante, en nuestro ejemplo q=0.33, para resolver su problema de decisión mostrado en la Figura 3b. Así, si el defensor estima que pD=0.25, su mejor acción será *d1* con una utilidad esperada de $1.18 = q\,[+1\ pD + (-1)\ (1-\ pD)] + (1-q)\,[+2] > 0$ (la utilidad de la acción *d0* para el defensor).

Obsérvese cómo ARA permite al defensor modelizar un atacante cuyas creencias sobre su posibilidad de resistirse con éxito son muy diferentes a las del defensor. En nuestro ejemplo, el defensor cree que no es fácil imponerse al país *A* en caso de resistirse (pD=0.25), pero también considera que el atacante no es consciente de su superioridad y estima que éste dará poca probabilidad a esta posibilidad ($pA \sim PA$ es una beta con valor esperado 0.8). Esto junto con una mejor estimación de las utilidades del atacante, permiten al defensor, en su opinión, predecir como poco probable (q=0.33) que el país *A* se resista a una intervención (*a1*), lo que hace más deseable la intervención (*d1*) para el defensor.

Discusión

En este capítulo hemos introducido con un ejemplo sencillo la metodología ARA centrándonos en situaciones secuenciales de Defensa-Ataque. El ARA se aplica a situaciones más generales. Por ejemplo, situaciones de toma de decisiones simultáneas, o con interacciones más complejas entre ambos agentes, o múltiples agentes. En el ejemplo hemos supuesto que el atacante maximiza su utilidad esperada. El ARA se puede también utilizar para modelizar agentes que emplean principios diferentes al de maximización de utilidad esperada (Ríos Insua et al, 2016), así como en otras situaciones presentadas en Banks et al (2015) y referencias relacionadas referidas a lucha contra dispositivos explosivos improvisados, seguridad urbana, lucha contra drones, protección de infraestructuras críticas, lucha contra piratería, detección de spam o ciberseguridad, entre otros.

Agradecimientos

El trabajo de David Ríos está financiado por los programas MTM2017-86875-C3-1-R del Ministerio de Economía e Innovación de España y la Cátedra AXA-ICMAT sobre Análisis de Riesgos Adversarios. Este trabajo también ha sido en parte fomentado a través del Programa H2020 de Investigación, Desarrollo Tecnológico y Demostración de la UE, bajo el acuerdo no. 740920 (CYBECO).

Bibliografía

D. Banks, J. Ríos, D. Ríos Insua (2015) *Adversarial Risk Analysis*, CRC Press, Boca Raton, FL.

T. Bedford, R.M. Cooke. (2001) *Probabilistic Risk Analysis: Foundations and Methods*, Cambridge University Press, Cambridge, United Kingdom.

E. Chen, D. Budescu, S.K. Lakshmikanth, B. Mellers, P. Tetlock (2016) Validating the Contribution-Weighted Model: Robustness and Cost-Benefit Analyses. *Decision Analysis*, 13, 128-152.

R. M. Cooke (1991) *Experts in Uncertainty: Opinion and Subjective Probability in Science*. Oxford University Press, New York, NY.

S. French, D. Ríos Insua (2000) *Statistical Decision Theory*. Wiley, New York, NY.

L.H.J. Goossens, R.M. Cooke, B.C.P. Kraan (1998) Evaluation of weighting schemes for expert judgement studies. *Proceedings of the Fourth International Conference on Probabilistic Safety Assessment and Management*, 1937-1942.

S. Hargreaves-Heap, Y. Varoufakis (1995) *Game Theory: A Critical Introduction.* Routledge, New York, NY.

J. Harsanyi (1982) Comment: Subjective Probability and the Theory of Games: Comments on Kadane and Larkey's paper, *Management Science*, 28, 120-124.

J.B. Kadane, P.D. Larkey (1982) Subjective probability and the theory of games. *Management Science*, 28(2): 113-120.

R.L. Keeney, H. Raiffa (1993) *Decision Making with Multiple Objectives Preferences and Value Tradeoffs*. Wiley, New York.

R. Myerson (1991) *Game Theory: Analysis of Conflict.* Harvard University Press, Cambridge, MA.

A. O'Hagan, C.E. Buck, A. Daneshkhah, J.R. Eiser, P.H. Garthwaite, D.J Jenkinson, J.E. Oakley, T. Rakow (2006) *Uncertain judgements: eliciting experts' probabilities.* John Wiley & Sons.

H. Raiffa (1968) *Decision Analysis: Introductory Lectures on Choices Under Undertainty.* Addison-Wesley, Menlo Park, CA.

H. Raiffa (1982). *The Art and Science of Negotiation.* Harvard University Press, Cambridge, MA, 1982.

D. Rios Insua, D. Banks, J. Rios (2016) Modeling opponents in adversarial risk analysis. *Risk Analysis*, 36(4), 742-755.

Jesus Ríos se graduó en Ciencias Matemáticas en la Universidad Complutense y obtuvo su doctorado en la Universidad Rey Juan Carlos. Desde el año 2010 el Dr. Ríos es miembro del personal de investigación del laboratorio de investigación de IBM en Nueva York. Su trabajo de investigación se centra en la aplicación de métodos de Inteligencia Artificial para resolver problemas de decisión y predicción. Antes de unirse a IBM, ocupó cargos de investigación en la Universidad de Manchester, Universidad de Luxemburgo, Universidad de Aalborg, SAMSI y Universidad de Concordia (Canadá). Recientemente recibió el premio DeGroot de la International Society of Bayesian Analysis.

David Ríos Insua es AXA-ICMAT Chair en Análisis de Riesgos Adversarios en ICMAT-CSIC y Numerario de la Real Academia de Ciencias Exactas, Físicas y Naturales. Es Catedrático (en excedencia) de Estadística e Investigación Operativa. Previamente ha sido profesor y/o investigador en Manchester, Leeds, Purdue, Duke, SAMSI, CNR-IMATI, Paris-Dauphine, Aalto, IIASA y UPM. Recientemente recibió el premio DeGroot de la International Society of Bayesian Analysis. Es director científico de Aisoy Robotics.

La gestión de riesgos en empresas industriales

Virginia Leal

Resumen

Las actuales condiciones de entorno, cada vez más complejas y cambiantes, han impulsado la implantación de marcos de gestión en las empresas que permiten responder ante la materialización de eventos de riesgo. La implantación de un sistema de gestión de riesgos dota a las organizaciones de mecanismos para la identificación, análisis, valoración y tratamiento de los riesgos de forma coherente con su estrategia y aporta una mayor certidumbre sobre la consecución de los objetivos y el cumplimiento de su visión a largo plazo.

En el caso de las empresas industriales, la gestión de riesgos permite anticiparnos a la ocurrencia de eventos gestionando no solamente sus efectos económicos, sino aquellas consecuencias que pueden afectar la reputación de la organización diseñando estrategias para ganar en legitimidad social, y velar, de forma prioritaria, por la seguridad y salud de las personas incluyendo nuestros proveedores, clientes y las comunidades locales de los entornos donde operamos.

Introducción

La capacidad de seleccionar la alternativa más conveniente en una situación de incertidumbre es una habilidad esencial en el ser humano. Esta capacidad ha permitido a la humanidad evolucionar a lo largo del tiempo en condiciones adversas de escasez de alimento, ausencia de cobijo o falta de protección. La toma de decisiones adecuadas en contextos complejos es uno de los factores que nos ha permitido triunfar como especie.

En esencia, la gestión de riesgos empresarial se fundamenta en esta misma capacidad humana. El riesgo es una variable esencial en cualquier actividad económica y se ha de convivir con él, pues no se puede eliminar por completo. Capturar una oportunidad en busca de un beneficio, sea éste del tipo que sea, siempre exige asumir un determinado nivel de riesgo del que habrá que dimensionar sus consecuencias.

Actualmente la gestión de riesgos es un elemento permanente en la agenda de la alta dirección de las corporaciones. Esto se debe a que el entorno geopolítico,

económico, tecnológico y social es cada vez más complejo y dinámico dando origen a nuevos riesgos o modificando riesgos conocidos que cambian rápidamente en su alcance o consecuencias. Ante esta situación, las empresas buscan formalizar marcos de gestión explícitos que permitan determinar el efecto potencial de estos riesgos sobre sus objetivos de negocio y que garanticen la consecución de la visión de la compañía en un medio/largo plazo. El interés de los equipos directivos por la gestión de riesgos radica en que, además de una herramienta para definir la estrategia de la compañía, es una vía para mejorar sus operaciones, y, en último término, ayuda a asumir con flexibilidad situaciones de estrés saliendo fortalecidos de la experiencia.

Adicionalmente, ciertos actores clave están impulsando la gestión empresarial fundamentada en riesgos. Los inversores, las agencias de *rating* y los reguladores desean tener un mejor entendimiento de cómo las compañías se enfrentan a sus riesgos y qué mecanismos disponen para su gestión en situaciones de estrés, como ha demostrado la crisis sanitaria del COVID-19. En concreto, los inversores quieren tener un conocimiento más profundo del grado de exposición de sus activos financieros en relación con el retorno económico a obtener. Algunas ventajas adicionales que reconoce la comunidad inversora son una distribución más efectiva del capital al racionalizar los recursos económicos dentro la organización o una mejora en la evaluación de las oportunidades. Asimismo, los reguladores quieren garantizar que las industrias dispongan de sistemas que mantienen los riesgos en los niveles de aceptabilidad, sobre todo en aquellos sectores con alta relevancia económica y social, como puede ser el energético.

Marcos de Gestión

Existen diferentes opciones a la hora de definir el marco formal de un sistema de gestión de riesgos. En lo que se refiere a normas internacionales, la mejor opción para una organización no financiera es la ISO 31000. Este modelo proporciona los principios y pautas fundamentales, y al no ser específica de ninguna industria o sector, puede adaptarse a las necesidades de cada organización. Tras la primera emisión de la norma en 2009, ha sido recientemente actualizada en febrero de 2018. Esta nueva versión pone un mayor énfasis en la necesidad de integrar la gestión del riesgo en los procesos de toma de decisión y destaca la importancia de los factores humanos y culturales.

La norma ISO 31000 describe los componentes de un sistema integrado de riesgos, es decir, los principios, el marco y los procesos a aplicar. El marco de gestión se

cristaliza a través de las políticas, los objetivos, los procedimientos, los planes, las responsabilidades, los recursos, los procesos y las actividades necesarias para la gestión de riesgos en todos los niveles de la organización. Estos componentes irán evolucionando con el tiempo a medida que el sistema de gestión madure, pudiéndose alcanzar un elevado grado de sofisticación tanto en herramientas como en métodos. De todos estos elementos, se recomienda en primer término, reflejar en un documento de alto nivel denominado política de gestión de riesgos, la actitud y el compromiso de la organización respecto a la gestión de riesgos. De él, emanarán las directrices generales que se concretarán en el resto de los componentes del sistema.

El proceso de gestión de riesgos aplica los elementos descritos en el marco de la norma ISO 31000 para ejecutar las siguientes actividades (Fraser y Simkins, 2010):

Establecimiento del contexto: Consiste en realizar una revisión profunda de los factores, tanto externos como internos, que pueden generar la aparición de nuevos riesgos o cambios en los mismos. En el análisis del contexto externo se ha de estudiar cualquier elemento fuera de la organización que puede ser fuente de incertidumbre, como nuevas demandas de los grupos de interés, tendencias emergentes en el mercado, nuevos desarrollos tecnológicos, aparición de desarrollos legislativos, cambios culturales o sociales, etc. Los factores relativos a cambios organizacionales, la implantación de nuevos sistemas productivos, las modificaciones en los procesos de gestión o en los sistemas de información, etc. formarán parte de los aspectos a considerar en el análisis de contexto interno.

Apreciación del riesgo: Comprende la identificación, el análisis y la evaluación del riesgo, lo que permite a la organización conocer los riesgos a los que está expuesta y determinar su valor. Para la tarea de identificación de riesgos se pueden emplear diferentes técnicas como el *brainstorming* o lluvia de ideas, el método *what if* o metodologías de análisis de escenarios. Asimismo, existe un amplio espectro de métodos de análisis y valoración de riesgos que abarcan desde técnicas matemáticas cuantitativas hasta la estimación cualitativa por un grupo de expertos (Chapman, 2011). La evaluación del riesgo determinará si el nivel de riesgo es aceptable o no para los niveles marcados en la organización. Los métodos de evaluación son también numerosos y pueden incluir el empleo de matrices de valoración, la definición de criterios estadísticos o el uso de indicadores. La aplicación de unas técnicas u otras dependerá de la experiencia y las capacidades del grupo gestor de riesgos y pueden adaptarse en función de la naturaleza del riesgo.

Tratamiento del riesgo: La gestión de riesgos ha de estar orientada al tratamiento. Si la fase de apreciación culmina con el diagnóstico de que el nivel de riesgo no

es aceptable, la organización debe impulsar medidas para reducir las posibles consecuencias negativas en caso de materialización del riesgo o reducir su probabilidad de ocurrencia. Existen cuatro estrategias de tratamiento del riesgo denominadas evitar, transferir, mitigar y aceptar. Evitar el riesgo significa eliminar el riesgo decidiendo no iniciar o continuar la actividad que lo origina; por ejemplo, limitar la inversión en activos situados en países alto riesgo evita el lucro cesante por interrupción de operaciones debido a la ocurrencia de guerras o conflictos sociales. La segunda de las estrategias se denomina transferir y consiste en traspasar o compartir el riesgo con otras partes; contratar un seguro para que, en el caso de que se produzca el hecho no deseado, el impacto económico para la compañía se atenúe o compartir la gestión con socios en negocios con alta incertidumbre son medidas categorizadas dentro de este epígrafe. La estrategia mitigar consiste en adoptar medidas para modificar el valor del riesgo y pueden ser muy diversas, como la contratación de personal especializado, la ejecución de inversiones o la implantación de nuevos procesos; la mayoría de las acciones de tratamiento del riesgo se encuadran dentro de esta categoría y deben priorizarse en función de criterios de reducción de riesgo y coste de la acción mitigante. La última de las estrategias se denomina aceptar. Consiste en mantener el nivel de riesgo tras una decisión informada y consensuada dentro de la organización, siempre y cuando su nivel se encuentre dentro de los límites tolerables.

Una de las claves en la implantación de un sistema de gestión de riesgos es la correcta aplicación del principio de responsabilidad diferenciada. El modelo formal más difundido es el denominado "de 3 líneas de defensa", ampliamente implantado en distintos sectores y que está recomendado por COSO (*Committee of Sponsoring Organizations of the Treadway*), el Instituto de Auditores Internos, el Comité de Basilea y multitud de instituciones de supervisión y regulación. El modelo identifica tres líneas fundamentales, la línea operativa, las funciones especialistas y auditoría interna y reconoce la importancia de la supervisión del consejo de administración y del comité de dirección de la empresa en la gestión de riesgos. La línea operativa en una empresa industrial está compuesta por aquellas áreas responsables del día a día de los riesgos en los activos productivos o las actividades comerciales. Las funciones especialistas son generalmente áreas corporativas y gestionan de forma específica un determinado tipo de riesgo, como puede ser los de seguridad y medioambiente, los financieros o los de naturaleza legal dando directrices al resto de la organización. Finalmente, la función de auditoría interna vela por la razonabilidad y suficiencia del diseño y evalúa el funcionamiento de los sistemas de control y gestión de riesgos de la organización.

Una de las primeras etapas en la implantación del proceso de gestión es la construcción de una taxonomía que permita la categorización de los riesgos identificados en la organización. Este ejercicio de clasificación aterriza en factores concretos el potencial universo de riesgos y dota a la empresa de un lenguaje común que facilita la comunicación entre las distintas partes involucradas. Disponer de un listado estructurado dividido en elementos y atributos permite, por ejemplo, sistematizar y agilizar los procesos de identificación y resulta de gran utilidad en las actividades de información a la dirección. Por último, es importante destacar que la taxonomía suele ser el primer pilar para la designación de los denominados "*risk owners*" o dueños de riesgos, es decir, aquellos puestos organizativos o áreas que serán responsables de la gestión del riesgo identificado.

Una forma sencilla de categorizar los riesgos en un primer nivel sería la siguiente:

Riesgos estratégicos: Riesgos asociados con los objetivos clave de largo plazo y con la dirección del negocio. Estos riesgos pueden surgir de las acciones de otros participantes en el mercado, de las oportunidades elegidas y de las decisiones tomadas por el propio negocio.

Riesgos operacionales: Riesgos vinculados con las operaciones habituales llevadas a cabo en el desarrollo del modelo de negocio. Incluye los riesgos relacionados con los procesos operativos que van desde el desarrollo de productos y servicios hasta la producción y la comercialización, incluyendo relacionados con los recursos humanos y organizativos, los sistemas de información, sistemas de aprovisionamiento, etc.

Riesgos financieros: Riesgos relacionados específicamente con los procesos, técnicas e instrumentos utilizados para la gestión de las finanzas de la compañía, así como aquellos procesos que suponen un mantenimiento efectivo de las relaciones financieras con los clientes.

Cada uno de estos epígrafes puede ser subdivido en otros elementos alcanzando la granularidad necesaria para cubrir todas las actividades y procesos de la compañía, así como describir todos los potenciales factores de entorno que pueden afectar a la organización. Por ejemplo, dentro de los factores operacionales comúnmente se incluyen los relacionados con la cadena de aprovisionamiento, la operación de los sistemas productivos y los procesos comerciales o los aspectos de cumplimiento normativo o regulatorio. La taxonomía de riesgos ha de ser un documento vivo que ha de ser revisado periódicamente y debe adaptarse a las modificaciones en los modelos de negocio o a los cambios de entorno para poder reflejar adecuadamente los riesgos a los que debe enfrentarse una organización. Por esto, es recomendable

que cada compañía construya su propia taxonomía. La categorización de los riesgos empleada por una empresa de telecomunicaciones será en muchos de sus términos diferente a la usada en una empresa del sector energético o la de un banco.

La gestión de riesgos no será efectiva si se percibe como un proceso aislado y ajeno a los mecanismos habituales de toma de decisión de las empresas. Todas las decisiones de negocio relevantes deberían ir acompañadas de un análisis de riesgos ya que mejora significativamente la calidad de la decisión y ofrece a la alta dirección una visión más realista de las posibles alternativas.

Integración de la gestión de riesgos en los procesos de decisión

Para mejorar la integración de la gestión de riesgos en los procesos de decisión debe haber un cambio en el tipo de información que recibe el equipo directivo. La simple inclusión de una sección denominada "Riesgos asociados y relación de acciones mitigantes" en la documentación recibida mejora la comprensión de los factores sujetos a incertidumbre y permite analizar los diferentes escenarios de ocurrencia orientando el esfuerzo hacia la mitigación. A partir de aquí, se puede evolucionar hacia fórmulas más avanzadas, como simular el grado de incertidumbre existente en el cumplimento de los objetivos estratégicos, presupuestarios o de otros indicadores clave de la empresa. Por ejemplo, el análisis de riesgos ayuda a sustituir los presupuestos estáticos por una distribución de posibles valores sujetos a incertidumbre. También ayuda a establecer indicadores de gestión basados en análisis de riesgos mejorando la probabilidad de cumplimiento. Otro ejemplo puede ser su empleo en los procesos de inversión. El uso de modelos de simulación para evaluar el atractivo de los proyectos permite a la empresa evitar muchas tomas de decisiones inadecuadas causadas por los métodos tradicionales de valoración como el valor actual neto. Con estas técnicas, se puede estimar la distribución de posibles resultados o cuantificar la probabilidad de obtener la rentabilidad objetivo, información relevante a tener en cuenta a la hora de valorar la oportunidad.

Cultura organizativa y gestión de riesgos

Las organizaciones, al igual que las personas, han de regirse por ciertos valores. Por ejemplo, en el caso de las empresas industriales, los valores esenciales han de ser el aseguramiento de la salud y la seguridad de las personas y la protección del medioambiente. Otro criterio fundamental es responder con transparencia a las

necesidades de los grupos de interés (clientes, proveedores, comunidades locales, autoridades, reguladores, etc.) promoviendo el desarrollo de negocios sostenibles que busquen el beneficio de la sociedad en su conjunto. Por lo tanto, si existiera algún conflicto entre estos valores y los resultados operativos, la organización tiene la responsabilidad de elegir los primeros, impulsando la dirección siempre esta elección.

Estos valores se traducen a través de la clasificación de los riesgos en función del denominado "apetito al riesgo". El apetito es la cantidad de riesgo que una organización quiere asumir para alcanzar sus objetivos. Teniendo en cuenta esta definición, existirán riesgos de "apetito cero", aquellos en los que la organización desea tener la mínima exposición. Estos riesgos no deben valorarse en función de criterios de coste/beneficio y deben mitigarse hasta niveles tan bajos como sea razonablemente posible. Dentro de estos riesgos, se incluyen aquellos que afectan a las dimensiones de cumplimiento, ética y conducta, salud, medioambiente, accidentabilidad, etc. La empresa ha de trazar claramente las líneas rojas ante la exposición a estos riesgos cuyos impactos van más allá de la mera afección económica. Por el contrario, existen riesgos intrínsecos a su actividad donde la cantidad de riesgo asumido por la organización dependerá de los potenciales beneficios. Estos riesgos son los propios del negocio, como pueden ser los de mercado o competencia, y se gestionarán en función de criterios de rentabilidad económica.

La gestión de riesgos no es sólo normativa, procesos, herramientas y técnicas. En definitiva, se trata de cambiar la cultura corporativa y la mentalidad de la alta dirección, la gerencia y de los empleados. Una cultura organizacional adecuada atraerá a individuos cuya postura ética inherente y su tendencia a asumir riesgos estará en línea con los estándares más elevados. El equipo de gestión de riesgos debe alentar a los empleados a plantear libremente las incertidumbres existentes y difundir las lecciones aprendidas tras la ocurrencia de hechos con consecuencias perjudiciales para la organización. En este cambio cultural, la actitud de los líderes resulta fundamental ya que deben ser proactivos a la hora de incentivar la identificación y prevención de riesgos por parte del personal. Este paso es crítico a la hora de transformar una organización "que culpa" en otra "que aprende" tras situaciones adversas sobreponiéndose a ellas gracias a las enseñanzas obtenidas de hechos pasados. Esta cultura organizacional redundará en una empresa con mayor capacidad de anticipación y garantizará su sostenibilidad a largo plazo.

Bibliografía

R.J. Chapman (2011). S*imple Tools and Techniques for Enterprise Risk Management*, John Wiley & Sons, West Sussex, United Kingdom.

J. Fraser, B.J Simkins (2010) *Enterprise Risk Management: Today´s Leading Research and Best Practices for Tomorrow´s Executives.* New Jersey, USA.

Virginia Leal Sanz ocupa el cargo de Gerente del Sistema Integrado de Riesgos. Tras un periodo en Arthur Andersen, se incorpora al equipo de Repsol en el año 2001. A lo largo de estos años ha despeñado diferentes cargos en áreas de la compañía como la Dirección de Tecnología o la Dirección de Medios. Desde el año 2013 trabaja en el Sistema de Gestión de Riesgos de Repsol, área encargada de la identificación y evaluación de los riesgos financieros y no financieros más significativos para organización con el objetivo de reportar los hechos significativos a la Alta Dirección. Es licenciada en Ingeniería Química por la Universidad Complutense de Madrid y Master en Negocio Energético por el Club Español de la Energía.

Tendencias en la modelización de riesgos de entidades financieras

Ignacio J. Carnicero González

Resumen

La evolución en los procesos de toma de decisiones dentro de la industria financiera quizá sea uno de los ejemplos pioneros y más longevos que podamos encontrar en el uso de técnicas estadísticas avanzadas. Se podría decir que el uso sistemático y masivo de datos para evaluar riesgos forma parte de la esencia misma del negocio bancario. Y su evolución más reciente, desde los procesos de automatización, digitalización y mejora de eficiencia, hasta la evolución de la gobernanza y normativa de modelos y prácticas, reglas de protección de consumidores, interpretabilidad etc. presentan peculiaridades y enseñanzas potencialmente útiles para muchos otros campos. El artículo intenta resumir las grandes tendencias vividas hasta el momento, los principales mitos existentes, errores cometidos etc…, a la vez que intenta señalar posibles líneas de trabajo futuro que permitan mantener la solidez de nuestro sistema financiero en un entorno cada vez más dinámico y complejo.

El dato como materia prima de la industria financiera

Cuando un niño intenta entender el mundo aprende usando analogías: las situaciones/cuestiones complejas suelen parecerse a otras más simples. Sin embargo, cuando alguien intenta usar esa técnica, comparando las instituciones financieras con el resto de sectores de la economía, suele encontrar muchas dificultades. ¿Qué fabrica? ¿Cuál es su materia prima? ¿Cómo consigue resultados? Siendo excesivamente simplista, el elemento fundamental de su existencia es gestionar valor (o su medida más conocida: dinero) a lo largo del tiempo. Para ello es imprescindible conocer bien los riesgos existentes, pero, como todo el que haya trabajado en la industria sabe, estos pueden ser casi infinitos: riesgo de crédito, de mercado, operacional, reputacional etc…Y desde muy temprano, los bancos han usado datos para evaluar sus principales riesgos.

Los datos, a diferencia de lo que sucede en otros sectores, han estado, por tanto, siempre en el centro de su proceso productivo. Pensemos, por ejemplo, en el riesgo de crédito. Desde los orígenes de la banca, un cliente con suficiente riqueza o garantías, con ingresos recurrentes… parecía, a los ojos de la lógica básica, un buen

candidato a recibir un préstamo. Es decir, si bien era el banquero quien tomaba la decisión final, solía apoyarse en datos objetivos, como la información sobre la situación patrimonial o la capacidad de repago del cliente, e incluso en nociones básicas de estadística, muchas veces de forma inconsciente.

Con el paso de los años y la extensión del préstamo entre diferentes estratos de la sociedad, el uso de técnicas estadísticas fue avanzando muy rápidamente. Los conceptos de "pérdida esperada" (como indicador del riesgo medio de una cartera de préstamos) o capital (pérdida extrema) aparecieron junto a la inclusión de los primeros modelos estadísticos de clasificación de clientes (modelos de *scoring* o de *rating*). El registro sistemático del historial de quiebras de los clientes, combinado con sus características en el momento de conceder un préstamo (sector económico, ratios financieros, etc…), permitieron identificar cuáles eran las variables significativas a la hora de estimar la probabilidad media de impago. Diferenciar correctamente los clientes en origen permitía mejorar sustancialmente el resultado económico de las carteras de préstamos.

Pero esto no sólo ocurrió con el riesgo de crédito, sino que se extendió al análisis de otros tipos de riesgos. Por ejemplo, también desde el origen de los mercados financieros se han intentado usar señales para predecir el comportamiento de los precios. Y por ello, conceptos como la volatilidad o la diversificación se extendieron con el uso masivo de los productos derivados (futuros/opciones…) en el último cuarto del siglo XX.

Toda esta evolución ha tenido como consecuencia natural el desplazamiento de la "opinión experta" por una "decisión informada" en muchas fases de la operativa financiera. El ejemplo más claro son los *scoring* de crédito para particulares, los cuales, basados puramente en experiencia histórica y criterios estadísticos, constituyen el criterio principal de concesión de un préstamo y, salvo excepciones muy particulares, no incluyen la "opinión" de los gestores de las oficinas o de analistas individuales.

Sin embargo, esta clara tendencia al uso generalizado de datos para la automatización de decisiones no ha sido un proceso continuo y uniforme, sino más bien oscilante y, a menudo, limitado a determinados ámbitos. La fe ciega en los datos ha chocado muchas veces con la realidad de encontrarnos con datos incompletos, sesgados o incorrectos, hipótesis equivocadas o simplemente factores no tenidos en cuenta. Esto ha provocado que sigan siendo necesarios numerosos mecanismos de control humano. Pongamos algunos ejemplos de este aprendizaje en forma de prueba y error que podrían ser útiles para otras industrias no financieras.

Exactitud vs simplicidad

Es una obviedad que la industria financiera es una de las más reguladas que existen. Tanto por el impacto masivo de sus acciones en nuestra sociedad, como por su exposición a los ciclos económicos y la naturaleza de los riesgos que afronta, siempre se ha intentado incluir reglas para conseguir un funcionamiento ordenado. Si bien en épocas tempranas la regulación estaba vinculada a aspectos más relacionados con la competencia (ruptura de monopolios) o con la protección de los consumidores (usura, garantía de depósitos). Con la extensión de técnicas avanzadas en la medición de riesgos las autoridades tomaron conciencia de que había que garantizar que los modelos usados en la toma de decisiones también fueran lo más adecuados posible, sin sesgos, prudentes, etc…Y en este sentido la crisis económica del 2008 puso a prueba la estructura de supervisión de las entidades financieras.

Uno de los ejemplos paradigmáticos de cómo han aprendido los supervisores ha sido la evolución de la normativa de capital de Basilea.[33] En pocas palabras, el negocio bancario siempre tiene en cuenta las pérdidas medias de su actividad (pérdida esperada por impagos/fraudes) y, en consecuencia, sus "precios" (margen, tipos de interés etc…) lo recogen como un coste más. Sin embargo, hay toda una serie de catástrofes de ocurrencia más extraordinaria (perdida inesperada) para las que los bancos deben mantener un capital mínimo que les permita soportar sus consecuencias sin quebrar. El éxito del primer intento de estandarizar los criterios para fijar esta cifra de capital de los bancos (denominado Basilea I), proporcionó un lenguaje común que favorecía un terreno competitivo más homogéneo. Si los supervisores aplicaban correctamente estas normas, se podían comparar los riesgos entre diversas instituciones. Si una cartera exigía unos requerimientos de capital mínimo similares a otra, es que sus riesgos también lo eran. Hasta aquí los datos usados eran relativamente sencillos, ya que se basaban en categorías simples de los activos (deuda pública, préstamos a corporaciones…). No obstante, su evolución posterior, Basilea II, abrió la puerta al uso de modelos internos de cálculo: cada institución, dentro de sus capacidades e información disponibles, podía adaptar el cálculo de sus riesgos de modo más adecuado al perfil de sus clientes, provocando aproximaciones heterogéneas o "personalizadas". En la medida en que el capital es escaso y por tanto una restricción, el uso de modelos internos pasó a ser una

33 Los Acuerdos de Basilea son los acuerdos de supervisión bancaria o recomendaciones sobre regulación bancaria emitidos por el Comité de Basilea de Supervisión Bancaria. Están formados por los acuerdos Basilea I, Basilea II y Basilea III. Reciben su nombre a partir de la ciudad de Basilea, Suiza, donde el CBSB mantiene su secretariado en la sede del Banco de Pagos Internacionales. Ver https://www.bis.org/

herramienta de ventaja competitiva, no sólo por parte de las instituciones, sino también de determinadas jurisdicciones, y, por tanto, el refinamiento metodológico fue visto desde el principio como una herramienta para optimizar (minimizar) el consumo de capital.

Y en este sentido la crisis puso en evidencia carencias importantes en algunos de estos enfoques metodológicos. Por un lado, la ciclicidad y el continuo cambio en el funcionamiento de las economías hacen que las series históricas no siempre recojan observaciones adecuadas para sucesos extraordinarios (por ejemplo, el efecto en cadena de la crisis inmobiliaria, el colapso de los mercados de deuda pública, etc…), lo que requiere mantener ciertos criterios conservadores a la hora de usar la información disponible. Por otro lado, la complejidad de los modelos usados (por ejemplo, en riesgo de mercado / valoración de productos financieros) ofrecía una sensación de "falsa seguridad" derivada de su perfección técnica; sin embargo, algunas de sus hipótesis base a veces no eran del todo adecuadas/realistas o bien no recogían elementos relevantes, como el riesgo de quiebra de una contrapartida o el riesgo de liquidez.

La reacción de los reguladores (Basilea III, con sus distintas y extensas variantes) ha sido la de una relativa vuelta a soluciones más sencillas para aquellos riesgos cuya modelización/datos era menos robusta (por ejemplo, el riesgo operacional), y una mayor tendencia general a la homogeneización. Por resumirlo en una frase: si no hay evidencia sólida para usar modelos sofisticados, mantengamos el análisis simple y comprensible. La reacción normativa tiende a reducir la complejidad, limitando la discrecionalidad no suficientemente justificada de los modelos de riesgo internos (de hecho, Basilea III limita el ahorro máximo de capital de los modelos internos frente a los métodos estándares al 27,5%), favorece unos mínimos comunes que ayuden a un terreno de juego más previsible y equilibrado, y recupera una mayor discrecionalidad supervisora. Este último elemento, es quizá el que más se aleja de un mundo totalmente cuantificable u objetivo y precisamente busca el cuestionamiento continuo. Supone reconocer que hay elementos como la estrategia de las entidades, su organización, riesgos nuevos, menos conocidos, u otras cuestiones que no se pueden medir correctamente pero que deben tenerse en cuenta. Con estas medidas se intenta evitar por ejemplo comportamientos deterministas que amplifiquen los efectos "manada" en procesos de crisis, al tiempo que intentan identificar aquellos modelos que pudieran ser menos sólidos.

Tendencias vs modas

Durante la última década, el término *big data* ha sido usado de un modo tan intenso, y muchas veces tan erróneo, que ha pasado a perder buena parte de su sentido. Realmente con ese término se suelen mezclar varios conceptos muy diferentes entre sí, tanto en definición como en su grado de desarrollo o utilidad:

- Fuentes de datos masivas: La materia prima sustancial de todo modelo de decisión son los datos, y la propia estadística nos dice que, en general, cuantos más datos, mejor. Sin embargo, aunque potencialmente la mayor parte de actividades humanas producen grandes volúmenes de datos (historial de compra, transacciones, vida laboral...) muchas veces estos datos no están aún disponibles, contienen numerosos sesgos/errores o, realmente, no se les puede considerar *big*. Por definir una regla básica, solo volúmenes de datos de miles de millones de registros se podrían considerar realmente *big data* y esto hace que pocas fuentes de datos actuales tengan un tamaño suficientemente grande como para ofrecernos tendencias estables y compensar el ruido de la información errónea.
- Capacidades computacionales: En este ámbito sí que podemos señalar que ha habido una importante mejora, y que sus efectos están siendo verdaderamente transformacionales. Los costes se han reducido dramáticamente y los tamaños de las bases de datos fácilmente tratables han crecido de modo exponencial. Esto ha permitido un servicio más rápido y eficiente a los clientes (plataformas digitales) que, en el mundo de los riesgos financieros, nos ha obligado también a que nuestros modelos de admisión o de fraude respondan con la misma calidad y velocidad. Si ahora un cliente puede solicitar desde su móvil un préstamo de modo sencillo y casi inmediato es también gracias a que las herramientas de control de riesgos se han sofisticado para proporcionar esa misma experiencia de cliente.
- Técnicas estadísticas: Pocas novedades importantes se han añadido en estos últimos años a lo que sería el *corpus* metodológico básico, pero la combinación de nuevas/mejores fuentes de datos y mayores capacidades computacionales sí han permitido sacar provecho a algunas de las técnicas más novedosas que, aunque ya conocidas, habían sido poco viables en la práctica hasta ahora: *machine learning*, teoría de grafos etc… son conceptos que ahora sí se pueden utilizar en procesos financieros diarios y no sólo en un laboratorio.
- Uso: También se ha consolidado un cierto consenso sobre el hecho de que las decisiones humanas pueden (y muchas veces deben) ser apoyadas por

sistemas de recomendación o análisis de datos automáticos. Estos últimos años también se ha avanzado mucho en entender los sesgos de los propios equipos de analistas de riesgo. Todo ello, y la fiabilidad de las predicciones basadas en datos, han favorecido una tendencia clara a aumentar el uso de estas técnicas.

Combinando todo lo anterior, la industria financiera ha pasado desde los intentos iniciales de descubrir un "santo grial" para predecir el futuro simplemente escuchando lo que nos dicen los datos, a una situación actual más realista que podríamos caracterizar con los siguientes principios:

- Lo verdaderamente diferencial parece ser más la existencia de datos nuevos y granulares de los clientes que las técnicas aplicadas.
- Las técnicas más dinámicas (por ejemplo, *machine learning*), por ahora sólo parecen dar mejores resultados que las "clásicas" para problemas que son cambiantes por naturaleza como, por ejemplo, la detección de fraude. Y es por eso que muchas veces, la mejora predictiva no compensa la mayor complejidad en su implementación.
- La normativa de privacidad de los clientes, sobre todo en países occidentales (por ejemplo, GDPR), está restringiendo mucho el uso de los datos que podrían ser diferenciales. Adicionalmente, se están incluyendo nuevas políticas para garantizar una toma de decisiones más transparente para el cliente como reglas de "no discriminación" en los algoritmos, o comprensibilidad de los mismos. Todo ello limita los potenciales beneficios de las mejoras técnicas en los resultados bancarios.
- *Strong data* frente a *big data*. Con este concepto se suele aludir a que las variables realmente importantes para algunos modelos (como los modelos de crédito) suelen ser relativamente pocas y suelen estar vinculadas a la experiencia histórica específica (en nuestro ejemplo, específica del proceso de concesión de crédito) y no tanto a comportamientos en otras actividades de la vida (redes sociales, cumplimiento legal etc…). Es decir, si las personas se comportan de modo muy diferente en situaciones diferentes, el *big data* puede perder parte de su valor.
- La calidad y la completitud en la captura de la información siguen siendo muy relevantes. Al menos en la industria financiera, la inversión en buenos sistemas (y procesos) de almacenaje de información es siempre rentable. Con entornos de información escasa, siempre es mejor capturar buena información que intentar limpiarla después.

Pongamos dos ejemplos diferentes que muestran la aplicación real de estos principios en la industria de préstamos al consumo:

Estados Unidos: La expansión de la denominada industria *fintech* o de préstamo al consumo bajo canales digitales durante la fase de recuperación económica iniciada en 2010, podría interpretarse a primera vista como una victoria del uso de los datos. Sin embargo, si bien esta afirmación puede ser cierta en cuanto a los datos para ofrecer un producto al cliente (experiencia de uso o aproximación comercial), no lo es en cuanto a modelos de riesgos. EEUU es un país cuyos procesos de concesión de crédito a particulares están muy centrados en el uso de la calificación de crédito proporcionada por FICO,[34] y esta nueva industria sigue usando fundamentalmente este criterio. Es decir, aunque se han incluido variables adicionales o análisis más sofisticados, el uso continuado de esta calificación evidencia que, al menos hasta el momento, las variables fundamentales en el comportamiento crediticio de los clientes siguen siendo las mismas, lo que parecería una prueba de que la aproximación al *strong data* puede tener sentido. Esto no impide que, por ejemplo, en materia de fraude, la información de actividad digital sí ha provocado claras mejoras predictivas.

China: En este caso, las conclusiones son algo diferentes. Una normativa claramente más laxa en términos de privacidad individual y un mercado potencial de grandes dimensiones, han permitido desarrollar de un modo muy rápido una serie de compañías de financiación al consumo similares a las norteamericanas pero, en este caso, con un mayor uso de fuentes de información alternativas. Concretamente, las fuentes derivadas de las compañías telefónicas (historial de llamadas, localización, dispositivos etc…) y las técnicas de reconocimiento facial aplicadas a dispositivos móviles han provocado que el fraude en esta industria sea mínimo. Quizá las mejoras en predicción de impagos sean menos impresionantes y más similares a las de otros países, y, además, aplicar algunas de estas técnicas bajo regulaciones como la europea sería complicado, pero la experiencia china nos deja al menos algunas enseñanzas relevantes:

- Las mejoras sustanciales en los modelos de fraude sostienen que sigue habiendo fuentes de datos muy útiles no explotadas. ¿Qué fuentes quizá no hayamos encontrado aún en otros mercados/industrias?

34 FICO score es una métrica introducida por a la compañía de análisis de riesgos crediticios FICO (*Fair, Isaac and Company*) usada por la mayoría de los bancos y entidades de crédito de EEUU. Aunque su composición exacta no es conocida, es una medida basada en pesos otorgados a diferentes componentes (historial de pago, tipos de créditos utilizados, solicitudes de préstamo, peso de la deuda, historial de crédito).

- La explosión de la industria financiera en este país también demuestra las ganancias en términos de eficiencia y productividad derivadas de usar decisiones y técnicas de aprendizaje automáticas, dentro de entornos homogéneos y con un tamaño relevante como es el mercado de financiación al consumo chino.

Reflexiones finales

Aunque el espíritu de este artículo parezca desmitificar en parte las ventajas del *big data* o de las técnicas automáticas de decisión, realmente el objetivo es justo el contrario: destacar sus fortalezas. Los seres humanos cometemos errores, tenemos sesgos y evidenciamos claras carencias para procesar volúmenes altos de información desordenada e identificar patrones. Y es en estas carencias donde deberíamos enfocar el uso de las nuevas técnicas: como un apoyo realmente diferencial, con resultados reales y ventajas clarísimas en términos de eficiencia, velocidad y de gestión en entornos cambiantes. La industria financiera está en un estado avanzado en el uso de datos, ya ha atravesado varias de las fases iniciales en este proceso, que nos lleva desde la figura todopoderosa del experto a la de los modelos automáticos y, como hemos intentado reflejar con ejemplos, también se han cometido errores que se han ido corrigiendo. Pero como tendencia de fondo, y como mejora para la asunción de riesgos, estabilidad financiera etc., las ventajas del uso generalizado de datos en la toma de decisiones son indudables.

No obstante, sí deberíamos añadir un comentario final a modo de conclusión: no se pueden obviar las dificultades de implementación. Para industrias nuevas (como las mencionadas *fintechs*) implementar sistemas y técnicas nuevas es relativamente sencillo. Pero en industrias más establecidas como la financiera, la transición de procesos de decisión, almacenaje de datos, criterios etc. es claramente más complejo. Y no solo por la migración de herramientas, sino también por el factor humano. Confiar en un algoritmo es un aprendizaje en sí mismo. Perder el control de una decisión es un gran cambio para muchas personas. Por eso no olvidemos que la transformación de equipos es uno de los aspectos más relevantes de este proceso. Seguiremos necesitando personas para formular correctamente las preguntas a responder, para sospechar que algo no funciona bien, para identificar que la realidad ha cambiado lo suficiente como para inventar nuevos modelos etc. Y ese cambio de decisores a diseñadores de decisiones es largo, difícil y…no es modelizable.

Ignacio J. Carnicero González es actualmente responsable de BBVA Global Risk Management Analytics. Su carrera en el grupo BBVA comenzó en 2007 en los orígenes del equipo Portfolio Management dentro de la división de Banca Corporativa e inversión (CIB), convirtiéndose en responsable del mismo en 2009, bajo el mandato de gestionar activamente el portfolio global de riesgo de crédito de dicha división. A finales de 2014, se incorporó al área de riesgos para liderar el equipo Analytics responsable del desarrollo de los modelos de riesgos para el grupo BBVA. Previamente, trabajó durante un año en consultoría financiera y durante siete años en diversos puestos de responsabilidad dentro del grupo Santander (Asset Management, Dirección Financiera etc…). Cuenta con una licenciatura en Ciencias Económicas por la Universidad de Valladolid, un postgrado en Economía y Finanzas por CEMFI (Banco de España), un curso de postgrado en Seguridad y Defensa por el Instituto Gutiérrez Mellado, y la certificación FRM (Financial Risk Manager).

Capítulo IV: Profundizando en el conocimiento

Aprendizaje Máquina: Evolución, presente y previsiones

Aníbal R. Figueiras Vidal

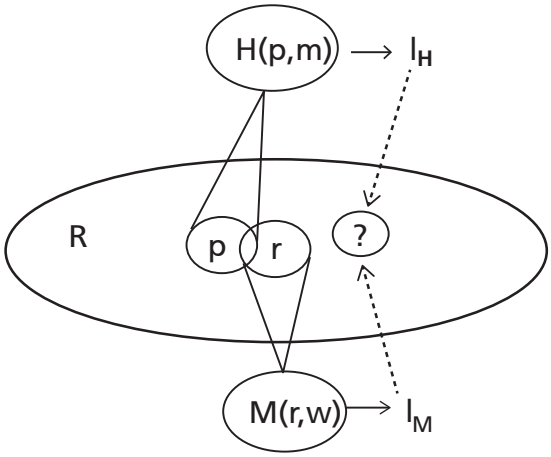
Resumen

En este capítulo se presenta el concepto y los fundamentos generales del Aprendizaje Máquina y se revisan de modo conciso su evolución y su estado actual, resaltando los aspectos de mayor relevancia teórica y aplicada. Una perspectiva de lo que cabe esperar en un futuro, con expresión de lo que el autor considera más importante, cierra la exposición.

Concepto y aspectos fundamentales

Si una persona desconoce un aspecto de la realidad "general" –incluyendo el futuro y aspectos no físicos–, puede utilizar un proceso de inferencia a partir de percepciones sensoriales y el contenido de una de sus memorias –típicamente, la de trabajo– mediante algún proceso mental, obteniendo una estimación del valor de una magnitud o una decisión entre un número de alternativas. Una máquina (un algoritmo) de aprendizaje opera de modo similar: a partir de unos datos –suministrados por sensores y registros– y de unos parámetros –cuyos valores han sido fijados mediante un adecuado proceso de entrenamiento– proporciona una inferencia: una estimación o decisión "máquina".

Gráfico 1. Inferencia humana y máquina



R: realidad; ?: aspecto desconocido; p: realidad percibida; r: datos registrados (o medidos); m: (una) memoria; w: parámetros; H: humano; M: máquina; I_H: inferencia humana; I_M: inferencia máquina.

La similitud expuesta no debe conducirnos a falsas extrapolaciones: en términos generales, los procesos que conducen a las inferencias son distintos. Tampoco existen en las máquinas emociones, sentimientos ni principios (aunque pueden simularse). El algoritmo máquina tiene valor instrumental: ayudarnos en tareas que resultan tediosas, arriesgadas, difíciles o hasta imposibles por nuestras limitaciones de percepción y cálculo. También ha de comprenderse que una máquina aprende lo que se le permite aprender. Y conviene señalar que un mismo algoritmo puede aplicarse a la resolución de distintos problemas de inferencia: basta sustituir los datos y los valores de los parámetros por los apropiados para el caso, como los humanos hacemos con lo percibido y el contenido de la memoria empleada.

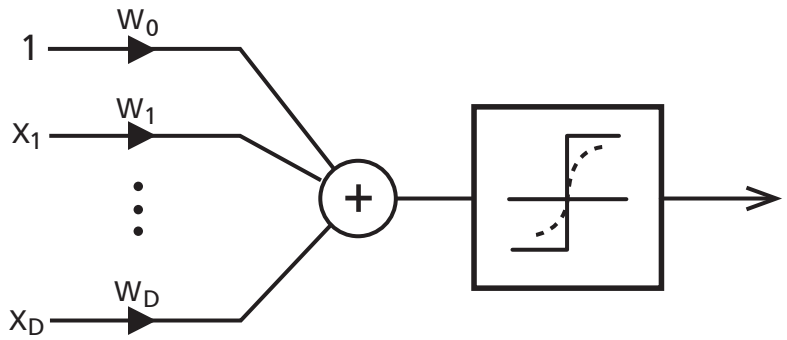
Es tradicional dividir las máquinas de aprendizaje en dos categorías (Bishop 2006): a) las generativas, que construyen un modelo (estadístico) del problema a resolver; y b) las discriminativas, que elaboran directamente una respuesta obtenida mediante transformaciones de los datos que se establecen para maximizar una medida de calidad (o minimizar una de riesgo). De las primeras se dice que son explicativas, aunque cabe oponer que ciertos modelos resultan difíciles de comprender incluso para especialistas. Se considera que las segundas no son explicativas. No obstante, debemos resaltar que, bajo ciertas condiciones, sí proporcionan

información inteligible, como son estimaciones de las probabilidades *a posteriori* de cada alternativa –aunque no aclaraciones sobre qué importa en su obtención. Piénsese que las personas procedemos así en muchas ocasiones para decidir: por ejemplo, en elecciones inconscientes. Además, en términos generales, las máquinas discriminativas suelen brindar mejores prestaciones. Dada esa ventaja, recibirán en este capítulo un tratamiento preferente, así como los problemas de decisión, con los que nos encontramos continuamente: qué pensamos, qué leemos, a dónde nos dirigimos, qué postura adoptamos, qué decimos; en realidad, decidimos en cada instante.

Evolución[35]

Se puede decir que las máquinas aprenden a partir de ejemplos: observaciones, normalmente con sus etiquetas, que dan respuesta a lo que se desconoce. Aunque la Estadística tradicional propuso métodos (generativos) para utilizar ejemplos con propósitos de inferencia, se acepta mayoritariamente que el origen del Aprendizaje Máquina fue el Perceptrón de Frank Rosenblatt para decisión binaria, nacido a finales de los 1950.[36]

Gráfico 2. Perceptrón de Rosenblatt



. x_1…x_D: valores observados; w_1…w_D: pesos; +: indicador de suma. El bloque final incluye la activación dura –signo– o blanda (a trazos).

35 Las notas históricas de (Duda *et al.* 2001) amplían lo que se dice en esta sección.

36 Se considera a Frank Rosenblatt el padre del Aprendizaje Máquina. Las limitaciones expresivas de un Perceptrón produjeron un estancamiento de años en el avance de las redes neuronales).

El Perceptrón buscaba unos parámetros, los pesos $\{w_{-d}\}$, que diesen lugar a que el valor sgn $(w_{-0}+w_{-1}\ x_{-1}+\cdots w_{-D}\ x_{-D})$, donde $\{x_{-d}\}$ son las variables que componen un dato y sgn la función signo (+1 si el valor de paréntesis es positivo, -1 caso contrario), identificase la decisión a tomar.

Rosenblatt comprendió que los algoritmos de búsqueda tradicionales, que minimizan secuencialmente una medida de error, no podían aplicarse para el entrenamiento porque la función signo no es derivable, y recurrió a los principios hebbianos o de aprendizaje por refuerzo (en este caso, negativo): ante un error, los pesos se modificaban en el sentido de corregirlo.

Marvin Minsky, investigador del MIT, mostró un decenio después las limitaciones del Perceptrón, no sólo algorítmicas, sino en cuanto a capacidad: ésta se reducía a tomar decisiones para problemas en que una combinación lineal de los datos servía de frontera entre el sí y el no. De poco valió el que otros investigadores propusieran soluciones basadas en una modificación menor, consistente en sustituir la función signo por una aproximación derivable, llamada activación blanda, lo que permitía un entrenamiento mediante algoritmos de búsqueda convencionales y, aún más importante, construir máquinas más potentes mediante apilamiento de unidades como el Perceptrón con activación blanda. Así se posibilitaban fronteras de formas arbitrarias, y el entrenamiento sólo requería una derivación ordenada (el algoritmo de Retropropagación). Nacían con ello los llamados Perceptrones Multi-Capa: agrupaciones en paralelo de perceptrones blandos, apiladas una sobre otra.

De este primer periodo de crisis y abandono casi completo de los trabajos en Aprendizaje Máquina no se salió hasta que otro profesor del MIT, David Rumelhart, y una larga lista de colaboradores internos y externos formalizaron en 1986 las posibilidades expuestas. Siguió un enfebrecido repunte de la investigación y las aplicaciones, animado además por los teoremas no constructivos sobre la capacidad ilimitada de los perceptrones con una capa oculta para establecer relaciones entre entrada y salida que George Cybenko (Cybenko 1989) y Kurt Hornik *et al*. (Hornik *et al*. 1989) presentaron en 1989.

A comienzos de la década de 1990 surgieron dificultades. Primera: el número de datos etiquetados prácticamente disponibles para entrenar perceptrones con una capa oculta limitaba su capacidad expresiva. Al mismo tiempo, los intentos de utilizar arquitecturas con más capas ocultas –lo que implicaba transformaciones más complejas, y subsiguientemente potenciales mejoras– fracasaban, salvo para

algunas formas sencillas, porque las derivadas se desvanecían al progresar en el algoritmo de Retropropagación (por saturación de las activaciones blandas). Llegó entonces la segunda recesión en el desarrollo del Aprendizaje Máquina.

Durante los 1990 surgieron las conocidas como Máquinas de Vectores Soporte, que realizaban transformaciones generales preestablecidas y las combinaban maximizando una medida de la separación entre las muestras. Pero también se continuó la investigación sobre perceptrones con una capa oculta (y sobre otras máquinas más tradicionales, como los árboles de decisión y regresión). Los avances con estas máquinas siguieron una dirección muy relevante: entrenar los algoritmos bajo condiciones diversas, de modo que una bien elegida agregación de las salidas resolviese mejor el problema bajo análisis. Se trata de los conjuntos de máquinas (Rokach 2010). Muchos han sido los éxitos de estos planteamientos, y hoy siguen despertando interés y atención. La razón fundamental de este éxito radica en su efectividad para combatir las limitaciones impuestas por el reducido número de ejemplos de entrenamiento en relación con la dificultad del problema a tratar.

Finalmente, a mediados de los 2000 se dieron los primeros pasos para diseñar y entrenar perceptrones con varias (o muchas) capas ocultas, que son las hoy denominadas Redes Neuronales Profundas (Bengio 2009). Por su importancia y actualidad, se tratarán en el apartado siguiente.

La situación actual

Hoy en día, casi todas las menciones al Aprendizaje Máquina se hacen refiriéndose a *Big Data* (Macrodatos, prefiere la RAE), véase (Su). Las definiciones de los problemas a los que alude dicha denominación hacen uso de "las uves": volumen de los datos, velocidad de su aparición, variedad de sus formas, variabilidad temporal… y, finalmente, valor –que ha de entenderse como el que tiene la solución del problema.

En opinión de quien escribe, es la v final la que delimita realmente el ámbito. Como quiera que la predicción meteorológica, la categorización de textos, la genómica o la medicina personalizada son valiosas, e intrínsecamente implican otras uves, justifican el empleo de los algoritmos y arquitecturas *Big Data*, recursos computacionales (del tipo paralelización e intercomunicación, gestión de memorias…) que permiten enfrentarse a ellas, y sobre los que no cabe extenderse aquí. Pero en otros muchos casos recurrir a esas tecnologías no solamente puede ser innecesario e improductivo, sino hasta perjudicial. Así, manejar un exceso de ejemplos para definir la frontera

de una clasificación lineal –un hiperplano– es un inútil gasto. E incluir variables sin relación estadística con la etiqueta es una forma de conseguir los perturbadores efectos del ruido.

Las Redes Neuronales Profundas iniciaron su actual explosión de investigación, desarrollo y uso con los trabajos de Geoffrey Hinton a mediados de los 2000, siguiendo una brillante idea: se diseña una máquina para resolver un problema distinto del que se desea atacar, pero que proporciona ventaja en la representación de los datos, y después se refina el resultado para la inferencia que se persigue. Un ejemplo de fácil comprensión es el de los Auto-Codificadores Apilados con Reducción de Ruido de Yoshua Bengio y sus colaboradores: son máquinas profundas que se diseñan capa a capa (lo que no presenta problema) para eliminar un ruido que se añade a las muestras de entrada, y, tras ello, se lleva a cabo un segundo entrenamiento orientado a la resolución del problema que se considere. El primer entrenamiento obliga a que las representaciones de la entrada conseguidas a través de las transformaciones consecutivas de las diversas capas contengan información muy refinada sobre los datos. Por tanto, partir de ellas para hacer los ajustes finales necesarios es más que razonable.

En el último decenio se han propuesto y probado con éxito diversos procedimientos para hacer posible el entrenamiento directo de las Redes Neuronales Profundas. Entre ellos están: las activaciones que impiden la anulación de las derivadas, los parámetros complementarios para controlar las saturaciones, la utilización de grupos de muestras para establecer los parámetros entrenables –pesos– en lugar de muestras una a una, y varias más. Por ello, cabe vaticinar una productiva progresión en el estudio y la aplicación de estas máquinas. Sabemos que, en estos momentos, ya acreditan prestaciones superiores a las de los expertos… incluso en tareas que los humanos tenemos por propias, como la identificación facial, el reconocimiento de emociones/sentimientos, la conversión texto-voz e incluso juegos de estrategia como el ajedrez o el go.

Concluiré esta sección con unos párrafos dedicados a un aspecto fundamental del Aprendizaje Máquina que, aunque haya aparecido esporádicamente en el pasado, solo se ha abordado sistemáticamente en el siglo XXI: la resolución de Problemas Singulares.

Se propone la denominación de Problemas Singulares para aquellos que requieren modificaciones *ad hoc* de los algoritmos de aprendizaje convencionales, porque la

directa aplicación de estos conduce a resultados decididamente subóptimos. Hay muy destacables familias de estos problemas, que aparecen frecuentemente en importantes aplicaciones reales: los de clasificación desequilibrada (Branco *et al.* 2016), en los que las muestras de las diferentes clases constituyen poblaciones de muy distintos tamaños –y que conducen a los algoritmos convencionales a decidir en exceso a favor de las clases mayoritarias, con un sesgo similar al de los humanos–, constituyen los más conocidos; lo que se comprende al darse cuenta de que incluyen la intrusión –física o en red–, la detección de defectos y averías, el reconocimiento de objetos, el diagnóstico (no sólo médico), la identificación de signos y símbolos, y otros muchos no menos importantes; incluso la detección de amenazas a instalaciones críticas. Otros problemas singulares son los de costes dependientes de la muestra –fraude, concesión de crédito, rotación de clientes, marketing directo…–, muchas veces también desequilibrados. Nótese que se trata de problemas presentes en el núcleo de los negocios y las finanzas.

La clasificación ordinal, aquella en la que las diferentes clases implican un orden que conviene mantener en lo posible, es necesaria para la asignación de prioridades en servicios de salud, asistenciales, y similares. Los problemas multietiqueta, en los que ha de procederse simultáneamente a varias clasificaciones –como tipos de objetos presentes en una imagen–, suponen otra familia digna de mención. Y otras varias que no es posible incluir en un texto tan breve. Como se ve, muchas y de interés; pese a ello, gran cantidad de las modificaciones introducidas en los algoritmos convencionales para tratar con ellas son puramente empíricas, lo que ni implica éxito, ni siquiera garantiza la mejora de prestaciones con respecto a tratamientos convencionales. Concebir, diseñar, evaluar y explotar algoritmos fundamentados –que sí ofrecen garantías– para estos problemas constituye uno de los mayores desafíos en Aprendizaje Máquina.

Una previsión del futuro

Para la lectura de esta sección no conviene olvidar que las predicciones de lo por venir incluyen inevitablemente los efectos de mucha incertidumbre –con variados orígenes– y de las carencias de quien predice. Aquí se reducirán los aspectos considerados a los que, en opinión del autor, tienen potencialmente mayor alcance.

Claro está que los avances en las tecnologías *Big Data* propiciarán una ampliación en número de grandes problemas que se traten con algoritmos de aprendizaje, además de una mejora de los resultados. Pero todavía mayor peso tendrá el progreso en

concepción, diseño y uso de Redes Neuronales Profundas: el *Big Learning* promete aún más. Y, para acercarse a él, no ha de renunciarse a agrupar metodologías de acreditado rendimiento, como la diversificación, junto con los deseables avances en sistemas distribuidos y en aprendizaje dinámico –no parece haberse hallado receta general que resuelva el permanente dilema exactitud vs. adaptabilidad: la rapidez de reacción no permite afinar en los resultados. Esta opinión nace de un punto de vista personal, y a la vez firme: las máquinas aprenden lo que se les enseña, e importa tanto el qué –los datos– cuanto el cómo, es decir, el proceso de entrenamiento. Tal es la guía principal a seguir.

La ampliación en varias direcciones de las formas de memoria a incorporar tiene importancia fundamental, no sólo para adaptarse a las características del problema –como emplear formas a corto y a largo plazo en entornos variantes– o para construir máquinas de funciones múltiples, sino incluso para analizar sus contenidos y deducir mecanismos para su gestión. En estadios del Aprendizaje Máquina próximos a la Inteligencia Artificial "fuerte" –con propósitos más generales que la resolución de problemas concretos–, la oportuna gestión de distintos tipos y niveles de memoria –tal vez de modo similar a como lo hace nuestro cerebro– supone una necesidad.

El mejor conocimiento sobre máquinas para Problemas Singulares ha de mantenerse como un objetivo clave. No se trata tan solo de hacer viable la resolución de muchos problemas importantes, sino también de progresar en la construcción de una taxonomía de los problemas de aprendizaje, con la esperanza de que llegue un momento en que cada problema pueda recibir un tratamiento individualizado.

Para concluir: no se debe regatear esfuerzo encaminado a avanzar en la comprensión de las máquinas de aprendizaje; esfuerzo que deberá incluir vías indirectas de interpretación, y hasta adaptaciones de algunos procesos mentales.

Bibliografía

C. M. Bishop. (2006) *Pattern Recognition and Machine Learning.* Springer, New York, NY, USA.

R.O. Duda, P. E Hart, D. G. Stork. (2001) *Pattern Recognition* (2nd ed.). Wiley, New York, NY, USA.

L.R. Rokach. (2010) *Pattern Classification Using Ensemble Methods.* World Scientific, Singapore.

Y. Bengio. (2009) *Learning Deep Architectures for AI.* Foundations and Trends for Machine Learning 2 (1), 1-127. Now Publishers, Hanover, MA, USA.

X. Su. *Introduction to Big Data.* https://www.ntnu.no/iie/fag/big/lessons/lesson2.pdf

P. Branco, L. Torgo, R.P. Ribeiro. (2016) A Survey of Predictive Modeling on Imbalanced Datasets. *AMC Computer Surveys*, 49, 31:1 – 31:50.

G. Cybenko. (1989) Approximation by Superpositions of a Sigmoidal Function. *Mathematics of Control, Signals and Sytems*, 2, 303-314.

K. Hornik, M. Stinchcombe, H. White. (1989) Multilayer Feedforward Networks Are Universal Approximators. *Neural Networks*, 2, 359-366.

Anibal Figueiras es doctor Ingeniero de Telecomunicación por la UPC (1976) y Catedrático de Universidad (1978). Presta actualmente sus servicios en la Universidad Carlos III de Madrid, dedicando su docencia e investigación a los fundamentos probabilísticos de las Máquinas de Aprendizaje y a la concepción y aplicaciones de nuevos diseños. Es Académico Numerario de la Real de Ingeniería desde 2000. Entre otras distinciones, se le ha otorgado el Doctorado "Honoris Causa" por las Universidades de Vigo (1999) y San Pablo de Arequipa (Perú) (2012), y ha recibido el Premio "Miguel Catalán" de la Comunidad de Madrid a la trayectoria científica (convocatoria 2018).

El potencial de la Inteligencia Artificial en el ámbito de la inteligencia estratégica y la seguridad

Carme Artigas

Resumen

La Inteligencia Artificial se ha convertido en la tendencia tecnológica de más impacto de la última década, gracias a los avances en la computación y el procesamiento masivo de datos, *Big Data*, que ha permitido el desarrollo de la analítica avanzada de datos y de sus múltiples aplicaciones en todos los ámbitos y sectores empresariales. Su uso es creciente, desde el reconocimiento avanzado de imágenes, el procesamiento del lenguaje natural que da lugar al desarrollo de *chatbots* y asistentes personales inteligentes, así como las interacciones con las máquinas a través del *Internet de las Cosas* y la robótica. En el ámbito de la inteligencia estratégica y la ciberseguridad, la Inteligencia Artificial supone la incorporación de nuevas metodologías de trabajo, a través del *Machine Learning* y la detección de patrones y anomalías, que va a permitir un salto cualitativo y cuantitativo en la capacidad de prevenir y detectar ataques en tiempo real. Sin duda un claro ejemplo de la innovación tecnológica en este campo.

Introducción

Vivimos inmersos en una explosión de información sin precedentes en la historia de la humanidad. Las estadísticas revelan que tan sólo en los dos últimos años se ha generado el 90% de la información existente hoy en el mundo y esta tendencia crece exponencialmente. El nacimiento de una sociedad donde las personas están siempre conectadas a diferentes dispositivos móviles y a través de redes sociales y en el que las máquinas han empezado también a conectarse entre sí para intercambiar información, está provocando la digitalización de numerosos aspectos de la actividad humana.

A su vez, y de manera creciente, las decisiones están cada vez más determinadas por la manera en la que las organizaciones obtienen, analizan y utilizan la información. Nunca como hasta ahora entidades de todo tipo han tenido acceso a cantidades ingentes de datos provenientes no sólo de sus sistemas de información internos sino de muchas otras fuentes. Es lo que se llama *Big Data* ("macro datos" o "inteligencia de datos"), grandes volúmenes de datos en múltiples formatos (texto, imágenes,

sonidos, movimientos al navegar por la red o ante un cajero automático, datos de sensores, etc.), generados no sólo por la interacción de personas en internet o redes sociales –que generan una huella digital–, sino por la interacción de cosas y objetos (móviles, *tablets*, mercancías con RFID, etc.) conectadas a la red.

Big Data no es un término preciso, sino más bien la caracterización de un fenómeno basado en la acumulación de grandes volúmenes de datos de cualquier tipología, que crecen exponencialmente y que, ya sea por su tamaño, volumen, requerimientos de rendimiento, necesidad de baja latencia u otras consideraciones tecnológicas, no pueden ser gestionados mediante sistemas tradicionales de almacenamiento, gestión y análisis de datos.

Sus aplicaciones son múltiples en todos los sectores y campos. Sus beneficios a nivel de negocio los podemos ver en cuatro ámbitos: la generación de nuevas fuentes de ingresos gracias a la personalización, a las mejoras en la eficiencia operativa, en la toma de decisiones y en la prevención de fraude y riesgo. Para los consumidores, el *Big Data* permite nuevas aplicaciones y una mejor experiencia de usuario. Algunos ejemplos en nuestra vida cotidiana incluyen desde los motores de recomendación de una tienda online, la concesión automática de créditos y la predicción meteorológica hasta el tratamiento personalizado del cáncer y los coches autónomos.

La Inteligencia Artificial, un nuevo paradigma

Los avances en la computación de datos masivos, *Big Data*, han permitido en paralelo el avance de técnicas avanzadas en el análisis de datos, más allá de lo que permitía hasta ahora la estadística tradicional. Es lo que llamamos Ciencia de Datos, disciplina que engloba las capacidades de computación, desarrollo de *software* y analítica de datos avanzada, como *Machine Learning*, *Deep Learning* y otras formas de Inteligencia Artificial.

Englobamos bajo el nombre de Inteligencia Artificial, la capacidad de las computadoras en emular la capacidad cognitiva humana. Si bien la Inteligencia Artificial se empezó a desarrollar en la década de los años 50, a partir de los años 80 su desarrollo se estancó ya que no hubo avances a nivel de computación que fueran significativos para esta disciplina. Sin embargo, a partir del desarrollo del *Big Data*, se ha producido un salto cuantitativo que ha permitido desplegar técnicas más avanzadas para el análisis de datos.

El proceso de aprendizaje comienza con observaciones o datos, como ejemplos, experiencia directa o instrucción, para buscar patrones en los datos y tomar mejores decisiones en el futuro en función de los ejemplos que brindamos. El objetivo principal es permitir que las computadoras aprendan automáticamente sin intervención o asistencia humana y ajustar las acciones en consecuencia.

Las máquinas, al igual que las personas, tienen diferentes formas de aprender dependiendo de los datos con las que podamos alimentarlas. Es decir, los datos son para las máquinas lo que la experiencia es para los humanos. Existen varias estrategias de aprendizaje por parte de las máquinas: aprendizaje supervisado, aprendizaje no supervisado y aprendizaje de refuerzo.

El *Machine Learning* o aprendizaje automático es una aplicación de la inteligencia artificial que proporciona a los sistemas la capacidad de aprender y mejorar automáticamente a partir de la experiencia sin ser programado explícitamente. El aprendizaje automático se centra en el desarrollo de programas informáticos que pueden acceder a los datos y utilizarlos para entrenar y aprender por sí mismos.[37]

La revolución del mundo digital, la conectividad móvil y los avances en computación de datos masivos han permitido el desarrollo de unas capacidades únicas en tres ámbitos principales en los que hasta ahora los humanos éramos mejores que las máquinas: en el reconocimiento de imágenes, en el reconocimiento de textos y en el reconocimiento de voz.

En la actualidad, el desarrollo de la Inteligencia Artificial actual se centra precisamente en estos tres ejes:

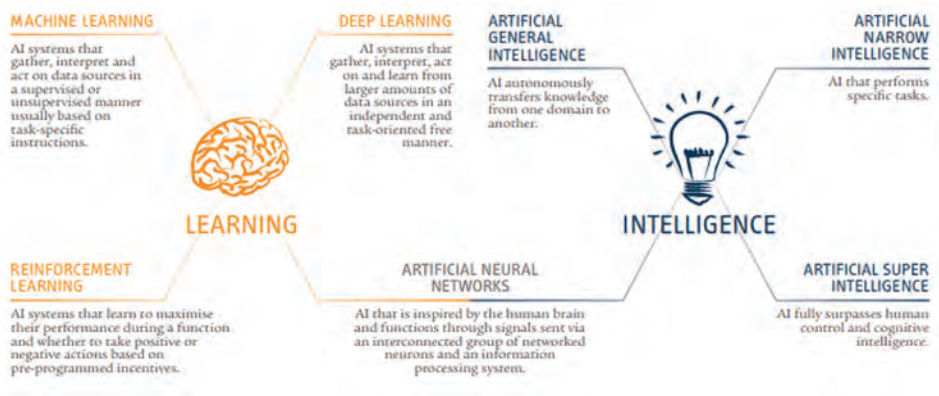
- Procesamiento del Lenguaje Natural (*Natural Language Processing*): permite que un agente cognitivo nos entienda cuando escribimos y sea capaz de entender y responder las preguntas que le hacemos por escrito en formato texto (ej. *chatbots*). Se apoya de las técnicas de análisis de texto (*Text Analytics*)
- Reconocimiento de imágenes (*Image Recognition*): permite desarrollar aplicaciones que a partir de una fotografía o video generado en un proceso, por ejemplo, en el acceso automático a un aeropuerto mediante pasaporte, reconocimiento facial o en el reconocimiento de una matrícula para acceder a un parking.

37 Para más información ver capítulo "Aprendizaje Máquina: Evolución, presente y previsiones".

- Reconocimiento del habla (*Speech Recognition*): permite que un agente cognitivo nos entienda cuando hablamos por voz (ej. asistentes personales inteligentes como Amazon Alexa, Siri de Apple o Aura de Telefónica).

Todos estos ejemplos de aplicación caen dentro de la categoría de Inteligencia Artificial específica (Narrow AI), refiriéndose a los algoritmos diseñados para resolver problemas concretos y específicos (reconocer una imagen, predecir un comportamiento, clasificar un grupo de elementos, calcular la probabilidad de un evento, detectar anomalías...). Para cada problema específico se necesita aplicar algoritmos de *Machine Learning* a partir de un conjunto de datos de entrenamiento específicos. Los expertos están de acuerdo en que harán falta muchas décadas para que el desarrollo de esta disciplina llegue a lo que llamamos Inteligencia Artificial General (General AI), que se refiere a sistemas complejos capaces de emular e incluso superar la capacidad de inteligencia humana para resolver tareas generales y diversas de manera amplia.

Figura 1. Clasificación de las distintas ramas de la Inteligencia Artificial



Fuente: EUISS (European Union Institute for Security Studies) 2018.

La intersección entre Inteligencia Artificial y Seguridad

El concepto de seguridad es muy amplio, ya que se puede definir en virtud de múltiples contextos y niveles. Se habla de la seguridad personal, la de un país, tanto en el nivel físico como en el del ciberespacio. El actual desarrollo de la Inteligencia Artificial permite encontrar aplicaciones que cubren todos estos ámbitos, si bien

no podemos ignorar que a medida que las amenazas y el contexto de seguridad evolucionan en función del desarrollo tecnológico, también lo deben las leyes y las regulaciones que los gobiernos deben implantar para mitigar los riesgos añadidos.

La ciberseguridad

El *software* que permite hacer funcionar nuestros ordenadores y aparatos inteligentes está sujeto a errores de programación de código, a la vez que vulnerabilidades de seguridad que pueden ser explotadas por delincuentes informáticos (*hackers*). Estas vulnerabilidades se pueden ramificar y expandir a gran escala, a nivel de una región o nación, como hemos visto recientemente en los ataques de virus informáticos a escala mundial.

La ciberseguridad ha pasado por distintas etapas o metodologías de detección de amenazas. La primera de estas fases de evolución fue la detección *basada en firmas*, donde las empresas utilizaban un conjunto de reglas para identificar amenazas inteligentes mediante la observación de patrones de eventos específicos de ataques conocidos y documentados. En este paradigma se emplean métodos de pre-procesamiento, como la inspección profunda de paquetes en el tráfico de la red, para encontrar posibles firmas en el tráfico de la red que se ha capturado. En un segundo estadio de evolución nos encontramos con la detección basada en anomalías, donde se buscan las acciones maliciosas no comunes que pueden ser detectadas anticipadamente e identificadas. A partir de modelizar el comportamiento de las redes, los sistemas, las aplicaciones, los usuarios y los dispositivos, se generan patrones de comportamiento de lo que sería un uso regular. Cuando se detecta una desviación o anomalía de ese patrón es cuando se puede anticipar el ataque.

Actualmente la seguridad basada en Inteligencia Artificial permite fusionar ambos planteamientos. Un sistema es capaz de aprender de manera autónoma a partir de incidentes de detección y gestión de amenazas, detectando patrones, identificando anomalías, reentrenándose permanentemente con los resultados y mejorando de manera continua.

Para ello hacen falta tres pilares fundamentales: la captura de datos, la extracción de características y la detección en tiempo real.

La otra gran ventaja de la Inteligencia Artificial aplicada a la ciberseguridad es la posibilidad de automatizar muchos de estos procesos, sin intervención humana, lo que supone una mejora importante para ayudar a mantener una red segura,

predecir patrones para evitar intrusiones, mejorar la gestión de identidades o el análisis masivo de muestras de *malware*.

Sin embargo, la otra cara de la moneda es que la Inteligencia Artificial también se puede convertir en una herramienta potente para los ciberdelincuentes al aumentar la eficacia y complejidad de sus ataques. Según el último índice de referencia cibernética internacional de Neustar, el 82% de los responsables de áreas de Seguridad de la Información (*Chief Information Security Officer*, CISO) de Seguridad de la organización (*Chief Security Officer*, CSO) a nivel mundial están preocupados ante la expectativa de que los ciberdelincuentes utilicen la inteligencia artificial contra su organización. Los impactos más temidos son la pérdida de datos y de la confianza de sus clientes.

Los algoritmos más utilizados para la detección de anomalías en ciberseguridad son los algoritmos de *Deep Learning*, que permiten la reacción ante los ataques en tiempo real. Con la ayuda de la inteligencia artificial los equipos de prevención de ciberseguridad pueden automatizar una gran parte de sus tareas, liberando más tiempo para focalizarse en estrategias defensivas. Los ataques sofisticados, basados en algoritmos de *machine learning*, que permitirían identificar vulnerabilidades y automatizar la selección de víctimas selectivas, aún no son muy frecuentes, pero los cibercriminales se están acercando a este escenario. De ahí que los equipos de seguridad deben ser capaces de combinar la inteligencia humana con las soluciones de inteligencia artificial para prevenir los riesgos de manera proactiva.

Seguridad y prevención del crimen

La analítica predictiva y otros tipos de herramientas basadas en inteligencia artificial se está aplicando hoy en día de manera eficiente por parte de la Policía local o estatal en muchos países. A partir de datos de movilidad, cambios en el tráfico, clima, eventos locales, geolocalización, datos históricos de incidencia de crímenes en el pasado, así como información socioeconómica y financiera, se puede predecir con bastante exactitud cuándo y dónde hay más probabilidad que se produzca, por ejemplo, un atraco a determinadas horas en un barrio en concreto. En muchas ciudades, el uso de cámaras es habitual y el análisis *a posteriori* de las imágenes de un evento criminal es clave en la mayoría de los casos para detener a los criminales, aunque no se usan tanto de manera predictiva. También es creciente el uso de drones en tareas de vigilancia y patrulla para detectar actividades criminales, siendo útil para reforzar la seguridad en puertos, aeropuertos, instalaciones industriales y otras infraestructuras críticas.

En países como Singapur o China, el uso de cámaras de vigilancia ha ido un paso más allá y el utilizar el análisis de imágenes en tiempo real para monitorizar cada actividad individual en la prevención de actos criminales, clasificando a los ciudadanos por niveles de peligrosidad y, a menudo, poniendo en cuestión los límites de la privacidad e intimidad personal.

Estos modelos se van perfeccionando cada vez con más datos, de manera que las Agencias de Inteligencia Estatales, accediendo a información de llamadas de teléfono, geolocalización, imágenes y al rastro digital de determinados sospechosos, pueden mejorar mucho en la predicción de estos ataques. Llevando estas aplicaciones a un terreno más sofisticado, la combinación de Inteligencia Artificial y Teoría de Juegos permite predecir ataques terroristas con mucha mayor precisión.

Aplicaciones en el campo de la inteligencia estratégica

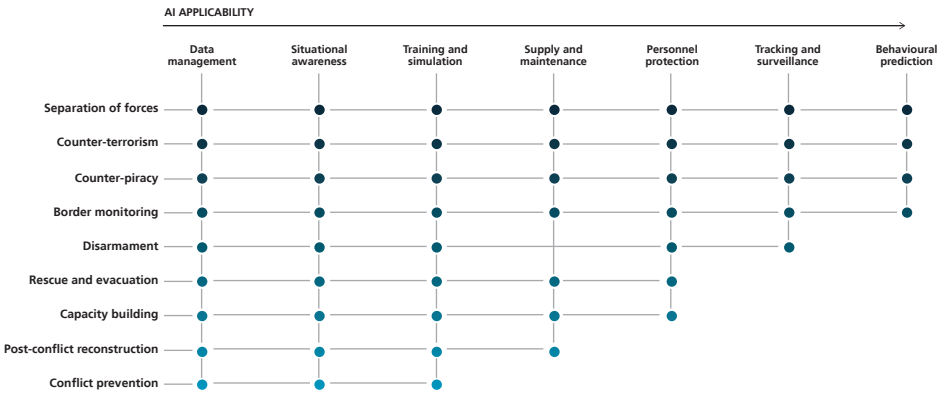
Poco se habla, sin embargo, en relación con la aplicación de la Inteligencia Artificial en el ámbito de la inteligencia estratégica. En este campo, el análisis de escenarios, factores e impacto supone la identificación de relaciones complejas, más propio de una futura inteligencia artificial general que la actual específica.

Cuando se tratan problemas en el mundo real, los resultados de nuestras acciones normalmente no se pueden saber con certeza. Esto se debe a que, por una parte, bien porque el mundo no es completamente observable, bien porque nuestro modelo del mundo es incompleto. Para poder tomar decisiones correctas en estas condiciones se debe incorporar la probabilidad en el proceso de razonamiento. En este contexto, los modelos Bayesianos se utilizan para estructurar y formalizar este proceso.

Los sistemas actuales de inteligencia artificial son capaces de leer los datos, procesarlos y desarrollar modelos predictivos, pero muchas veces no son interpretables. El ejemplo más actual son los modelos de *Deep Learning*, que sabemos que funcionan acertando sus predicciones con una gran fiabilidad, pero sin embargo no se sabe exactamente cómo funcionan y por tanto no son interpretables hoy en día. Por ello, el papel del ser humano es aún vital para la interpretación de los datos y para la toma de decisiones razonada. Las limitaciones actuales del uso de la inteligencia artificial residen en su especialización en problemas concretos o muy dependientes del contexto. El progreso en un dominio no es directamente trasladable a otro dominio distinto e incluso cercano.

Algunos beneficios, que sin duda aporta la inteligencia artificial, se encuentran en el ámbito de la detección, preparación y protección. Ello se traduce en la capacidad de integrar múltiples fuentes de datos para poder definir mejor una dinámica de crisis o conflicto en una región. Aunque muchas causas y características de conflictos son multifactoriales y difíciles de prever, sí que pueden complementar a los sistemas actuales de monitorización. El análisis de propensión a crisis y conflictos con *Big Data* e inteligencia artificial puede reducir los tiempos de reacción entre la detección y la acción temprana.

Un análisis detallado de las publicaciones científicas académicas sobre este campo en todo el mundo realizado por la editorial *Elsevier* permite ver los ámbitos principales de aplicación de la inteligencia en este campo tan complejo, cubriendo desde la prevención de conflictos, la piratería, el desarme, la evacuación, el contraterrorismo, etc.



Fuente: Scopus/Elsevier 2018

Consideraciones finales

Queramos o no, la inteligencia artificial está aquí para quedarse. Analistas como Gartner[38] predicen que en el 2020 las tecnologías de inteligencia artificial estarán embebidas en la mayoría de los productos de *software* y en los servicios, lo que de manera inevitable cambiará la manera en la que vivimos, trabajamos y hacemos negocios.

38 Para una información más detallada ver el informe "Predicts 2021: AI and the Future of Work"

A pesar de que estamos antes las primeras etapas de su desarrollo, las tecnologías de *Machine learning* o *Deep learning* ya han mostrado su eficacia y beneficios en multitud de industrias y empresas de diferentes sectores (banca, compañías aseguradoras, comercio minorista, producción, logística, salud, educación, ciberseguridad …).

La expansión masiva de la aplicación de la inteligencia artificial traerá sin duda nuevas oportunidades y nuevos riesgos, por lo que la aplicación de estas nuevas técnicas en los ámbitos de la predicción y prevención de la seguridad en todas sus dimensiones será de gran importancia.

La inteligencia artificial tiene un gran potencial para afectar a los equilibrios de poder tanto al nivel de la economía global como en el de la competición militar. Su desarrollo plantea varios retos y cuestiones en áreas como las relaciones internacionales, la privacidad y la seguridad.

La manera en cómo los países se diferencien competitivamente, los nuevos equilibrios políticos y las alianzas entre el sector público y privado determinarán los nuevos liderazgos.

De momento, la batalla por el dominio tecnológico se centra hoy en día entre China y EEUU, estando Europa aún rezagada. Sin embargo, el foco para Europa está en garantizar los principios éticos de la inteligencia artificial antes de que se desarrolle sin control, evitando la dictadura de los algoritmos y su discriminación, abordando el problema de la supresión de puestos de trabajo como consecuencia de la robotización, así como la protección de la privacidad de datos.

Estas diferencias en los valores pueden impactar de manera diferente en el desarrollo de cada país, que podrá ser distinto según sus prioridades (límites a la privacidad, intimidad, la consideración de derechos civiles digitales, etc.). Estamos ante un futuro incierto en el que la manera en la que seamos capaces de conciliar estos avances tecnológicos con el progreso económico y social es lo que va a definir un nuevo orden mundial y una nueva sociedad.

Carme Artigas es experta internacional en Big Data e Inteligencia Artificial. Es la cofundadora de Synergic Partners, una compañía creada en 2006, pionera en Big Data a nivel europeo y que fue adquirida por Telefónica en el 2015. Ha ejercido como CEO desde su fundación hasta la reciente integración de la compañía en el Grupo Telefónica a finales del 2018, y es miembro de Data Innovation Network de la Universidad de Columbia IDSE (NYC), así como embajadora por la Universidad de Stanford del Programa Women in Data Science. También pertenece a la Junta Directiva de la Asociación Española de Directivos (AED) y es ponente en distintos foros nacionales e internacionales de Big Data, como Strata + Hadoop World, así como profesora colaboradora en varios programas Máster en nuevas tecnologías, Big Data e innovación. Previamente a la fundación de Synergic Partners, fue CEO de Ericsson Innova, primer fondo de capital riesgo de Ericsson a nivel europeo, y presidenta de WIVA (International Wireless Internet Venture Association).

Aprendizaje profundo en el área de Inteligencia

Francisco Soler Flores

Resumen

Los trágicos acontecimientos ocurridos en España, EEUU y algunos países de la UE a principios de siglo marcan un antes y un después en el tratamiento de la seguridad y el funcionamiento de las agencias de Inteligencia en todo el mundo. El siglo XXI comienza siendo el de los datos y su tratamiento masivo, el *Big Data* en la era de Internet abre las posibilidades de la explotación de la información de forma masiva y las grandes compañías de Internet contribuyen ofreciendo recursos y soluciones para este propósito. La Inteligencia Artificial, aunque surgida a finales de los años 50, experimenta en la segunda década del siglo XXI un espectacular avance apoyada en el *Big Data*, el progreso en la capacidad de cómputo y el Aprendizaje profundo. El objetivo de este capítulo es conocer qué es el Aprendizaje profundo, cómo funciona y qué posibilidades ofrece dentro del ámbito de la seguridad y las agencias de Inteligencia.

Inteligencia Artificial

La Inteligencia Artificial es un campo con un rápido crecimiento y con implicaciones en el campo de la seguridad nacional que está captando la atención de empresas internacionales, líderes en el sector comercial, intelectuales de la defensa, politólogos, etc. Un ejemplo claro es el Departamento de Defensa de EEUU, el cual está desarrollando aplicaciones de Inteligencia Artificial para un amplio rango de funciones. Pero esto, por supuesto, no es algo aislado. El 20 de Julio de 2017, el Gobierno Chino lanzó una estrategia en que detalló su plan para conseguir el liderazgo en IA para 2030; dos meses después, Vladimir Putin anunció públicamente la intención de Rusia de trabajar en tecnologías de IA. Por otro lado, Elon Musk, en su momento CEO de SpaceX y fundador de OpenaAI y Tesla, envió una carta firmada por 116 líderes internacionales en el sector de la tecnología pidiendo a las Naciones Unidas la prohibición de las armas autónomas. En la carta advertía de que las armas autónomas alimentadas por la IA permitirán que los conflictos armados se produzcan a una escala nunca vista antes, con una celeridad inimaginables para las personas, apelando así a los medios a prevenir una carrera de armamentos y proteger a los civiles de un posible mal uso.

La Comunidad de Inteligencia de EEUU analiza millones de registros de datos al día para resolver crímenes y prevenirlos de manera proactiva. Sin embargo, el proceso de revisión de estos datos, que en algunos casos incluye imágenes, imágenes satelitales y documentos, requiere mucho tiempo a los analistas. Intentando responder a esto, la comunidad de inteligencia de EEUU lleva ya tiempo desarrollando metodologías de Inteligencia Artificial para ayudar a los analistas a evaluar los datos de manera más eficiente y efectiva de la mano, por ejemplo, de la Agencia de Proyectos de Investigación Avanzados de Defensa, DARPA, responsable del desarrollo de nuevas tecnologías para uso militar.

El área de inteligencia artificial, dentro de las ciencias de la computación, existe desde mediados del siglo XX, formalmente cuando en el verano de 1956 se celebró la famosa conferencia de Dartmouth cuyos participantes fueron los principales investigadores en IA durante los años sucesivos. Aunque ya Alan Turinng especulara sobre la posibilidad de crear máquinas que pensaran en 1950, no es hasta este evento cuando se consolida el concepto en la comunidad científica.

Así John McCarthy, junto con Marvin Minsky, acuñó el término de Inteligencia Artificial en la década de los 50 y las Redes Neuronales aparecen ya con Frank Rosenblatt en 1958 con el diseño de un prototipo: el perceptron. De esta forma las ideas clave de las Redes Neuronales Profundas sobre lo que se hablará en el siguiente punto, ya se conocían a finales del siglo XX, así como algunos de los algoritmos fundamentales.

De esta forma se define la inteligencia Artificial como la disciplina que tiene como objetivo desarrollar máquinas que se comporten como si fueran inteligentes, que desde el punto de vista de la mente humana suele entenderse como adquisición de conocimiento y sus diferentes formas, la experiencia, o la capacidad para resolver problemas nuevos.

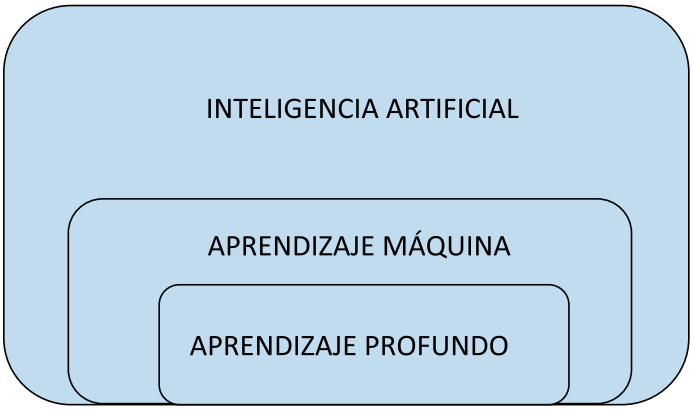
Actualmente, esta área ha atraído una espectacular atención mediática debido a los importantes avances conseguidos en el contexto de la investigación científica y de las grandes tecnológicas de internet, con ejemplos como la utilización de Aprendizaje profundo o Redes Neuronales Profundas para clasificar imágenes con un éxito rotundo hasta superar actualmente al humano en esta tarea o el uso de modelos de Aprendizaje por Refuerzo para ganar en juegos extremadamente complejos como el Go.

Existen diferentes teorías y modelos alrededor de la Inteligencia Artificial como la Lógica Proposicional, los modelos de razonamiento bajo incertidumbre como las Redes Bayesianas, el Aprendizaje máquina, las Redes Neuronales o los modelos de aprendizaje por refuerzo.

Centraremos este capítulo en las Redes Neuronales y el Aprendizaje profundo que han supuesto un gran avance en el área de la inteligencia artificial en estos últimos años, la cual en este mismo momento está mostrando una gran capacidad y posibilidades debido al aumento de los datos accesibles, al *Big Data*, así como al aumento de la capacidad computacional de los ordenadores gracias al uso y la idoneidad de las Unidades de Procesamiento Gráfico para el cálculo matricial o tensorial, punto importante del procesamiento de redes neuronales.

La Inteligencia Artificial engloba el área de *Machine Learning* donde se incluye a su vez el *Deep Learning* (Fig 1).

Figura 1. Aprendizaje profundo e Inteligencia Artificial



Aprendizaje máquina, redes neuronales y aprendizaje profundo

El aprendizaje es una tarea compleja para los seres humanos. Cuestiones como el aprendizaje de vocabulario, términos técnicos o algo tan aparentemente sencillo como un poema puede resultar complicado para una persona. Sin embargo, los

ordenadores pueden ayudarnos a resolver tareas de aprendizaje aún más complejas como las que requieren de capacidades matemáticas y análisis de datos.

Arthur Samuel (IBM) definió el aprendizaje automático o *Machine Learning* como el campo de estudio que dota a los ordenadores de la capacidad de aprender a resolver problemas para los que no han sido explícitamente programados. Así, en el área de Inteligencia Artificial, el aprendizaje se entiende como un proceso por el cual un ordenador es capaz de mejorar su habilidad en la resolución de un problema a través de la adquisición de conocimiento que obtiene a través de la experiencia. Esta área se estructura principalmente en:

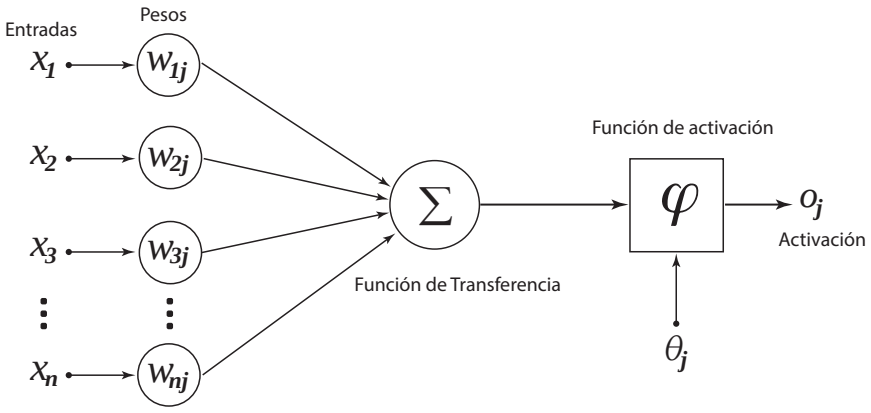
- Aprendizaje supervisado: construido a partir de un conjunto de datos de entrada o conjunto de entrenamiento con ejemplos de cada una de las clases que se pretenden modelizar
- Aprendizaje no supervisado: su objetivo es encontrar patrones en los datos de entrada que permitan construir un modelo de distribución de probabilidad.

Las Redes Neuronales fueron desarrolladas para intentar simular el sistema nervioso humano con tareas de Aprendizaje Máquina tratando las unidades computacionales con un modelo de aprendizaje similar al de las neuronas humanas. Así, las Redes son formalmente un modelo matemático-computacional inspirado en una simplificación del comportamiento del cerebro humano. Esta simplificación se basa en las redes de neuronas biológicas con un número estimado de cien mil millones de neuronas conectadas a través de los llamado *axón* o fibra nerviosa que permite transmitir las señales eléctricas entre las neuronas llamadas dendritas. Cada neurona puede establecer conexión desde con una docena de neuronas hasta con cientos de miles.

Las redes neuronales son teóricamente capaces de aprender cualquier función con suficientes datos de entrenamiento. Así, los componentes que una red neuronal artificial intenta emular de la biológica son:

- Procesamiento paralelo
- Memoria distribuida
- Adaptabilidad al entorno

Figura 2. Esquema Red Neuronal



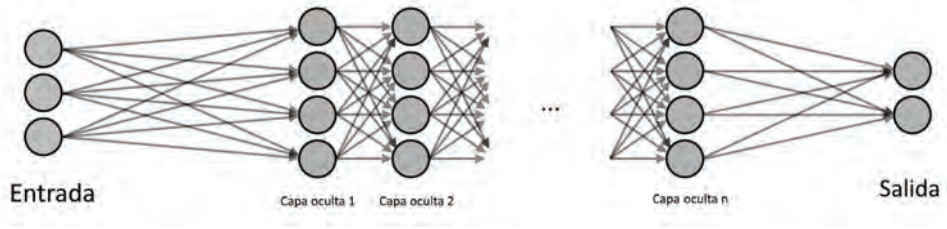
En un lenguaje sencillo, y siguiendo el esquema de la Figura 2, una sola neurona pasará un mensaje a otra neurona a través de esta interfaz si la suma de las señales de entrada ponderadas de una o más neuronas (suma) es suficiente (excede un umbral) para provocar la transmisión del mensaje. Esto se denomina activación cuando se excede el umbral y el mensaje se pasa a la siguiente neurona. El proceso de suma puede ser matemáticamente complejo. La señal de entrada de cada neurona es en realidad una combinación ponderada de potencialmente muchas señales de entrada, y la ponderación de cada entrada significa que esa entrada puede tener una influencia diferente en cualquier cálculo posterior y, en última instancia, en la salida final de toda la red. Además, cada neurona aplica una función o transformación a las entradas ponderadas, lo que significa que la señal de entrada ponderada combinada se transforma matemáticamente antes de evaluar si se ha excedido el umbral de activación. Esta combinación de señales de entrada ponderadas y las funciones aplicadas suelen ser lineales o no lineales.

Así el funcionamiento simplificado de una red neuronal es, a partir de un conjunto de variables o datos de entrada, estimar una salida. El enfoque podría asemejarse a un clásico "fuerza bruta" porque el ordenador buscaría ajustar los parámetros que se combinarían con los datos de entrada para proporcionar la salida, pero diferentes algoritmos de optimización ayudan a encontrarlo de manera óptima utilizando métodos clásicos del Análisis Numérico como el del Gradiente Descendiente o

modificaciones y evoluciones de los mismos como *Batch Gradient Descent, Mini-batch Gradient Descent* o *Stochastic Gradient Descent*.

El enfoque de red neuronal clásico no responde a problemas complejos que el Aprendizaje profundo o las Redes Neuronales Profundas pueden abordar. Si la red neuronal tiene varias capas ocultas o intermedias (Figura 3) entre la entrada y la salida se denominan redes neuronales profundas. Este número de capas puede llegar a ser en algunos casos incluso superior a cientos de miles. De esta forma, las estructuras algorítmicas permiten modelos que están compuestos por múltiples capas de procesamiento para aprender representaciones de datos con múltiples niveles de abstracción que realizan una serie de transformaciones lineales y no lineales de forma que a partir de los datos de entrada generan una salida próxima a la generada. El entrenamiento –aprendizaje de los parámetros de estas redes– requiere de conjuntos de datos muy grandes y de considerables capacidades de cómputo.

Figura 3. Esquema Red Neuronal Profunda



Para responder a esta necesidad de cálculo (Redes Neuronales & Deep Learning, Fernando Berzal 2018) de parámetros en redes neuronales, formadas por grandes cantidades de neuronas que operan en paralelo, la arquitectura paralela de una GPU (Unidad de procesamiento gráfico) es más eficiente que una CPU (Unidad Central de Procesamiento clásica) permitiendo implementar algoritmos que procesan grandes bloques de datos en paralelo. La estandarización en 2007 de la API de programación de las GPUs NVIDIA, denominada CUDA (*Compute Device Architecture*) reduce la ejecución de los cálculos de semanas a horas. Además, la librería de NVIDIA CUDNN (CUDA *Deep Neural Networks*) permite ejecutar algoritmos de Aprendizaje profundo usando sus GPUs, algo que han aprovechado entornos de trabajo como *TensorFlow* para ofrecer la posibilidad de desarrollar este tipo de modelos de manera más ágil y sencilla.

Así, de cara a desarrollar soluciones en el ámbito del Aprendizaje profundo, las grandes compañías de internet como Google, Facebook, Microsoft o Amazon han desarrollado entornos de trabajo como *keras, theano, mxnet, tensorflow* que hacen más cómodo y sencillo el trabajo con redes neuronales profundas. Además, la computación en la nube proporciona la infraestructura necesaria bajo demanda como alternativa de pago por uso de una infraestructura física en las instalaciones de las compañías. Por otro lado, estas mismas empresas han contribuido a proporcionar entornos de desarrollo como *Azure Machine Learning* o *Google Colaboratory*

La combinación de las neuronas y las diferentes conexiones entre capas proporcionan la topología o arquitectura de la red. Existen diferentes topologías o arquitecturas de redes neuronales profundas que se utilizan para diferentes propósitos. Algunos ejemplos son:

- Redes convolutivas o convolucionales (CNN): clasificación de imágenes, identificación de objetos y segmentación semántica (clasificación *pixel* a *pixel*)
- Redes secuenciales o recurrentes (RNN), lstm: análisis de series temporales, detección de anomalías, predicción de texto, *automatic translation*
- Redes Generativas Antagónicas, de adversarios o redes GAN. Permiten generar datos de manera sintética tales como imágenes nuevas a partir de un conjunto de entrenamiento
- Las recientemente populares redes tipo *Transformers con Megatron-Turin (Efficient Large-Scale Language Model Training on GPU Clusters Using Megatron-LM*, Deepak Narayanan *et al*. 2021) de Microsoft y NVIDIA o GPT3 de OpenAI (*Language Models are Few-Shot Learners*, Tom B. Brown et al 2020) como modelos de lenguaje o las redes de cápsulas que permiten identificar objetos en imágenes de manera completa y no por aparición parcial. (Zhang, Q., Wu, Y. N., & Zhu, S. C. (2018, June). Interpretable *convolutional neural networks*. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 8827-8836).)

Ante la complejidad computacional del entrenamiento de una red neuronal profunda, existen además alternativas al entrenamiento completo de la red con técnicas de *Transfer Learning* que permite la reutilización de lo aprendido en un entrenamiento previo y tras una simplificación de la arquitectura reducir enormemente el coste computacional.

En estos últimos años se ha generado una gran cultura de publicaciones abiertas de forma que muchos investigadores publican sus resultados inmediatamente sin esperar la aprobación de la revisión por pares habitual, por ejemplo arxiv.org de la Universidad de Cornell, lo que ayuda a que se genere una gran cantidad de *Open Source* asociado a estos trabajos y accesible a través de plataformas de la comunidad como Github. Además de las grandes tecnológicas como Google, Facebook, Microsoft o Amazon, universidades como Berkeley, Standford o Toronto son especialmente activas.

Aprendizaje profundo en los servicios de inteligencia y detección de comportamientos anómalos

La disponibilidad de datos, en este caso de grabaciones de video, video en *streaming*, registros de personas en diferentes actividades, comentarios en redes sociales, captación de audio, escritos en medios de comunicación o redes sociales hacen posible que algunos de los ejemplos de aplicaciones de las Redes Neuronales Profundas como la conversión de voz a texto, el reconocimiento facial, la detección de emociones, la clasificación de documentos, la traducción automática o la generación sintética de imágenes sean posibles y así se pueda proporcionar a las Agencias de Inteligencia herramientas para hacer más eficiente su trabajo, ofreciendo soluciones para problemas no resueltos o resolver tareas habituales de los analistas de Inteligencia de manera automática.

La posibilidad de las Redes Neuronales Profundas de identificar objetos en imágenes y reconocer patrones permite abordar cuestiones como la detección e identificación de presencia y un ejemplo claro de esta aplicación es el conocido sistema de vigilancia del gobierno chino que ayudado de más de 600 millones de cámaras, su sistema SkyNet lanzado originalmente en 2015 con el objetivo de localizar y capturar fugitivos y políticos corruptos, que evolucionó transformándose en algo mucho más avanzado, permitiendo identificar el sexo de la persona, edad, color de piel, características de la ropa y hasta rasgos únicos. Además, es capaz de reconocer vehículos por marca, modelo, color, así como tipo de vehículo y saber si está siendo conducido o aparcado. Esto no deja de ser aplicaciones del Aprendizaje profundo intuitivas enfocadas a problemas de seguridad e Inteligencia pero, adicionalmente, se puede hablar de aplicaciones como la detección de presencia o identificación de escenarios anómalos a partir de imágenes, combinando la capacidad del Aprendizaje profundo para extraer características de las imágenes y clasificarlas. De la misma forma, esta extracción de características aplicadas a los documentos o las transcripciones de

voz permite la identificación de textos o conversaciones pertenecientes a categorías lejanas a las habituales o incluidas dentro de un área de riesgo.

Así, las posibilidades de predicción de las Redes Neuronales Recurrentes abren la viabilidad de su aplicación a la estimación del siguiente paso dentro de un contexto de alerta, emergencia o peligro y a su vez permitiría establecer escenarios para solventar esto.

La detección de anomalías en series temporales es otro de los campos de aplicación de las redes neuronales profundas. La detección de *outliers* o valores anómalos es un campo de trabajo habitual de los analistas de inteligencia, ser capaz de identificar estos valores es de un gran valor enfocado a la predicción y toma de decisiones. La aplicación de los modelos de Aprendizaje profundo en los sectores tecnológicos y empresariales abre la puerta a la aplicación de nuevos casos de uso en el área de inteligencia. Por ejemplo, la conducción autónoma de vehículos ha de tener en cuenta cuestiones como la identificación de carril, la identificación de peatones u obstáculos en la carretera y la toma de decisiones en intersecciones o semáforos. Estos son hoy en día problemas resueltos, encontrándose la conducción autónoma actualmente en el punto donde se decide entre las decisiones automáticas y la ética, problema también presente en la aplicación de la IA en entornos de Inteligencia. Además, los avances estos últimos años han sido muchos, como demuestra la reciente patente de Google para la detección de anomalías en video.

Con todo esto, los líderes de las agencias de Inteligencia ven la Inteligencia Artificial como una herramienta para aumentar las capacidades de los analistas humanos, no como un reemplazo para ellos. Así, Mike Bender, director del *National Security Agency's Lab for Analytic Science*, afirmó que las posibles aplicaciones de la IA están superando las capacidades de muchos analistas, quienes necesitarán capacitación para aprovechar los conjuntos de herramientas.

Así, teniendo en cuenta las capacidades del Aprendizaje profundo y sus utilidades reconocidas en la comunidad científica y área empresarial podemos ver aplicaciones claras dentro del área de seguridad e inteligencia como:

- Vigilancia automática: el avance de la videovigilancia y las posibilidades del Aprendizaje profundo se ven reflejadas en sistemas comerciales que por ejemplo ofrece Amazon con sus Deep Lens, las cuales permiten ejecutar modelos de aprendizaje profundo de manera local en sistemas de cámara con la posibilidad sencilla de, adicionalmente, tomar acciones a partir de las anomalías identificadas

- Extracción de información anómala: las ventajas del Aprendizaje profundo en la extracción de características en imágenes y documentos o equivalentemente transcripciones de voz permite la identificación de intenciones, análisis de sentimiento y encontrar parecidos, diferencias o anomalías entre estos conjuntos de datos.
- Enriquecimiento de información: la información deteriorada o de calidad baja es posible enriquecerla o reconstruirla utilizando técnicas de Aprendizaje profundo. Los casos de audio o imágenes son ejemplos resueltos actualmente
- Anomalías en datos de fuentes abiertas: el análisis de datos de fuentes de datos abiertas tales como las redes sociales permite la identificación de registros anómalos que la seguridad puede tener en cuenta de manera automática.
- Los modelos aprendizaje por refuerzo basados en Aprendizaje profundo permiten aprender comportamientos en base a estímulos dirigidos y reconducir conductas de riesgo

Los avances logrados en estos últimos años en el área de Aprendizaje profundo y la gran inversión por parte de gobiernos y compañías especializadas en el área de Inteligencia Artificial no hacen más que alentar y mostrar esperanza en cómo la Inteligencia Artificial va a ayudar a resolver tareas dentro del área de Inteligencia de manera más eficiente y con la posibilidad de automatización para simplificar el trabajo de los analistas, reducir su carga de trabajo y destinar su esfuerzo a tareas que todavía han de estar en manos de personas. Además, los continuos avances y progresos permitirán sin duda mejoras en la eficiencia, superando en muchos casos al humano, pero sin duda no sustituyéndolo por ahora, en esta época de la Inteligencia Artificial Débil o Estrecha de propósito específico y todavía sin haber llegado a una Inteligencia Artificial General o fuerte, campo actualmente emergente o la distante Súper Inteligencia Artificial.

Bibliografía

Allen, G., & Chan, T. (2017). *Artificial intelligence and national security*. Cambridge (MA): Belfer Center for Science and International Affairs.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. AI magazine, 27(4), 12-12.

A. M. Turing (1950). Computing machinery and intelligence-AM Turing. Mind, 59(236), 433. Torres, J., & Torres, T. V. (2018). DEEP LEARNING Introduccion practica con Keras. *Barcelona, Creative Commons.* Publishing, Amazon

McCarthy, J., & Hayes, P. J. (1981). Some philosophical problems from the standpoint of artificial intelligence. In Readings in artificial intelligence (pp. 431-450)

Flynn, J. R. (2007). What is intelligence?: Beyond the Flynn effect. Cambridge University Press

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105)

Chen, J. X. (2016). The evolution of computing: AlphaGo. Computing in Science & Engineering, 18(4), 4-7

Soler Flores, F. (2014). Estimación de sucesos poco probables mediante redes bayesianas. Tesis doctoral

F. Berzal (2018). Redes Neuronales y Deep Learning. Independently published (14 de noviembre de 2018)

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Deep learning (Vol. 1). Cambridge: MIT press

Ertel, W. (2017). Reinforcement Learning. In Introduction to Artificial Intelligence (pp. 289-311). Springer, Cham

Samuel, A. L. (1988). Some studies in machine learning using the game of checkers. II—recent progress. In *Computer Games I* (pp. 366-400). Springer, New York, NY.

Mitchell, T. M. (1997). Machine learning. 1997. *Burr Ridge, IL: McGraw Hill, 45* (37), 870-877.

García, T. R., Cancelas, N. G., & Soler-Flores, F. (2014). The artificial neural networks to obtain port planning parameters. *Procedia-Social and Behavioral Sciences*, 162, 168-177.

Flórez, R., & Fernández, F. J. (2008). Las redes neuronales artificiales, fundamentos teóricos y aplicaciones prácticas. *La Coruña, Oleiros, España: Netbiblo.*

RAMON Y CAJAL S. Contribucion al conocimiento de la neuroglia del cerebro humano. Trab Lab Invest Biol (Madrid) 11: 255–315, 1913.

Aggarwal, C. C. (2018). Neural networks and deep learning. *Cham: Springer International Publishing.*

Liu, L., Wu, Y., Wei, W., Cao, W., Sahin, S., & Zhang, Q. (2018, July). Benchmarking Deep Learning Frameworks: Design Considerations, Metrics and Beyond. In *2018 IEEE 38th*

International Conference on Distributed Computing Systems (ICDCS) (pp. 1258-1269). IEEE.

Salas-Molina, F., & Pla-Santamaria, D. (2018, February). Coding oriented learning in economics, business and finance. In *Modelling in Science Education and Learning* (Vol. 11, No. 1, pp. 55-64). Universitat Politècnica de València

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural networks, 61, 85-117.

Zhang, Q., Wu, Y. N., & Zhu, S. C. (2018, June). Interpretable convolutional neural networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 8827-8836).)*

Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. Journal of Big Data, 3(1), 9.

Nouwens, M., & Legarda, H. (2018). Emerging Technology Dominance: What China's Pursuit of Advanced Dual-Use Technologies Mean for the Future of Europe's Economy and Defence Innovation. *International Institute for Strategic Studies/Mercator Institute for China Studies China Security Project*, 2018-12.

González M. N.., Soler, F.., Camarero, A. C., & Ansorena, Í. L. (2012). Treatment of Outliers to Study Railway Vibrations Transmission. *Ingeniería y ciencia*, (16), 191-219.) (Ly, L., & Tsai, Y. H. R. (2018). Autonomous Exploration, Reconstruction, and Surveillance of 3D Environments Aided by Deep Learning. *arXiv preprint arXiv:1809.06025.*

Koizumi, Y., Saito, S., Uematsu, H., Kawachi, Y., & Harada, N. (2019). Unsupervised Detection of Anomalous Sound Based on Deep Learning and the Neyman–Pearson Lemma. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27(1), 212-224.

Jain, A., Gupta, N., Mujumdar, S., Mehta, S., & Madhok, R. (2018, July). Content Driven Enrichment of Formal Text using Concept Definitions and Applications. In *Proceedings of the 29th on Hypertext and Social Media* (pp. 96-100). ACM.).

Sun, X., Zhang, C., Ding, S., & Quan, C. (2018). Detecting anomalous emotion through big data from social networks based on a deep learning method. *Multimedia Tools and Applications*, 1-22.

Pérez-Benito, F. J., Villacampa-Fernández, P., Conejero, J. A., García-Gómez, J. M., & Navarro-Pardo, E. (2019). A happiness degree predictor using the conceptual data structure for deep learning architectures. *Computer methods and programs in biomedicine*, 168, 59-68.

Kong, X., Ma, K., Hou, S., Shang, D., & Xia, F. (2018). Human Interactive Behaviour: A Bibliographic Review. *IEEE Access.*

Efficient Large-Scale Language Model Training on GPU Clusters Using Megatron-LM, Deepak Narayanan *et al.* 2021

Language Models are Few-Shot Learners, Tom B. Brown *et al* 2020

Francisco Soler Flores es doctor Ingeniero Informático en el área de Inteligencia Artificial y Licenciado en Matemáticas. Es profesor asociado en las áreas de Computación e Inteligencia Artificial en la Universidad Carlos III, profesor en diferentes Masters en el área de Aprendizaje Profundo en la Universidad de Alcalá, profesor en la Universidad Internacional de la Rioja y, ha sido investigador y profesor en grado, master y cursos de especialización en diferentes universidades públicas y privadas como la Universidad Politécnica de Madrid, Universidad Europea o la Universidad de Almería y profesor invitado en la Universidad de Valladolid o la Universidad de Castilla La Mancha. Ponente invitado en congresos y jornadas en el ámbito de la Inteligencia Artificial, Analytics y Big Data. Cuenta con más de 150 referencias en congresos y revistas internacionales de prestigio. En el ámbito de la empresa privada, creador de la compañía de Software Libre y Open Source Insolux, lideró la creación y desarrollo del área de Big Data-Analytics en la consultora tecnológica Altran, siendo además responsable de Operaciones en la compañía especializada en Analytics e Inteligencia Artificial Tessella del mismo grupo, fue miembro del equipo ejecutivo de Inteligencia Artificial en everis Europa y actualmente forma parte de la división de Inteligencia Artificial en Accenture como AI Senior Manager en España, Portugal e Israel.

La revolución de los datos

Juan Murillo Arias

Resumen

La huella digital de ciudadanos y empresas es cada vez más extensa y granular. Permite no solo hacer análisis descriptivos de las interacciones que están dejando esta huella en forma de datos, sino que a menudo permite el pronóstico de su evolución futura con mayor o menor grado de acierto. Estas capacidades aumentadas dotan a las sociedades de capacidad de anticipación, pero a la vez abren la caja de los dilemas éticos cuando la huella que estamos analizando baja al detalle del dato personal. Además, esta huella se distribuye de forma desigual: en la dicotomía entre empresas y sector público resultan ser las primeras las que cuentan con fuentes de datos de mayor resolución espacial y temporal. También se abre una brecha entre sectores fragmentados, como el financiero, frente a sectores monopolísticos, como los servicios de redes sociales. Una desventaja similar podemos percibir entre sistemas sometidos a control democrático y estados autoritarios, contando los segundos con una mayor capacidad de concentración de información sobre los ciudadanos, y por tanto de poder y consolidación del autoritarismo.

Estelas digitales para pronosticar el futuro

Quienes formamos parte de la última generación que guardará memoria de la época previa a internet y a la telefonía móvil tenemos muy presente la larguísima lista de acciones que llevábamos a cabo sin asistencia tecnológica: desde comunicarnos por correo postal, a localizar un punto de encuentro en un callejero en papel, pagándolo todo con dinero en efectivo, compartiendo nuestras fotografías, lecturas y opiniones políticas sólo de forma presencial en un círculo reducido de familiares y amigos. Hoy en día, al llevar a cabo cada una de estas acciones individuales en el plano digital, dejamos un rastro de migas de pan, un flujo creciente de datos que describen con gran detalle nuestra actividad cotidiana, nuestros intereses y hasta nuestro poder adquisitivo o la capacidad de influencia que tenemos entre un conjunto de seguidores. Y aunque la huella del 99,9% de la población sea irrelevante a efectos de garantizar la seguridad nacional, la información agregada de la sociedad en su conjunto permite no solo hacer análisis descriptivos sobre cualquier sector de actividad, sino que posibilita anticipar tendencias respecto a evoluciones futuras. Al menos esta es la teórica meta que persiguen entidades como el IARPA.[39]

39 Iniciativa del Departamento Nacional de Inteligencia del gobierno de EEUU que impulsa la investigación en materia de prospectiva.

Y sin embargo es prudente mantener cierto escepticismo al respecto de estas nuevas capacidades, fundamentalmente por las siguientes razones técnicas:

La primera radica en las fuentes utilizadas: por un lado, los datos constituyen una sombra parcial de una realidad mucho más compleja. Además, los sistemas sobre los que dejamos huella digital son inconexos. Así, como individuos, la empresa que nos presta servicio de fibra óptica, televisión digital y telefonía móvil puede saber qué tipo de contenidos de ocio preferimos, o nuestra ubicación a lo largo del tiempo mediante la geoposición de nuestro terminal móvil (si le hemos otorgado consentimiento para ello); por su lado, la red social en la que compartimos las fotografías de nuestras últimas vacaciones conocerá nuestras preferencias de ocio, y la plataforma en la que adquirimos el último libro sabrá nuestros gustos literarios. Estos datos alimentarán modelos con los que los departamentos de inteligencia de negocio de todas estas empresas inferirán nuestra propensión al consumo y orientarán sus futuras acciones de carácter comercial. Sin embargo –por la citada dispersión de huellas, y porque todos estos servicios los usan distintos miembros de una misma unidad familiar– todas ellas sin excepción tendrán dificultades para tener una visión completa de nuestra actividad específica individual y de nuestras acciones futuras.

Ante esta visión incompleta se puede argüir que los errores de pronóstico –falsos positivos y falsos negativos a nivel individual– pierden importancia y se compensan unos a otros cuando se trata de sacar conclusiones a nivel macro sobre cambios de tendencia. Nos encontramos aquí frente a la segunda dificultad, ya no solo técnica sino también metodológica: inevitablemente y por su propia definición *un modelo es siempre una simplificación esquematizada de la realidad que pretende representar*. Debido a la mencionada parcialidad de las fuentes de datos empleadas, pero también por la limitada capacidad de los sistemas actuales de computación, reducimos la realidad a un número de variables mucho menor que el conjunto de todas las que en realidad tienen influencia sobre el fenómeno estudiado. Pensemos en la baja fiabilidad de los modelos numéricos actuales para pronosticar el tiempo meteorológico con un grado de incertidumbre aceptable a más de tres semanas vista: nos topamos con una realidad no lineal en la que una ligera desviación en las condiciones de contorno iniciales desemboca en fenómenos totalmente opuestos e igual de probables. De forma análoga pero aún más marcada, la realidad social y económica conforma un sistema gobernado por interacciones complejísimas y difícilmente modelables. En nuestros esfuerzos por simplificar este sistema tendemos a infravalorar la ocurrencia de *cisnes negros*, sucesos improbables, pero de altísimo impacto, como ha señalado *Nassim Nicholas Taleb* en sus múltiples obras. Tanto ante

la crisis financiera de 2008 como frente a la primavera árabe desencadenada en 2010 los datos estaban ahí, pero los sistemas de aviso fallaron a la hora de alertar e informar sobre la magnitud de estos fenómenos. Por último, los modelos basados en aprendizaje automático tienen una asombrosa facilidad para detectar correlaciones ocultas en una cantidad ingente de datos, pero también grandes dificultades para encontrar las causas de estas correlaciones, como muy bien explica Judea Pearl en *The Book of Why.*

Asimetrías en el acceso a fuentes de información

Incluso considerando las anteriores salvedades reconozcamos que, efectivamente, tenemos mucha mayor capacidad que nuestros predecesores para anticipar sucesos gracias a la explosión de datos y a las herramientas analíticas actuales. La clave entonces radica en alinear los intereses y capacidades de los tres sectores clave que pueden hacer posible que este potencial redunde en el máximo beneficio común: gobiernos, ciudadanos y empresas.

Para ello hay que tener en cuenta en primer lugar que la huella digital se distribuye de forma desigual en distintas dimensiones: en la dicotomía entre empresas y sector público resultan ser las primeras las que cuentan con fuentes de datos de mucha mayor resolución espacial y temporal que los gobiernos. La Comisión Europea ha visto que esto en sí es un problema y creó en 2019 un grupo de trabajo del que formé parte para encontrar soluciones que alineen incentivos a la compartición de datos entre ambos sectores.

También se da una gran desigualdad entre sectores fragmentados –como el financiero o el energético– frente a sectores monopolísticos, como los servicios de redes sociales, cuyo número de usuarios se cuentan por cientos o miles de millones. A partir de huellas fragmentadas es más difícil obtener una visión de conjunto acertada y, sin embargo, si todos estamos de acuerdo en que en economía los monopolios distorsionan el mercado y éste pasa a ser cautivo, en el acceso a la información sus efectos son igualmente negativos, al dejar en manos de unos pocos decisores privados el control del conocimiento. La revolución digital ha tenido justo ese efecto, el de potenciar el surgimiento de grandes actores globales con mayor información que la mayoría de los gobiernos, cuya visión es ante todo local.

Por otro lado –desde el punto de vista de los derechos civiles– esta nueva capacidad abre la caja de pandora de los *dilemas éticos* cuando la huella que estamos analizando

baja al detalle del dato personal. La realidad es que las herramientas tecnológicas actuales harían las delicias de *Gerd Wiesler*, el agente de la Stasi protagonista de la película *La vida de los Otros*, que perseguía disidentes en la extinta RDA. Aquí las sociedades han de encontrar el punto de equilibrio entre seguridad frente a privacidad. En un artículo ya antiguo *Anja Kaspersen*, directora de Geopolítica y Seguridad del *World Economic Forum*, se preguntaba si era lícito que los gobiernos buceasen en la huella digital de los ciudadanos para garantizar su seguridad, cuando es patente por un lado que los grupos terroristas se valen de los nuevos canales de comunicación para coordinar sus acciones y captar acólitos, pero también es evidente que estas nuevas capacidades ponen en manos de gobiernos autoritarios herramientas de control social inéditas.

Esto nos lleva a hablar de otra inequidad más en el acceso a fuentes de datos: la que se da entre sistemas democráticos garantistas en materia de privacidad, y estados autoritarios, contando los segundos con una mayor capacidad de acceso y concentración de información, y por tanto de poder, dado que juegan en su geografía con una ventaja monopolística incluso superior a la que en sector privado tienen los grandes operadores de servicios digitales. Para salvar esta desventaja, la administración de George W. Bush puso en marcha en 2007 el programa secreto PRISM, en virtud del cual la NSA obtenía acceso –previo pago de millones de dólares– a los datos alojados en los servidores de Microsoft, Yahoo, Google, Facebook, PalTalk, AOL, Skype, YouTube y Apple, y cuya existencia fue filtrada por Edward Snowden en el año 2013. Cuando trascendieron los esfuerzos del gobierno de un país democrático como EEUU por emular el acceso a datos personales (que damos por hecho que se dan por parte de regímenes no democráticos) el escándalo mediático fue mayúsculo, pero en su reflejo hacia la presar el foco del debate viró –en un uso magistral de los marcos conceptuales descritos por *George Lakoff*– y no se habló tanto del comportamiento del gobierno ni de las compañías que habían vendido los datos de sus usuarios, sino de la deslealtad del exempleado de la CIA y de la NSA al hacer pública esta información; tanto es así que hoy no hablamos del "*caso PRISM*", sino del "*caso Snowden*".

Tras aquellos acontecimientos el *Pew Research Center* preguntó directamente a ciudadanos estadounidenses si su comportamiento en el uso de servicios digitales había cambiado a raíz de saber que el gobierno había accedido a esta traza digital, y la respuesta fue afirmativa en un 25% de los casos. Este mismo instituto hizo extensiva otra encuesta al resto del mundo. La pregunta esta vez se dirigía a pulsar en qué medida la reputación de EEUU había sido afectada por el escándalo, y la

conclusión fue que el impacto había sido menor en la opinión pública. No fue sin embargo desdeñable este daño en la confianza entre aliados tradicionales, como han sido EEUU y Europa: en julio de 2020 el Tribunal Superior de Justicia de la UE falló a favor del abogado austríaco Maximilian Shrems, que había denunciado que Facebook ejercía transferencias internacionales de datos de sus usuarios hacia EEUU, una jurisdicción en la que la Ley de Inteligencia e Investigación en el Extranjero (FISA por sus siglas en inglés) y las órdenes ejecutivas del gobierno atribuyen a la NSA, la CIA y el FBI capacidades de vigilancia sobre ciudadanos europeos, permitiéndoles requerir o incautar datos personales sin justificación. En un momento en el que multitud de empresas operan a ambos lados del atlántico, esta sentencia supuso un antes y un después, y llevó al levantamiento de nuevas barreras a las transferencias de datos entre la UE y EEUU, al dejar de considerarse este país un destino seguro para los datos de los ciudadanos europeos.

Marcos geopolíticos muy diferentes

¿Podría afirmarse que la sociedad está dispuesta a pagar con su privacidad unas mayores garantías de seguridad? ¿o que simplemente se resignan porque el uso de unos servicios que perciben como gratuitos compensa el inconveniente de la escucha activa? Sin duda esto es precipitado e inexacto, y la respuesta sería dispar según distintos interlocutores y contextos culturales. En este sentido pueden distinguirse tres orbes en función del marco legal imperante en materia de privacidad y de uso responsable de datos:

- El estadounidense: de facto el dominante según métricas de mercado, pues la combinación de tres factores fundamentales –universidades de primer orden, acceso a financiación de proyectos de innovación mediante fondos de capital riesgo con alta tolerancia a las pérdidas, y un marco legal inicialmente más permisivo en el uso de datos– ha dado lugar a los actores globales digitales que todos conocemos.
- El europeo: quizá a la zaga en desarrollos basados en datos y en generación de empresas digitales de escala global, pero en vanguardia a la hora de crear un marco legal garantista para la ciudadanía. Tanto es así, que la nueva legislación del estado de California –CCPA– está abiertamente inspirada en el Reglamento General de Protección de Datos de la UE. Además, la Comisión Europea se ha anticipado a sentar unas bases éticas y regulatorias que guíen los usos y aplicaciones de la Inteligencia Artificial.40

40 Proposal for a Regulation of the European Parliament and of The Council laying down harmonised rules on Artificial Intelligence.

- El chino: la administración de Xi Jinping ha trazado un plan estratégico cuyo objetivo final es erigirse como la primera potencia en materia de inteligencia artificial en el año 2030, para lo cual destinará una inversión de 150.000 millones de dólares. Los sectores abarcados son múltiples: salud, medio ambiente, mercado laboral, pero sin duda uno de los que se llevará gran parte de este presupuesto será el de la seguridad nacional, en sus dos vertientes: exterior (aplicaciones militares) e interior (sistemas de monitorización y control de la ciudadanía).

En definitiva, ante la pregunta *¿cuál es el lugar idóneo para percibir los beneficios que las herramientas basadas en inteligencia artificial nos va a traer?* La respuesta depende del interlocutor: en el caso de una empresa, probablemente EEUU; desde el punto de vista de los derechos individuales de los ciudadanos, en la UE; y del lado de un servicio estatal, seguramente China.

De pronosticar el futuro a manipularlo

Hasta ahora hemos hablado de la utilidad de la tecnología y de las nuevas fuentes de información para comprender el pasado y anticipar el futuro. Toca abordar además la capacidad que las herramientas tecnológicas han creado para influir y tratar de dar forma alternativa a ese futuro, una forma que vendrá predefinida por los incentivos que tengan quienes hagan uso de estos nuevos instrumentos.

En su saga *Fundación* (iniciada en 1942), *Isaac Asimov* acuñó el término *psicohistoria*: una técnica de control de la evolución social basada en una combinación de psicología, historia y estadística. En aquella serie literaria de ciencia ficción el control estaba orientado positivamente: buscaba acortar de 30.000 a 1.000 años de oscuridad el tiempo de recuperación tras un futuro colapso civilizatorio (sin duda la inspiración fue la obra de Edward Gibbon, *Historia de la decadencia y caída del Imperio Romano*).

En 2018 la filtración de un vídeo de "design fiction" titulado *The Selfish Ledger,* generado por y para comunicación interna entre el equipo Google, al trascender hizo saltar bastantes alarmas, por hacer explícitas las nuevas capacidades de un actor global a la hora de influir en la evolución de nuestra sociedad, incluso cuando los ejemplos de aplicación que menciona son en principio loables: lucha contra el cambio climático, cuidado de la salud, etc.

Cabe distinguir aquí dos elementos distintos:

- La polarización no intencionada: la selección de contenidos responde en primera instancia al interés de los individuos por ver reforzadas sus creencias previas cuando acceden a la información (sesgo de afirmación), pero a este viejo efecto, por el cual tradicionalmente los lectores adquirían únicamente los periódicos de su misma tendencia política, se suman ahora los incentivos empresariales por retener a los usuarios de servicios informativos digitales y redes sociales mediante la creación algorítmica de filtros que recrean burbujas informativas. La radicalización de contenidos hace que los usuarios pasen más tiempo haciendo uso de estos servicios, como han analizado los investigadores Zeynep Tufekci[41], o Jan Lorenz y Peter Holtz, entre muchos otros, y este tiempo de atención cautiva es una de las métricas que guía los ingresos por publicidad de estos servicios.
- La polarización intencionada: responde a la voluntad de determinados agentes por ejercer el control social, un fenómeno nada nuevo de hecho, con profundas raíces históricas y todo un corpus académico que lo documenta. Lo novedoso son las herramientas que lo posibilitan: en el empleo de datos masivos y en la focalización de la audiencia radican las mayores novedades respecto a la propaganda tradicional. La tecnología actual permite primero perfilar mediante técnicas de psicografía a los usuarios de redes sociales para identificar los perfiles indecisos más propensos a ser influidos, y puede a continuación centrarse en aquellos capaces de inclinar la balanza en los lugares clave donde un cambio de tendencia local puede tener mayor influencia global, como se vio con el escándalo de *Cambridge Analytica*, (aunque hay que ser también cautos con la influencia real[42] que estas técnicas de segmentación pudieron tener en el referéndum del *Brexit* en junio de 2016, o en las elecciones estadounidenses de noviembre de ese mismo año, pues su influencia ha podido ser magnificada por los medios de comunicación, ante la sorpresa inicial de aquellos resultados electorales). Lo que sí han demostrado investigadores como Sinan Aral es que las mentiras y noticias falsas, por su efecto novedoso y su componente de supuesto escándalo, tienden a propagarse por las redes sociales a una velocidad superior al de las verdades; sencillamente resultan más atractivas y activan las emociones primarias de la audiencia de una forma más directa. Además,

41 Expuesto en su charla TED: "We're building an artificial intelligence-powered dystopia just to make people click on ads".

42 Artículo en la revista *Nature: The scant science behind Cambridge Analytica's controversial marketing techniques.*

la automatización en la publicación de estas noticias falsas mediante bots permite escalar su dispersión a un coste muy bajo.

Desde 1945 hemos convivido con la amenaza de la bomba atómica como un arma de disuasión a la que muchos países aspiran, unas pocas potencias han accedido, pero que en realidad nadie quiere utilizar. El poder de manipulación de la opinión pública a través de las nuevas herramientas y canales digitales es un riesgo diferente, porque sus efectos son apenas visibles mientras se gestan pero, lamentablemente, son en cierto modo eficaces a escala masiva, y están al alcance de muchos a bajo coste.

En definitiva, la humanidad afronta cuatro grandes retos en este nuevo escenario:

1) Mantener la libertad individual y el control sobre el acceso a nuestra huella digital.
2) Mantener bajo control democrático las capacidades de vigilancia masiva, a la vez que se garantiza la seguridad no opresora.
3) Utilizar la inercia globalizadora y seguir universalizando el acceso a la información rigurosa para procurar que el sistema de democracias liberales se haga extensivo a países en la actualidad autocráticos.
4) Limitar el poder de los monopolios para mantener una diversidad suficiente en las posibilidades de acceso a la información, manteniendo viva la llama del pensamiento crítico pero sosegado, para orientarnos en la era de la posverdad y los radicalismos.

Bibliografía

Nassim Nicholas Taleb (2007) *The Black Swan: The Impact of the Highly Improbable.* Random House and Penguin Books. New York, USA.

Judea Pearl and Dana Mackenzie (2018) *The Book of Why: The New Science of Cause and Effect* Basic Books. New York, USA.

George Lakoff (2004) Don't Think of an Elephant!: Know Your Values and Frame the Debate. White River Junction. Chelsea Green Pub. Columbia, USA

Daniel Geschke, Jan Lorenz, Peter Holtz (2018) *The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers*. British Journal of Social Psychology. London, UK.

Mark R. Grandstaff (2006) *Propaganda and Mass Persuasion: A Historical Encyclopedia, 1500 to the Present.* The Journal of Military History, Virginia, USA

Juan Murillo es Ingeniero de Caminos, Canales y Puertos especializado en Urbanismo, y cuenta con un Executive MBA por la Escuela de Organización Industrial. Ha desarrollado su carrera en el ámbito del planeamiento urbano, y en el análisis y diagnóstico socioeconómico de municipios en fase de transformación: destinos turísticos y áreas de reconversión objeto de proyectos de mejora y dinamización para impulsar su sostenibilidad económica y medioambiental. En el ámbito de la investigación ha colaborado en diversos artículos científicos en cooperación con el MIT Senseable City Lab y otras instituciones académicas. Ha sido miembro del equipo de Data Strategy & Data Science Innovation de BBVA. Entre sus intereses personales se cuentan la historia y la prospectiva, la ciencia y la tecnología, o la regulación digital. En la actualidad es asesor en estrategia urbana y transformación digital, y dirige el área de Sostenibilidad y Ciudades Inteligentes de OdiseIA.

Inteligencia Artificial, artificial

Eduardo Lazcano

La inteligencia como cualidad humana se fundamenta en la capacidad de "unir los puntos", de comprender las cosas mediante la asociación de aspectos que no parecen ser asociables a simple vista. La inteligencia espacial utiliza la interpretación de las sombras y el momento del día para determinar la ubicación cardinal y también es capaz de reconocer un objeto desde diferentes perspectivas relacionando elementos del mismo desde diferentes ángulos. De este modo, el proceso siempre consiste en observar, en relacionar y en proyectar conclusiones. Adicionalmente, se pueden generar predicciones que no son más que escenarios plausibles creados mediante nuevas combinaciones de viejos elementos.

Los modelos de análisis están siendo sometidos a múltiples disrupciones

Refiriéndonos a la inteligencia que nos ocupa, el proceso es el mismo: observar, analizar y proyectar. El problema es que cada elemento de esta secuencia está siendo sometida a una serie de disrupciones que han terminado por cuestionar su fiabilidad.

En primer lugar, el número de elementos a analizar o puntos a unir ha crecido exponencialmente debido a la revolución en la capacidad de generar datos no solo a través de la sensorización de objetos inanimados, sino también mediante la interpretación de los comportamientos humanos. Uno de los grandes retos es definir qué se mide. Las métricas tradicionales ya no sirven para los propósitos actuales y por ello, hace tiempo que el mundo de la empresa dejó de realizar segmentaciones sociodemográficas para centrarse en otras actitudinales e incluso persono-técnicas, considerando que no somos la misma persona todas las horas del día, sino que mutamos según las circunstancias.

Lo importante no es cómo se mide, sino qué se mide. Pat Riley fue el entrenador del equipo de baloncesto Los Angeles Lakers entre 1981 y 1990 ganando cuatro títulos de la NBA. Riley no solo quería un equipo técnicamente cualificado; lo que quería por encima de todo era un equipo intenso y para ello definió las métricas que iba a manejar para evaluar el rendimiento de sus jugadores. En lugar de dar prioridad a los porcentajes de tiro, se la daba a los rebotes, que son más consecuencia de

la intensidad. Pero también cuantificaba las veces que los jugadores chocaban las manos, se tiraban al suelo a por un balón o incluso las que protestaban al árbitro. Porque todas ellas miden mejor la intensidad. Cuando tienes claro lo que pretendes, es mucho más fácil encontrar los indicadores adecuados y esos indicadores son los puntos que luego deberemos unir.

El segundo aspecto de la cadena de valor que ha sufrido una disrupción es la del análisis. Si la cantidad de información ha crecido exponencialmente –en Telefónica predicen que en 2020 habrá 40 zettabytes[43]–, no lo ha hecho al mismo ritmo la capacidad de análisis. Corremos por tanto un claro riesgo de caer en la infoxicación, definida como el bloqueo por exceso de información. Cuando Al Capone fue requerido por la agencia tributaria norteamericana, les facilitó toda la información. Y cuando digo toda, es absolutamente toda. De este modo saturó a los investigadores ralentizando la instrucción del caso y bloqueando los recursos de la institución. El cerebro dedica el 20% de su tiempo y energía a recopilar información y el 80% a analizarla y procesarla. Es la distribución inversa a la que tenemos en las instituciones y en las empresas en las que empleamos la mayoría de los recursos en obtener información, mucha de ella desechable en última instancia. El futuro de la inteligencia pasa, pues, por nuestra capacidad de dimensionar el esfuerzo entre ambas tareas y de ser capaces de realizar una compilación selectiva de información y un análisis suficiente.

En el aspecto de la proyección de conclusiones, el cambio es radical. Si decíamos que la inteligencia espacial nos permite rotar una forma en la mente para proyectar qué aspecto tendría una vez girada, lo que está sucediendo hoy en día es que las formas se están transformando durante el giro haciendo impredecible su aspecto al final. Como paradigma de ello, vemos que las investigaciones electorales están fallando como nunca, cuestionando la fiabilidad de los estudios declarativos. Sucede porque tomamos las decisiones inconscientemente o, mejor dicho, desde la parte inconsciente de nuestro cerebro. El inconsciente procesa el pensamiento de la misma manera que el consciente, con los mismos prejuicios, experiencias y sesgos, aunque no seamos capaces de percibirlo conscientemente.

John Bargh[44] hace la analogía siguiente: la parte consciente de nuestro cerebro procesa a 55 *bits* por segundo (bps) y nuestro inconsciente lo hace a 11 millones de

43 Un zettabyte es una unidad de almacenamiento de información cuyo símbolo es el ZB, equivale a 10^{21} bytes.

44 John Bargh. (2007) *Sociology Psychology and the Unconcious: The Automaticity of Higher Mental Processes (Frontiers of Social Psychology), Psychology Press, United Kingdom*

bps. Esto significa que nos manejamos en el día a día con un procesador de juguete mientras nuestro comportamiento menos predecible lo hace con un superordenador. Cuando conocemos a alguien, aunque no lo sepamos, estamos procesando su olor. Lo que sucede es que, como es una información que consideramos poco relevante, la dejamos en nuestro inconsciente para liberar capacidad de nuestro limitado procesador consciente. Que no lo percibamos conscientemente, no quiere decir que no suceda. Robert Kurzban[45] explica cómo el inconsciente engaña sistemáticamente al consciente por motivos de supervivencia, por ejemplo, para hacernos creer que somos mejores profesionales de lo que somos. Dice Kurzban que la parte consciente es como la secretaria de prensa de la Casa Blanca que no sabe si se han torturado o no a presos en Guantánamo porque cuando comparezca ante la prensa, debe ser capaz de decir que "no le consta" con credibilidad.

Las encuestas e investigaciones basadas en metodología declarativa están siendo afectadas por esa disonancia entre ambos aspectos de nuestro cerebro. Cuando nos preguntan qué vamos a votar, muchas veces no nos sentimos cómodos diciendo la verdad y, aunque inconscientemente tenemos claro lo que vamos a hacer, esperamos hasta el final para justificarnos con una decisión aparentemente impulsiva. Lo mismo sucede cuando tenemos que pronunciarnos sobre temas sensibles como la inmigración o la igualdad.

Los modelos proyectivos son menos fiables que nunca, especialmente si se basan en estudios declarativos, pero la distorsión es más grande que eso. En una entrevista al científico de datos D.J. Patil[46] fue preguntado por lo heterogéneo de su carrera y él contestó negando la mayor. Siempre se había dedicado a lo mismo, a generar modelos predictivos en contextos de incertidumbre. Así, en meteorología se puede predecir la temperatura que va a hacer dentro de 15 días un grado arriba o un grado abajo o se puede ser incapaz de predecir qué tiempo va a hacer dentro de dos horas. Pues según Patil, el mundo ha entrado en un periodo de 20 años en el que no se sabe qué tiempo va a hacer dentro de dos horas. De hecho, la impredecibilidad es un factor determinante que no podemos controlar y sólo podemos adaptarnos a su existencia.

45 Robert Kurzban (2011) *Why everyone (else) is a hypocrite, Princeton University Press, Estados Unidos*

46 D.J. Patil se formó como matemático en la Universidad de San Diego, California y se doctoró en la Universidad de Maryland College Park en matemáticas aplicadas, desarrollando modelos predictivos meteorológicos para la NOAA (Administración Nacional Oceanográfica y Atmosférica). Después lideró un programa de detección de bioamenazas para la Oficina de Sistemas Avanzados del gobierno de los Estados Unidos. Posteriormente fue Chief Scientist en Linkedin y en la era Obama en la Casa Blanca, antes de volver al terreno académico.

En definitiva, la cantidad de información que hay que procesar y la dificultad de establecer patrones proyectivos hacen que la aspiración de predecir requiera de una reformulación de nuestras habilidades computacionales a la hora de interpretar el masivo flujo de información. Sin duda, ahí está una de las vías de desarrollo de la innovación en términos de inteligencia.

Inteligencia proactiva vs Inteligencia reactiva

Sin embargo, puede que haya otras rutas para la innovación en inteligencia menos secuenciales. Es quizás el momento de trabajar a dos velocidades diferentes, pero al mismo tiempo. La primera velocidad es la de analizar lo que tenemos, "lo que hay". Seleccionar los datos adecuados y procesarlos de forma inteligente para extraer conclusiones. Es lo que podríamos llamar "la vía computacional". Podríamos compararla con la innovación incremental. La segunda vía se basaría en trabajar sobre "lo que no hay". La innovación disruptiva es lo que aún marca la diferencia entre el hombre y la máquina. La máquina ya es capaz de crear, pero sobre lo que ya hay y esa innovación incremental ya se está desarrollando de forma vertiginosa.

Sería interesante reflexionar sobre si estamos demasiado focalizados en "detectar riesgos". Partimos de un entorno impredecible en el que las cosas están en permanente cambio y en la que, además 'los malos' tienen menos restricciones a la hora de operar.

La diferencia entre películas como *Furia de Titanes* en la que los protagonistas luchan con un cangrejo gigante y películas como *Avatar* es que en la primera el guionista puso un cangrejo luchando con personas porque era lo máximo que su imaginación podía concebir que pudiese realizarse en una película. El guionista de *Avatar* partió de un folio en blanco en el que todo era posible. La interpretación de riesgos es proyectiva y por tanto se apoya en lo que ya ha ocurrido, lo que conocemos. Sin embargo, los enemigos de la sociedad viven en un paradigma sin límites para la imaginación.

Es por ello que valdría la pena pensar qué es más eficiente, si basarnos en el acierto de nuestras predicciones o pensar en proyectar escenarios proactivos que nos permitan tomar la iniciativa. Es decir: acertar los riesgos o plantear oportunidades.

Cuando un defensa se enfrenta a un jugador como Messi, sabe que no hay forma de controlarlo por su variedad de recursos. Entonces, lo que hace es tratar de anticiparse estableciendo un sistema de coberturas que se anticipan a su vez a su

eventual fallo. A nivel empresarial sucede exactamente igual. Las industrias que plantean estrategias reactivas están volcadas en frenar las sangrías ocasionadas por las innovaciones de los competidores, siempre a rebufo. Sólo las que se anticipan con una posición, se apoderan del liderazgo y asumen la responsabilidad de dar el golpe primero, pueden tomar el control –en ocasiones– de la situación.

Daniel Solana[47] explica que, ante un problema complejo, si no sabes por dónde empezar, lo que tienes que hacer es empezar. Es decir, hay problemas y cuestiones que no se dibujan fácilmente llevándonos al bloqueo y en esa situación, lo mejor es hacer un movimiento porque el propio movimiento generará una reacción que ayudará a dibujar mejor el problema.

Cuando vamos a comprar una casa para vivir, estudiamos los mapas, las zonas que nos gustan, pero hasta que no pisamos la calle y vivimos la zona in situ, no sabemos realmente si nos va a gustar. Del mismo modo, cuando dos boxeadores comienzan el combate, la única forma de averiguar la estrategia del oponente es soltar el primer golpe. Ahí se dibujará el planteamiento del rival.

Este escenario de impredecibilidad no durará toda la vida, pero lo cierto es que en la realidad sociopolítica estamos viendo que los planteamientos proactivos, por muy extremos o incluso disparatados que parezcan, están teniendo mejores efectos para sus objetivos que los planteamientos más reactivos. El miedo a equivocarnos hace que nos sintamos más cómodos en la posición de no hacer nada hasta que no quede más remedio. De este modo, con el tiempo, se ha asimilado que la inacción no implica culpa. Si haces algo y te equivocas, es grave. Si no haces nada y ocurre algo malo, la culpa es del otro, del que hizo algo.

Si decidimos desarrollar una línea de innovación proactiva, explorando lo que no hay y planteando escenarios desde la creatividad y la empatía, estamos entrando en el terreno de las habilidades humanas y por eso surge el planteamiento de la Inteligencia Artificial artificial. Es decir, se trata de desarrollar la tecnología de inteligencia artificial hasta sus últimas capacidades y, una vez ahí, hackearla con un pensamiento puramente humano –de ahí el segundo "artificial"–. La consultora de innovación Soulsight realiza dinámicas bajo la teoría del décimo hombre que consiste en que, cuando en un grupo de 10 personas, todos están de acuerdo en que se va a dar una tendencia, uno de ellos debe defender lo contrario con el fin de explorar escenarios disruptivos.

47 Daniel Solana (2015) *'Desorden', LID Editorial, España*

El instinto de supervivencia tiende a preservar el libre albedrío y una de las formas de hacerlo es rebelarse contra aquellos desarrollos tecnológicos o sociales que hacen parecer el comportamiento del ser humano como si fuese predecible. En cuanto se establece algo como predecible, de alguna manera, tendemos a hackearlo y demostrar nuestra capacidad de tomar nuestras propias decisiones. Esto se manifiesta de forma especialmente clara en el contexto publicitario y empresarial.

Al ser humano le gusta ser sorprendido

Muchas de las grandes campañas o desarrollos innovadores de producto fueron lanzados a pesar de que toda la investigación previa predecía que no iba a gustar a la audiencia, porque los estudios son capaces de extraer lo que la gente espera y le parece coherente, pero no tanto lo que le sorprendería. Esto se ve acentuado cuando se trata de modelos declarativos donde se pone de manifiesto la disonancia entre la percepción consciente e inconsciente de la realidad. Así, el mencionado fracaso continuo que vivimos hoy en día en las predicciones publicadas en las encuestas electorales.

La tendencia innovadora en el mundo empresarial en los últimos años se está focalizando en la interpretación de la realidad desde el punto de vista humano. Grant McCracken[48] es un antropólogo y cuenta que en una ocasión fue invitado a participar en una sección del *Show de Oprah*, en la que él mismo y una decoradora de interiores visitarían la casa de los Sullivan, el típico hogar americano, para describir lo que allí se encontrasen. Imaginemos la situación:

Llaman a la puerta, conectan en directo y entran. La decoradora comienza a juzgar lo que allí ve: «Bueno, es obvio que esta familia no tiene sentido del diseño. ¡Ninguno! Fíjate esas cortinas. Corte incorrecto, tamaño incorrecto, color incorrecto…».

La señora Sullivan retrocede arrepintiéndose de haber llamado a Oprah e imaginando a todas sus amigas pegadas al televisor.

Mientras McCracken se asoma a la habitación de al lado donde están Dan y Danielle, el padre y la hija, haciendo tiempo jugando a subirla a caballito e interpretando el baile de *Pocahontas*.

48 Grant McCracken (2009) *'Chief Culture Officer', Basic Books, Estados Unidos*

En ese momento, la diseñadora entra en la cocina espantada: «¡Oh! Mira estos muebles. Mira, de verdad. Parecen presionados contra la pared. No hay sentido de la distribución.» Entonces McCracken interrumpe: «¿Sabes por qué están presionados contra la pared? Para dejar espacio y que tengan sitio para hacer el baile de *Pocahontas*. ¿Quieres verlo Oprah? Pasen, pasen…», e invita al padre y la hija a que hagan la *performance*.

En ese momento, McCracken tomó consciencia de que el concepto de hogar para esa señora estaba en función de los muebles que allí había y su disposición, mientras que, para él, el concepto de hogar estaba en función del comportamiento de sus habitantes.

La clave de la innovación disruptiva en términos de inteligencia podría estar en saber interpretar la realidad desde el comportamiento de las personas, desde sus motivaciones y sus pasiones. No es lo que se ve, sino lo que se interpreta.

Una vez estuve dando una formación en una multinacional del pequeño electrodoméstico. El grupo era sensacional y tocamos este punto como una de las claves de la transformación digital. Entonces señalé uno de sus productos que estaba en una estantería y pregunté: «¿Eso qué es?». «Una plancha para el pelo» contestaron todos. Otros decían el modelo, otros el nombre técnico. «¿Sabéis que veo yo? Veo la autoestima de una chica de 20 años que no duerme fuera de casa si no lleva su plancha para el pelo porque le da seguridad. En esa licuadora veo el compromiso vital de un padre de 40 por mantenerse joven para disfrutar con salud de su hijo recién nacido. En esa plancha veo la pereza de una mujer que trabaja de sol a sol y, cuando acaba, tiene que ponerse a hacer las labores del hogar mientras su pareja no ayuda».

Así el iPhone en su día no parecía una gran idea bajo el argumento de que nada haría mejor una foto que una cámara de fotos. Las empresas interpretábamos que las partes desarrolladas individualmente siempre darían mejor calidad que lo que se pudiese hacer en un conjunto. Lo que ocurre es que había un factor mucho más importante para la gente que el de la calidad del resultado y era la portabilidad, la posibilidad de tener muchas cosas en un solo dispositivo. El éxito de las navajas suizas no era que el abrelatas fuese el más efectivo ni que el destornillador fuese el más cómodo. La clave estaba en la posibilidad de tenerlo todo en el mismo sitio y no tener que preocuparte más que de una cosa. Esa lectura de un comportamiento humano es lo que disparo la que hoy en día es una de las empresas más grandes del mundo.

Hay que interpretar la realidad desde la empatía

Mientras no seamos capaces de ver las cosas incorporando el foco de las personas, no seremos capaces de innovar en la dirección adecuada. Pero no es solo la interpretación empática de la realidad, sino la visión focal que tenemos a la hora de interpretarla. Hay gente que, ante un problema, hace *zoom in* y busca más detalles. Hay otra que, ante los problemas, hace *zoom out* y busca contexto. Hay gente que, cuando habla de cámaras de fotos, hace *zoom in* en desarrollar la mejor cámara de fotos. Hay gente que hace *zoom out* e interpreta para qué necesitamos la cámara de fotos llegando a la conclusión de que es más importante tenerla a mano en el momento en que queremos inmortalizar una vivencia que reproducirla en una calidad infinita.

Nos encontramos a muchos bancos invirtiendo cantidades ingentes de dinero en "reinventar la banca" cuando lo que se está reinventando en realidad es la relación que nosotros los ciudadanos queremos tener con el dinero. De ahí el auge de la economía colaborativa o que la gente más joven no sienta la necesidad de tener la casa o el coche en propiedad o el hecho de que no quieran trabajar en un banco, cosa que aún sorprende a los departamentos de recursos humanos de las grandes entidades financieras. Lo mismo sucede con el mundo de la prensa, que sigue trabajando para reinventar los medios de comunicación cuando, en realidad, lo que se ha reinventado es la relación que nosotros, la sociedad, queremos tener con la verdad.

La verdad no existe y no la queremos. Que Amancio Ortega done 3.200 millones de euros para comprar 2.900 máquinas que detectan el cáncer es un hecho, pero no una verdad. Que sea un acto de generosidad infinita o que sea una limosna humillante es la interpretación que hacemos del hecho y es a eso a lo que llamamos verdad. Dado que todos tenemos vocación de opinar sobre todo lo que vemos, la verdad como interpretación colectiva de la realidad, no existe.

La sociedad queremos entretenimiento, lo que han entendido muy bien la prensa del corazón y la deportiva. Si se le pregunta al entrenador del Real Madrid si teme al FC Barcelona y éste contesta un desinteresado "No", el titular al día siguiente será: "No temo al Barça". El proceso consiste en transformar un hecho irrelevante en algo que despierte el interés y las emociones de la audiencia.

La otra cosa que demandamos al margen de la verdad es la caja de resonancia. Queremos que nos digan lo que queremos escuchar y por eso es más viable un

proyecto que tenga una audiencia deseosa de escuchar lo que les quiere contar y una clientela deseando pagar por ello, que cualquier diario que busque la quimera de la verdad.

Es por tanto mi propuesta –sin perjuicio de una línea de innovación incremental más secuencial, más focal, reactiva y más proyectiva– el desarrollo de una línea más disruptiva, basada en la lectura de las motivaciones más instintivas y humanas para plantear de forma proactiva la introducción de estímulos que provoquen reacciones que dibujen mejor el problema.

De esta forma, la primera línea de trabajo dispondrá de más información para complementarse en un proceso sinérgico con la segunda.

Se llamará Inteligencia Artificial, Artificial.

Eduardo Lazcano es consultor independiente en proyectos de transformación, comunicación o desarrollo de equipos, además de su actividad docente y de investigación. Es autor de "Comunicación Emocional" (LID, 2017), La Gestión del Talento en la Era Digital (Foxize, 2014) y el primer Estudio de Consumer Centricity en España (IPMARK, 2016). Con anterioridad desarrolló su carrera profesional en Movistar, Pernod Ricard (compañía francesa propietaria y distribuidora de marcas de bebidas espirituosas como Absolut, Beefeater, Ballantine's o Chivas), y Deoleo (líder mundial propietaria y distribuidora de aceite de oliva con marcas como Carbonell, Hojiblanca, Koipe, Bertolli o Carapelli) en funciones de Brand Story Director.

Capítulo V: Casos de estudio

Seguridad inteligente con *deep learning* para detección de armas en imágenes de vídeo

Francisco Herrera, Siham Tabik, Alberto Castillo y Francisco Pérez

Resumen

El progreso exponencial de la tecnología está presente en la inteligencia artificial en campos como el aprendizaje automático o la visión por computador. Una técnica algorítmica específica de inteligencia artificial llamada redes neuronales convolucionales surge como solución para el procesamiento en problemas complejos de análisis de imágenes. Dentro de sus aplicaciones existen multitud de trabajos, entre otros la conducción autónoma y el reconocimiento facial. Debido a su potencial y resultados, han tenido un gran impacto también en el ámbito de la seguridad. Su aplicación ha supuesto el primer sistema en tiempo real para la detección de armas en videovigilancia. En este capítulo se presentan los estudios realizados para este problema y los resultados alcanzados.

Introducción: seguridad inteligente con videovigilancia e inteligencia artificial

Estadísticas publicadas por *Eurostat*[49] y por la Oficina de Naciones Unidas contra la Droga y el Delito revelan tasas de criminalidad preocupantes en la mayoría de los países europeos. De hecho, un informe de Eurostat de 2017 afirmaba que los sentimientos de inseguridad de los ciudadanos han aumentado en los últimos años debido a la delincuencia urbana y el terrorismo.

La seguridad y la protección de los ciudadanos contra la delincuencia urbana o el terrorismo se ven afectados por diversos factores. Por ejemplo, entre zonas de una ciudad con diferentes condiciones económicas y sociales, o situaciones urbanas que se consideran eventos de "alto riesgo", como los conciertos de música, los centros comerciales... Las multitudes, por ejemplo, se concentran en grandes eventos o en infraestructuras críticas, y representan un tipo diferente de desafío. Es

49 http://ec.europa.eu/eurostat/web/crime/database.

importante identificar y abordar comportamientos anómalos específicos o individuos sospechosos.

Hasta ahora, los principales sistemas de detección de armas estaban basados en detectores de metales [9] como los que se encuentran en aeropuertos y en eventos públicos en lugares cerrados. Estos sistemas pueden resultar poco efectivos para las particularidades del mundo actual. A pesar de esto, su robustez los hace necesarios en zonas de paso de determinados lugares (aeropuertos, edificios públicos, ...). Existe, pues, la necesidad de un sistema que refuerce y mejore los sistemas actuales para ser utilizados en una mayor variedad de entornos.

Una solución innovadora y efectiva sería dotar a las cámaras de vigilancia de inteligencia. Concretamente, las redes neuronales profundas o modelos de aprendizaje profundo [3], conocidas como *Deep Learning*, están mostrando tener un altísimo potencial en el reconocimiento de patrones espaciales en imágenes. En particular, las redes neuronales convolucionales (CNNs), una clase de redes neuronales profundas, constituyen actualmente el estado de arte en las principales tareas del campo de visión por computador como clasificación de imágenes, detección de objetos en imágenes y segmentación de objetos en imágenes. De hecho, desde 2012, todas las categorías de la prestigiosa competición *Large Scale Visual Recognition Challenge*[50] (ILSVRC) han sido ganadas por las CNNs [4].

La detección de armas y de objetos pequeños usando CNNs en el contexto de vídeo vigilancia es una tarea compleja debido a la alta variabilidad en cuanto a iluminación, materiales (las armas con una importante reflectancia superficial empeoran la calidad de las imágenes en ambientes de media y alta iluminación) y movimiento (los movimientos bruscos pueden producir emborronamiento en las imágenes). Además, varios objetos de uso diario como el smartphone se manejan de una forma parecida a algunas armas, lo que complica aún más la detección. Construir sistemas robustos y fiables frente a ese nivel elevado de variabilidad supone un desafío tecnológico. En este capítulo abordamos este problema y mostramos los avances y resultados recientes de un grupo de investigación de la Universidad de Granada para la detección de armas en vídeo.

A lo largo del mismo analizaremos qué son y como aprenden los modelos de aprendizaje profundo especializados en imágenes, es decir, las redes neuronales

50 Se trata de una competición mundial, organizada por el Vision Lab de Standford, que premia a los mejores algoritmos para la detección de objetos y clasificación de imágenes.

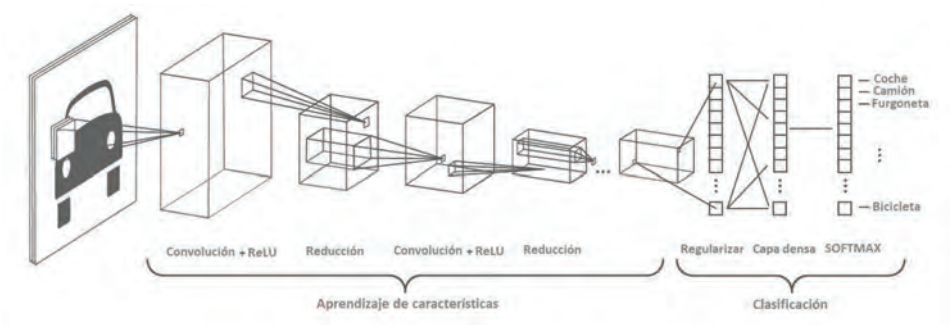
convolucionales. Se describen a continuación la detección de objetos en imágenes, los sistemas que se aplican para la detección de armas de fuego y armas blancas. Para finalizar, se plantean consideraciones futuras en seguridad inteligente.

Aprendizaje profundo y redes neuronales convolucionales ¿Qué son y cómo aprenden?

El aprendizaje profundo, área dentro de la Inteligencia Artificial, emula la estructura de neuronas del cerebro humano y su interconexión masiva. Realiza una agregación de múltiples operaciones simples sobre los datos de entrada y utiliza complejos algoritmos para el ajuste de esta estructura neuronal artificial. Así, consigue aprender una serie de patrones que de forma generalizada explican los datos que se pretenden analizar para generar nuevas inferencias [3].

Un subconjunto de modelos de aprendizaje profundo especializados en imágenes son las CNNs (figura 1), que han conseguido un gran avance en la clasificación y en la detección de objetos en imágenes. Estos modelos, basados en una combinación de operaciones matriciales simples para la extracción de patrones complejos en la imagen obtienen una serie de características principales y aprenden a partir de su entrenamiento en una colección amplia de ejemplos. Entrenando este tipo de modelos sobre un conjunto de datos de imágenes se obtienen sistemas automáticos de identificación de objetos pre-entrenados. Estos sistemas, gracias al conocimiento que han aprendido, son capaces de, dados nuevos ejemplos, discernir qué objeto está presente en éstos.

Figura 1. Visión conceptual de una red neuronal convolucional



Para el entrenamiento de estos modelos de clasificación es necesaria una amplia base de datos de imágenes. De esta forma, tenemos un problema de aprendizaje supervisado con el que se pueden realizar dos tipos de tareas importantes en el procesamiento de imágenes:

- Clasificación de imágenes: Proceso que discrimina la clase a la que pertenece la imagen, siendo cada clase un objeto conocido.
- Localización de objetos sobre una imagen: Proceso de selección de regiones sobre la que se basa la segmentación y la detección. La primera tiene como objetivo seleccionar los píxeles pertenecientes a múltiples objetos de la imagen y la segunda proporciona una ubicación y un tamaño de ventana que engloba al objeto en la imagen.

En la Universidad de Granada se ha creado el primer sistema de videovigilancia para la detección de armas en imágenes de vídeo haciendo uso de estas técnicas de aprendizaje profundo [5].

Detección de armas en vídeo utilizando CNNs

En esta sección se introduce qué es la detección de objetos con modelos de aprendizaje profundo, además se describen aplicaciones y técnicas enfocadas a la detección de armas. Se presentan los trabajos de detección de armas de fuego, fusión de imágenes para la visión binocular, detección de armas blancas, concluyendo con la importancia de las bases de datos y el pre-procesamiento como base del aprendizaje en detección.

¿Qué es la detección de objetos?

La detección de objetos en una imagen es la tarea encargada de localizar en la misma objetos proponiendo ventanas de tamaño variable en diferentes localizaciones dentro de la imagen. Posteriormente se determina si esas regiones se corresponden con el resultado deseado.

Un modelo de detección basado en aprendizaje profundo integra dos etapas (figura 2):

- Un algoritmo de proposición de regiones con el que, en la imagen de entrada se pretenden obtener las zonas de la misma que son más propensas a contener al objetivo mediante la proposición de ventanas de tamaño variable según patrones de conjunto de píxeles.

- Analizar con una red CNN si las zonas propuestas por este primer algoritmo de proposición de regiones son el objeto en cuestión. Estos modelos en su fase de entrenamiento aprenden a buscar el objeto en imágenes, reajustándose al cometer errores.

Figura 2. Etapas del modelo de detección de objetos, a la izquierda proposición de regiones y a la derecha la detección final de la red CNN

Estos modelos de detección tienen aplicaciones múltiples en diversos ámbitos relacionados con la seguridad como el diseño de un sistema de detección de armas para videovigilancia en tiempo real.

Detección de armas de fuego utilizando aprendizaje profundo

Según la revista MIT *Technology Review*, una de las aplicaciones de inteligencia artificial más novedosas en el mes de marzo del año 2017 fue la creación del sistema inteligente de detección de armas de fuego en tiempo real desarrollada por investigadores de la Universidad de Granada [5]. Este sistema inteligente que recibió el Premio Security Forum I+D+i 2017 se basa en un modelo de detección basado en CNNs.

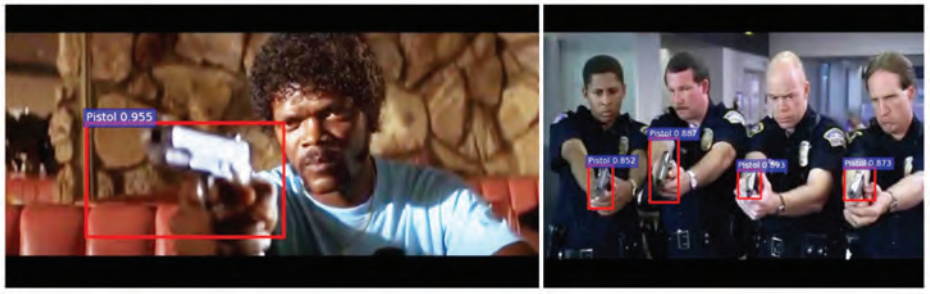
Este sistema de inteligencia artificial activa un aviso cuando detecta la presencia de un arma de fuego en una escena de vídeo. El algoritmo entrenado analiza cada imagen del vídeo en dos fases. Primero a partir de cada imagen genera una serie de regiones candidatas a contener un arma. Determinadas estas regiones se clasificarán dando una salida con la detección del objeto y su localización en la imagen. Para el entrenamiento de este modelo, ha sido muy importante la creación de una base de

datos de calidad, con la que el algoritmo basado en la CNN aprende las características del arma. La base de datos creada estaba compuesta de 3000 imágenes de pistolas con múltiples contextos y donde se usaban el objeto en distintas situaciones. Algunas de estas imágenes pueden verse en la Figura 3.

Figura 3. Imágenes usadas para entrenar el modelo de detección de armas de fuego

Tras el entrenamiento, el modelo adquiere la capacidad de distinguir las pistolas del resto de objetos empuñados por una persona. Posteriormente, el sistema procesa una secuencia de vídeo para localizar la presencia de pistolas en las imágenes activando una señal de alarma con la detección proporcionando un cuadro de color. La Figura 4 muestra un ejemplo de la salida que proporciona el sistema.

Figura 4. Escenas de las películas *PulpFiction* y *Mr. Bean* con la detección del sistema



Las aplicaciones de esta tecnología son múltiples:

- Se podría encontrar en un vídeo las escenas donde se visualicen pistolas sin necesidad de supervisar horas de grabación por una persona.
- Un sistema de cámaras de seguridad podría activar una alerta de la presencia de pistolas sin necesidad de intervención humana. Por ejemplo, un joyero que sufra un atraco con pistola en su joyería no tendría que arriesgar su vida intentando pulsar un botón que avise a la policía porque el sistema ya se encargaría de hacerlo.
- En un escenario de seguridad con múltiples cámaras, se podrían analizar todos los frames de los vídeos de forma que no se omitiese ninguna posible aparición de armas.

Como ejemplo de uso, se ha mostrado el funcionamiento del sistema inteligente de detección de pistolas en diferentes escenas de películas muy populares de los años 90, James Bond: *The World is Not Enough*, Pulp Fiction, Mission Impossible (*Rogue Nation*) and Mr. Bean. A pesar de la baja calidad de los vídeos usados, el detector proporciona una precisión alta y un número de falsos positivos (objetos clasificados como pistola erróneamente) bajo en todos ellos. Los vídeos están disponibles a través del enlace.[51]

Extensiones del modelo inicial de detección de armas

Una extensión de este sistema de detección ha sido publicada en [6] donde se aborda la mejora del sistema de detección de armas de fuego a través del pre-procesamiento de las imágenes basado en la información de profundidad obtenida por visión binocular.

Esta técnica simula la visión binocular humana para determinar la distancia entre la cámara y cada objeto del escenario. La técnica desarrollada combina la información obtenida por ambas cámaras en un mismo instante de tiempo para la obtención de la información de profundidad, lo que en términos técnicos se denomina mapa de disparidad. Esto requiere una configuración muy concreta del *hardware* en cuanto a posicionamiento de las cámaras, ángulo de direccionamiento, o algoritmo de fusión de información como principales requisitos.

El mapa de disparidad nos proporciona la información necesaria con la que podemos discernir zonas de no interés para la búsqueda del arma. Esas zonas de no interés

51 https://github.com/SihamTabik/Pistol-Detection-in-Videos.

serán descartadas de la imagen a procesar por el modelo de detección a través de operaciones binarias para la creación de una máscara que determina zonas que perderán sus características para no ser seleccionadas en la fase de proposición de regiones, evitando posibles falsos positivos por elementos del entorno (figura 5).

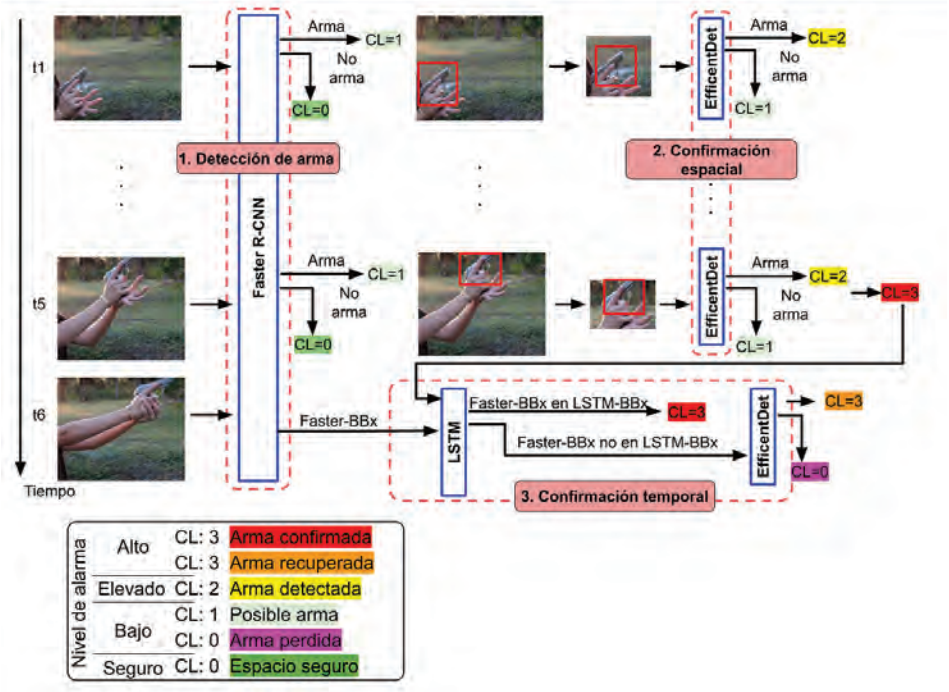
Figura 5. Imagen con la detección original (izquierda) y con el resultado del pre-procesamiento basado en fusión de imágenes por visión binocular (derecha)

Otra extensión ha sido publicada en [10] donde se aborda un sistema de multi-confirmación para reducir el número de falsos positivos a través del uso de la información temporal y espacial de las detecciones realizadas en frames consecutivos, metodología MULTICAST.

Esta técnica se basa en tres etapas (figura 6):

- Detección del arma: Esta etapa analiza el frame buscando una posible arma.
- Confirmación espacial con una CNN: Cuando la etapa anterior ha detectado una posible arma, se realiza una confirmación espacial confirmando la detección.
- Confirmación temporal con una LSTM: Se utiliza la posición de las detecciones previas para predecir la trayectoria del arma en futuros frames. Esta etapa recupera posibles falsos negativos producidos en la primera etapa.

Figura 6. Metodología MULTICAST basada en tres etapas, detección de arma, confirmación espacial y confirmación temporal



Detección de armas blancas

Un problema teóricamente similar a la detección de armas de fuego, pero diferente en la práctica debido a las diferentes características de cada tipo de arma, es la detección de armas blancas.

La implantación en sistemas de videovigilancia de los modelos de detección de objetos en imágenes a menudo encuentra la aparición de condiciones exógenas al problema debido a que la toma de las imágenes usadas para abordar un problema suele hacerse en condiciones no controladas.

En el diseño del sistema automático de detección de armas blancas en tiempo real, publicado en [2], se plantea el problema de la luminosidad sobre la detección de armas blancas debido al ruido en la imagen producido por la reflectancia de las armas. El sistema implementa la técnica DaCoLT (*Darkening and Contrast at*

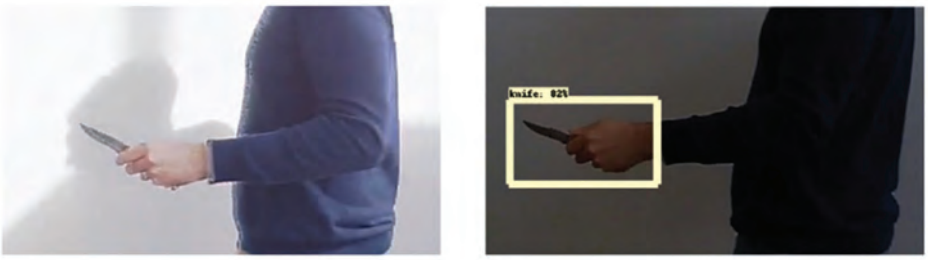
Learning and Test stages) para mejorar el rendimiento de detección en condiciones de luminosidad extrema.

El modelo CNN es entrenado con un conjunto de más de 2.000 imágenes que contienen una o más armas blancas en situaciones donde se empuñan de diferentes formas y el contexto donde se desarrolla es variado. Se pueden observar algunos ejemplos en la figura 7.

Figura 7. Imágenes de ejemplo usadas para entrenar el modelo de detección de armas blancas

Con el objetivo de localizar un arma, este sistema tiene la capacidad de procesar una secuencia de vídeo y lanzar una alarma más rápido que una persona centrada en la realización de esa misma tarea. Sobre un estudio realizado con 19 vídeos, el sistema de detección es capaz de lanzar la alarma al encontrar una secuencia de 5 imágenes consecutivas con arma en una media de 0,41 segundos. Los vídeos están disponibles a través del enlace https://github.com/alcasla/Automatic-Cold-Steel-Detection-Alarm.

Figura 8. Ejemplo de diferente resultado de detección según la condición de luminosidad



El estudio de luminosidad muestra la degradación y la variabilidad de la detección en distintas condiciones de luminosidad en escenarios interiores. El comportamiento variable según la condición de luminosidad (Figura 8) es un factor a controlar cuando se pretende abordar este problema en condiciones reales mediante modelos de detección. Para resolver la degradación en el rendimiento del sistema se propone un enfoque basado en pre-procesamiento (mejora cualitativa de los datos previo a su utilización por un algoritmo) que impulsa una mejora del aprendizaje generando nuevas imágenes artificiales. Aplicando una variación en su luminosidad, las imágenes mejoran en condiciones de luminosidad extrema mediante el ajuste de la luminosidad junto a la mejora del contraste para el procesamiento del vídeo.

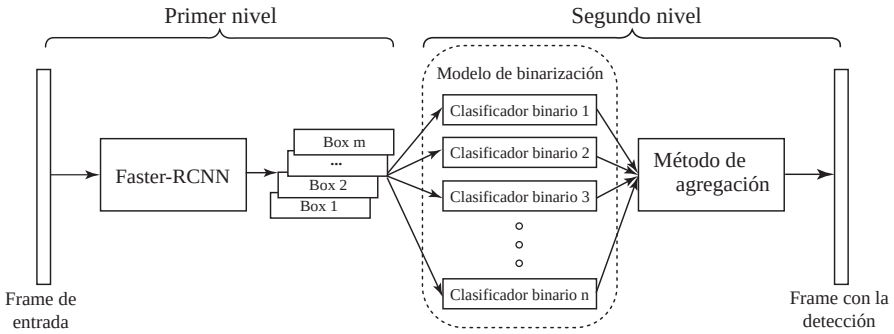
Detección de múltiples objetos

El sistema de detección de armas propuesto podría ser implementado en lugares como joyerías, bancos o como, por ejemplo, centros comerciales. Esto supone que las personas usas otros objetos como pueden ser los smartphones o incluso los monederos y podrían causar una detección errónea de un arma por la forma en la que se puede manejar. Es por ello por lo que se necesitaría tener un detector capaz de distinguir entre armas y objetos comúnmente manejados con las manos.

En el trabajo [11] se realiza la metodología ODeBiC basada en deep learning para mejorar la detección de objetos pequeños que se manejan de forma similar a un arma. Esta técnica se basa en dos niveles (figura 9):

1. En el primero nivel el detector propondrá regiones donde con un umbral se tendría sospecha de que pudiese aparecer un objeto de interés.
2. En el segundo nivel una técnica de binarización aplicaría todos sus modelos de clasificación para mediante una combinación de las salidas de todos los clasificadores dar una salida a cada una de las regiones propuestas.

Figura 9. Metodología ODeBiC para la detección de objetos pequeños que se manejan de forma similar a un arma



Otras propuestas en la literatura especializada

Los autores de [12] construyeron un conjunto de datos utilizando imágenes de un circuito cerrado de televisión (CCTV) instalado en la Universidad de Sevilla y datos sintéticos generados por el motor del juego Unity. Demostraron que la R-CNN más rápida basada en la red FPN (pirámides de características) y ResNet-50 puede utilizarse como detector de armas en CCTV casi en tiempo real.

Varios trabajos abordaron la reducción del número de FP y FN en la detección de armas en entornos de videovigilancia realistas utilizando enfoques de pre-procesamiento [13] y post-procesamiento [14]. En este último los resultados experimentales mostraron que el análisis de las armas de mano detectadas mediante un autocodificador mejora la precisión global de la detección en los vídeos.

Otros trabajos proponen utilizar la fusión de información para mejorar la detección de armas de mano en vídeos [15,16,17]. En el último, los autores utilizaron la estimación de la pose del esqueleto de las personas para detectar la amenaza en una imagen. Diseñaron un modelo de clasificación de varias etapas, en concreto, una primera CNN determina si una persona y una pistola están presentes en una imagen. En caso afirmativo, una segunda CNN estima la pose de la persona y, por último, una red neuronal pre-alimentada evalúa el nivel de amenaza basándose en las posiciones conjuntas de la estimación de la pose del esqueleto de las personas de la etapa anterior. El principal inconveniente de este enfoque es que no realiza una tarea de detección, sino que sólo clasifica fotogramas individuales. Del mismo modo, en [18], los autores utilizan primero la estimación de la pose de la persona para localizar las

zonas de las manos en la imagen y luego analizan la presencia de la pistola en esas regiones. La principal limitación de este enfoque es la fuerte dependencia entre la estimación de la pose y la detección de la pistola, si la estimación de la pose falla, la detección de la pistola también falla.

Bases de datos y pre-procesamiento: la base del aprendizaje para detección

Los modelos de aprendizaje profundo, aun empleando complejos y robustos modelos de detección, requieren también de un conjunto de datos bien diseñado y con imágenes de calidad. Este último requisito deriva de la forma de aprendizaje y funcionamiento de los modelos de detección de objetos, basados en la extracción de patrones presentes en conjuntos de píxeles en las imágenes. Por tanto, la calidad del conjunto de datos representada por aspectos como la nitidez, la precisión del color, o el brillo de la imagen, entre otros muchos aspectos, son requisitos fundamentales no siempre presentes.

El diseño de la base de datos de imágenes con el que se entrena al modelo CNN es una tarea compleja. Requiere de un concienzudo proceso de selección de imágenes de acuerdo a la resolución del problema abordado, teniendo en cuenta perspectiva, forma de sostener el objeto (en el caso de armas), oclusión del mismo, contexto o fondo, y elementos susceptibles de ser confundidos.

Un aspecto fundamental es la calidad de las imágenes utilizadas tanto para entrenar un modelo como las imágenes de vídeo procesadas para la búsqueda del arma objetivo. Es imprescindible el uso de técnicas de pre-procesamiento para la mejora de las imágenes, como el pre-procesamiento guiado por luminosidad o la eliminación del fondo basado en la visión binocular que hemos presentado, utilizando la salida de estas técnicas como entrada del modelo de detección.

Consideraciones futuras en la seguridad inteligente mediante vídeo vigilancia

Los sistemas de seguridad inteligente con inteligencia artificial integrada dejan ver un futuro en el que la automatización de los sistemas de videovigilancia proporcione herramientas complementarias a las empleadas en la actualidad. El rendimiento mostrado por los buenos resultados conseguidos deja ver el potencial y proporciona confianza para poder dar soporte en lugares conflictivos. Otro aspecto a favor de estos sistemas es el bajo coste que supone su implantación, ya que sólo necesita de un sistema de cámaras conectado a un servidor que utiliza unidades de procesamiento GPUs.

El potencial de las técnicas de inteligencia artificial deja en el horizonte una serie de problemas abiertos en el ámbito de la seguridad y la videovigilancia, incluyendo:

Videovigilancia con drones. El uso de drones ofrece una nueva perspectiva de zonas abiertas con la que se podría tener un rango de visión mayor para una prevención más temprana.

Comportamiento de multitudes. El análisis del comportamiento humano ha sido tradicionalmente abordado en entornos no-multitudinarios. El comportamiento de una multitud es más que la mera suma de comportamientos individuales, ya que presenta características colectivas que pueden describirse en términos generales (por ejemplo, "una multitud pacífica", "una multitud agitada" o "una aglomeración") [8]. Las técnicas actuales no son apropiadas para analizar multitudes y esta problemática apenas ha comenzado a abordarse en la literatura [7]. Se pueden utilizar los algoritmos de aprendizaje profundo para la detección de comportamiento anómalos en el contexto de la videovigilancia. Entre otros objetivos, se pueden utilizar para realizar el tracking de personas o grupos e identificar comportamientos anómalos en entornos de vídeo vigilancia multitudinarios (aglomeraciones, caídas, estampidas, masas agitadas o estresadas, saltos de vallas, etc.).

Seguridad en fronteras. El paso de vehículos en fronteras supone un riesgo cuando la matrícula no se corresponde con el modelo, falsificación que habilita el paso a vehículos. Una solución puede pasar por la identificación automática del modelo del vehículo a través de un sistema inteligente, para así cruzar la información de matrícula con el modelo de vehículo utilizando la base de datos policial.

Comentarios finales

El gran avance tecnológico de la Inteligencia Artificial que está teniendo lugar hoy en día puede y debe tener grandes repercusiones en el ámbito de la seguridad.

Son muchas las aplicaciones que hacen uso de modelos de aprendizaje profundo y, más en concreto, de CNNs. Una de estas aplicaciones es la creada por la Universidad de Granada con la que se ha diseñado un sistema de seguridad para detección automática de armas en imágenes de videovigilancia. Con este sistema se puede lanzar una señal de alarma en tiempo real cuando aparece un arma, ya sea una pistola o un arma blanca, en una de las imágenes de los vídeos de videovigilancia.

De cara a desarrollos futuros, cabe destacar los avances tecnológicos en redes de comunicaciones, especialmente el paradigma de Mobile Edge Computing (MEC)

[1]. La tecnología MEC permite optimizar las tareas de procesamiento de los datos generados por los sensores conectados en la red (e.g., cámaras de vigilancia, smartphone) ejecutándolas en los dispositivos físicos más cercanos donde se crearon en lugar de enviar los datos a través de largos recorridos de la red para que lleguen a centros de datos y nubes de computación. Los objetivos de este paradigma son reducir la congestión de la red de comunicaciones y permitir el procesamiento en tiempo real de los datos en los sensores. En este sentido, parece interesante integrar los modelos de aprendizaje profundo para detección de armas y objetos bajo el paradigma MEC.

Bibliografía

N. Abbas, Y. Zhang, A. Taherkordi, T. Skeie. (2018) Mobile Edge Computing: A Survey. IEEE Internet of Things Journal, 5, 450-465.

A. Castillo, S. Tabik, F. Pérez, R. Olmos, F. Herrera. (2019) Brightness guided preprocessing for automatic cold steel weapon detection in surveillance videos with deep learning. Neurocomputing, 330, 151-161.

I. Goodfellow, Y. Bengio, A. Courville. (2016) Deep Learning. MIT Press, Cambridge, United States.

J. Hu, L. Shen, G. Sun. (2017) Squeeze-and-excitation networks. arXiv preprint arXiv:1709.01507, 7.52

R. Olmos, S. Tabik, F. Herrera. (2018) Automatic handgun detection alarm in videos using deep learning. Neurocomputing, 275, 66-72.

R. Olmos, S. Tabik, A. Castillo, F. Pérez-Hernández, F. Herrera. (2019) A Binocular Image Fusion Approach for Minimizing False Positives in Handgun Detection with Deep Learning. Information Fusion, 49, 271-280.

V. Murino, M. Cristani, S. Sah, S. Savarese. (2017) Group and Crowd Behavior for Computer Vision, Academic Press, Cambridge, United States.

O.J. Urizar, E.I. Barakova, L. Marcenaro, C.S. Regazzoni, M. Rauterberg. (2017) Emotion estimation in crowds: a survey. 8th International Conference of Pattern Recognition Systems (ICPRS 2017), 1-6.

I. Uroukov, R. Speller. (2015) A preliminary approach to intelligent x-ray imaging for baggage inspection at airports. Signal Processing Research, 4, 1–11.

52 https://arxiv.org/abs/1709.01507.

R. Olmos, S. Tabik, F. Pérez-Hernández, A. Castillo, F. Herrera. (2021) MULTICAST: MULTI Confirmation-level Alarm SysTem based on CNN and LSTM to mitigate false alarms for handgun detection in video-surveillance. arXiv:2104.11653.[53]

F. Pérez-Hernández, S. Tabik, A. Castillo, R. Olmos, H. Fujita, F. Herrera. (2020) Object detection binary classifiers methodology based on deep learning to identify small objects handled similarly: Application in video surveillance. Knowledge-Based Systems, 194, 105590.

González, J. L. S., Zaccaro, C., Álvarez-García, J. A., Morillo, L. M. S., & Caparrini, F. S. (2020). Real-time gun detection in CCTV: An open problem. Neural networks, 132, 297-308.

Mahajan, R., & Padha, D. (2018, December). Detection of concealed weapons using image processing techniques: A review. In 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC) (pp. 375-378). IEEE.

Vallez, N., Velasco-Mata, A., & Deniz, O. (2021). Deep autoencoder for false positive reduction in handgun detection. Neural Computing and Applications, 33(11), 5885-5895.

Truong, T., & Yanushkevich, S. (2020, December). Detecting subject-weapon visual relationships. In 2020 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 2047-2052). IEEE.

Basit, A., Munir, M. A., Ali, M., Werghi, N., & Mahmood, A. (2020, October). Localizing firearm carriers by identifying human-object pairs. In 2020 IEEE International Conference on Image Processing (ICIP) (pp. 2031-2035). IEEE.

Abruzzo, B., Carey, K., Lowrance, C., Sturzinger, E., Arnold, R., & Korpela, C. (2019, November). Cascaded neural networks for identification and posture-based threat assessment of armed people. In 2019 IEEE International Symposium on Technologies for Homeland Security (HST) (pp. 1-7). IEEE.

Ruiz-Santaquiteria, J., Velasco-Mata, A., Vallez, N., Bueno, G., Álvarez-García, J. A., & Deniz, O. (2021). Handgun detection using combined human pose and weapon appearance. IEEE Access, 9, 123815-123826.

Francisco Herrera (EurAI Fellow 2009, IFSA Fellow 2013) es catedrático de Ciencias de la Computación e Inteligencia Artificial de la Universidad de Granada y director del Instituto DaSCI (Instituto Andaluz de Investigación en Ciencias de la Información e Inteligencia Computacional). Es Académico de la Real Academia de Ingeniería. Ha dirigido 51 tesis doctorales y publicado más de 550 artículos en revistas, recibiendo más de 104.000 citas (Scholar Google, H-index 156). Es

53

coautor de varios libros en diferentes campos de la inteligencia artificial, y es editor principal de la revista internacional "Information Fusion" (Elsevier). Es miembro de la redacción de una docena de revistas. Ha sido seleccionado como Investigador Altamente Citado en los campos de Ciencias de la Computación e Ingeniería, respectivamente, desde 2014 hasta la actualidad (Clarivate Analytics).

Siham Tabik, investigadora Ramón y Cajal en el Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada. Es doctora en Informática por la Universidad de Almería. Su investigación se centra en los modelos de aprendizaje automático con deep learning para problemas reales de vídeo seguridad.

Francisco Pérez Hernández, estudiante de doctorado en el departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada. Obtuvo el título del Grado en Ingeniería Informática y del Máster en Ciencia de Datos e Ingeniería de Computadores en el 2016 y 2017 respectivamente en la Universidad de Granada. Su línea de investigación se centra en la realización de estrategias de pre y post-procesado en deep learning para problemas multiclase en el ámbito de la seguridad.

Alberto Castillo Lamas, estudiante de doctorado en el departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada. Obtuvo el título de grado en 2015, realizó el Máster Profesional en Ingeniería Informática en 2016, y se especializó con el Máster en Ciencia de Datos e Ingeniería de Computadores en 2017 en la Universidad de Granada. Su línea de investigación se centra en el desarrollo de modelos deep learning junto a pre-procesamiento y post-procesamiento de imágenes, para la detección objetos y clasificación multi-instancia de escenarios complejos.

El proyecto Good Judgement: lecciones de un estudio de pronóstico

Eva Chen

Resumen

Entre el 2011 y el 2015, la Actividad de Proyectos de Investigación Avanzados de Inteligencia (IARPA, según sus siglas en inglés) de los Estados Unidos financió un proyecto de muchos millones de dólares para investigar los mejores métodos de elicitar, agregar y comunicar pronósticos probabilísticos sobre sucesos geopolíticos. El proyecto de Estimación Agregativa de Contingencias (ACE) puso a cinco organizaciones académicas a hacer pronósticos diarios sobre temas tales como "¿probará Corea del Norte un misil de largo alcance antes del 1 de junio del 2015?". El proyecto *Good Judgement*, basado en la Universidad de California Berkeley y la Universidad de Pensilvania, utilizó la sabiduría de ciudadanos de a pie para hacer pronósticos exactos que no sólo superaron a los del control y los de las otras organizaciones, sino que también a una línea de base interna elaborada por analistas de inteligencia que tenían acceso a información clasificada. Este artículo explica la labor que llevó a cabo *Good Judgement* para crear y combinar sus tres ingredientes clave: la formación, la selección y la agregación.

Introducción

Durante mucho tiempo, la investigación sobre el juicio humano se ha centrado en los sesgos cognitivos y en la incomprensión de las probabilidades (Kahneman y Tversky, 1973, 1984; Slovic y Fischhoff, 1977; Tversky y Kahneman, 1974). Las personas cometen constantemente errores sistemáticos en sus estimaciones de probabilidades por culpa de apoyarse demasiado en la heurística intuitiva (Kahneman y Tversky, 1977, 1982; Morewedge y Kahneman, 2010; Tversky y Kahneman, 1974). La formulación de juicios de probabilidad precisos es fundamental en ámbitos tales como el derecho, las finanzas, la medicina y la política (Croskerry *et al*., 2003; Jolls y Sunstein, 2004). Lo que es más importante, los gobiernos tienen que ser capaces de realizar pronósticos exactos que podrían afectar la vida de millones de ciudadanos, tales como cuánto invertir en una medicina antiviral o cuántas armas desplegar en una zona de conflicto.

La Actividad de Proyectos de Investigación Avanzados de Inteligencia (IARPA) quiso abordar estos temas a base de financiar un torneo de pronósticos geopolíticos, de cuatro años de duración. El torneo se llamaba la Estimación Agregativa de Contingencias[54] (ACE), y estaba diseñado para poner a prueba las mejores estrategias de elicitación y agregación de juicios de probabilidad. Cinco grupos basados en distintas universidades – incluyendo el proyecto *Good Judgement* [55] (GJP)– compitieron entre sí en el intento de desarrollar los métodos más precisos para hacer pronósticos sobre sucesos geopolíticos. IARPA estableció unas reglas básicas para el torneo:

- Preguntas: La mayoría eran preguntas binarias (por ejemplo, ¿abandonará el presidente de Siria, Bashar Hafez al-Assad, su cargo en los próximos tres meses?), otras eran preguntas multinomiales (por ejemplo, ¿quién será presidente de Rusia en el 2012?), otras eran preguntas condicionales (por ejemplo, ¿empezará el juicio de Ahmed Shafik antes del 1 de enero del 2013, si vuelve a Egipto antes de entonces?). Este conjunto de preguntas (unas 120 por año) se dividió en pequeños subconjuntos y fueron publicadas a lo largo de un periodo de nueve meses para cada uno de los cuatro años. La duración media de cada pregunta fue de 100 días.
- Analistas: Cada grupo de investigación tenía la responsabilidad de reclutar analistas para que hiciesen estimaciones probabilísticas sobre estas preguntas. Los analistas debían tener al menos 18 años, tener al menos el título de grado, y ser ciudadanos de los Estados Unidos; todas las características típicas de un analista de inteligencia. (Los analistas del GJP podían seleccionar las preguntas que quisiesen responder y actualizar sus previsiones todo lo que quisiesen mientras la pregunta estuviera sin resolver).
- Estimaciones de probabilidad: Cada día, los grupos de investigación tenían que presentar pronósticos agregados sobre el conjunto de preguntas. Tenían la posibilidad de presentar diez pronósticos agregados para cada pregunta, lo cual les permitía probar varios métodos distintos.
- Criterios de evaluación: el único criterio utilizado para medir resultados fue el índice de Brier (Brier, 1950), que determina la precisión de los pronósticos probabilísticos y fomenta la formulación de creencias verdaderas. Se calcula como la suma de desviaciones al cuadrado entre los pronósticos y la realidad (donde la realidad se codifica con un 1 en caso de que ocurra el suceso y 0 en caso contrario), con un rango de 0 (mejor resultado) a 2 (peor resultado).

54 https://www.iarpa.gov/index.php/research-programs/ace.

55 https://en.wikipedia.org/wiki/The_Good_Judgment_Project.

- *Benchmark:* IARPA quería un punto de referencia para poder juzgar el nivel de precisión de los distintos grupos en el contexto de las reglas del torneo. Para esto, contrataron a un grupo neutral para que reclutase a analistas parecidos que hiciesen pronósticos sobre las mismas preguntas a la misma vez. De este modo, utilizando la media sin ponderar para agregar los pronósticos, IARPA pudo comparar los métodos de los distintos grupos de investigación con los del grupo utilizado como punto de referencia.

El proyecto *Good Judgement*

El GJP fue concebido por Philip Tetlock y Barbara Mellers, de la Universidad de Pensilvania, y Don Moore, de la Universidad de California Berkeley. Juntaron a un grupo de psicólogos, informáticos, politólogos y economistas para que abordasen el reto propuesto por IARPA: *¿Cuáles son los mejores métodos para elicitar y agregar estimaciones probabilísticas a lo largo del tiempo?*

Dado que la literatura sobre los juicios de probabilidad es muy contradictoria (Dawes et al., 1989; Tetlock y Mellers, 2002) y que GJP tenía que afrontar el reto de realizar pronósticos longitudinales y en tiempo real, el proyecto decidió no adoptar ningún método establecido en concreto, sino adoptar la filosofía de las pruebas A/B para encontrar maneras de mejorar la precisión de los pronósticos.

Formación

Los estudios sobre cómo formar a las personas para que formulen juicios probabilísticos consistentes son muy controvertidos. En algunos casos, la formación demostró no ser efectiva, y en otros, conllevó a que se formulasen juicios menos exactos (Arkes, 1991). GJP decidió desarrollar y testar un programa de formación que se centrase en el razonamiento probabilístico, la proyección de escenarios, y la teoría política. El programa evolucionó a lo largo de los cuatro años, a medida que se testaban aspectos del mismo utilizando como base de referencia a un grupo de analistas que no recibió ninguna formación (Chang *et al*., 2016).

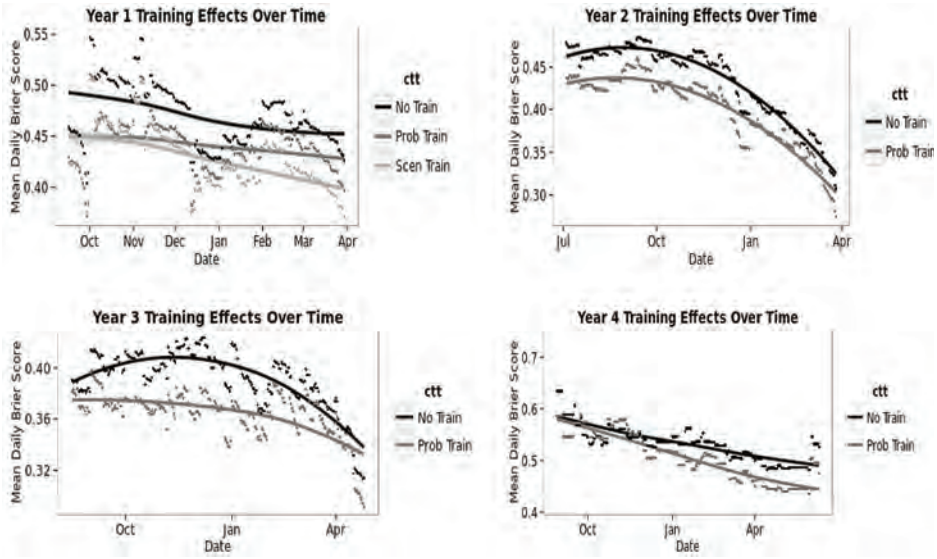
- Durante el primer año, el programa de formación consistió en dos módulos: razonamiento probabilístico y formulación de escenarios. El módulo sobre formulación de escenarios exigió a los analistas que desarrollasen probabilidades coherentes y lógicas, rigiéndose por la regla de la suma de probabilidades, así como que explorasen y cuestionasen sus propias

suposiciones, identificasen factores causales clave, considerasen el mejor y el peor de los casos, y evitasen sesgos que resultasen del exceso de corrección. La formación de razonamiento probabilístico se centró en conceptos como distinguir entre la calibración y la resolución, el uso de clases para hacer comparaciones, el uso de frecuencias base, utilizar y hacer el promedio de principios derivados de la sabiduría de grupo, encontrar y elaborar modelos predictivos estadísticos y matemáticos, utilizar series temporales de datos y datos históricos, y estar atentos a los sesgos cognitivos típicos en la mayoría de las personas.

- Durante el segundo año, se combinaron la formación en razonamiento probabilístico y la formación en formulación de escenarios.
- Durante el tercer año, se añadió algo de teoría política para ayudar a los analistas a entender las dinámicas entre actores políticos clave, determinar la influencia de las normas e instituciones internacionales, contrastar distintas perspectivas políticas, y estar atentos a potenciales escenarios imponderables.
- El cuarto año fue muy parecido al tercer año, pero se mejoraron los métodos de formación, incorporando más gráficos, evaluaciones de conocimiento interactivas, y un sistema de evaluación de aprendizaje.

A lo largo de la competición, la formación en razonamiento probabilístico afectó positivamente el nivel de precisión de los analistas. Los principios que tuvieron más efecto se derivaron del razonamiento probabilístico, específicamente del uso de clases para hacer comparaciones y de frecuencias base. Hubo diferencias importantes en términos de práctica deliberada entre los analistas que habían participado en el programa de formación y los que no, y el modelo de mediación se sostuvo durante tres de los cuatro años (ver Figura 1). Ambos grupos mejoraron su nivel de precisión a lo largo de estos años, pero los analistas que habían recibido formación se mantuvieron por delante de los que no durante todo el tiempo. Estos resultados resaltan la importancia de la práctica deliberada. Se confirma, por lo tanto, que hacer pronósticos es una habilidad que se puede desarrollar con el tiempo, y no un fenómeno que depende del azar.

Figura 1. Efecto del programa de formación a lo largo del tiempo



Selección

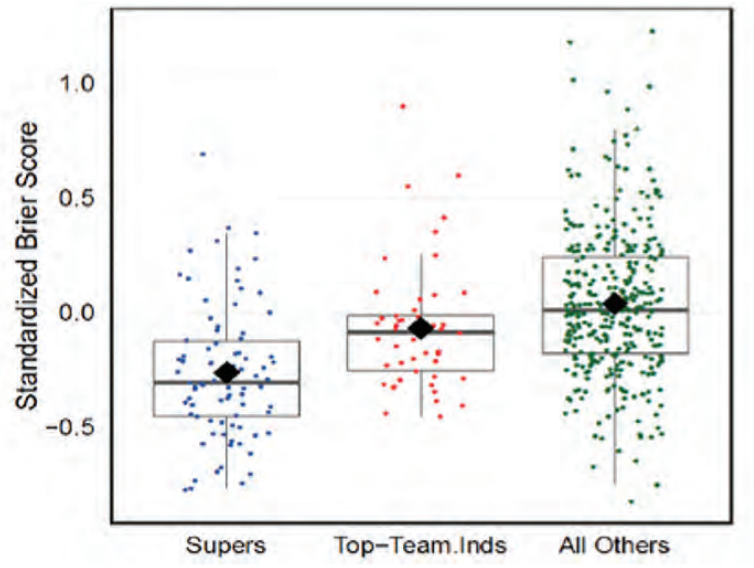
¿Es la idea del "Superpronosticador", una persona que consistentemente consigue predecir el futuro, un mito? Lo es. Los superpronosticadores no eran videntes, sino participantes en los cuales se fijó el GJP debido a sus habilidades para formular juicios probabilísticos más precisos que los de los demás sobre una gran cantidad de preguntas. A base de seleccionar a los mejores 2% de analistas de cada condición (es decir, los mejores 5 analistas de cada una de las 12 condiciones experimentales), se formaron equipos de Superpronisticadores en el segundo año para maximizar las sinergías que podían aportar estos analistas. La idea era que les estimularía mucho trabajar con otros analistas de alto rendimiento, que se beneficiarían de la profecía autocumplida de que se les describiese como Superpronosticadores, y que serían más inteligentes que los demás (Mellers *et al.*, 2015; 2017).

Los resultados demostraron que juntar a los mejores analistas en equipos mejoró el nivel de precisión de los pronósticos más que las dos manipulaciones experimentales — proporcionar formación y montar equipos — juntas o por separado. Superaron con creces a los otros equipos en los años que quedaban de torneo. Con el objetivo de demostrar la potencia de los equipos de Superpronosticadores, su nivel de precisión en el segundo y tercer año se comparó con el de otros dos grupos:

- Individuos de Alto Rendimiento: estos eran los analistas más precisos, detrás de los Superpronosticadores, que formaban parte de los equipos normales. Eran analistas de alto rendimiento que no fueron seleccionados al equipo de Superpronosticadores por un margen muy pequeño, y que trabajaron con otros analistas de alto rendimiento.
- Otros: Todos los demás analistas sirvieron de línea de base. Dado que los analistas tenían la opción de seleccionar las preguntas a las cuales quisiesen responder, sólo se incluyó en este grupo a aquellos que participaron en más de 25 preguntas.

La Figura 2 demuestra que tanto el equipo de Superpronosticadores como el de Individuos de Alto Rendimiento superó al grupo de Otros, es decir, ninguno experimentó una regresión a la media. Sin embargo, el equipo de Superpronosticadores superó con creces al de Alto Rendimiento, demostrando así la sinergía tan potente que creó el juntar a analistas tan afines.

Figura 2. Rendimiento de los Superpronosticadores comparado al de los otros grupos[56]



Índice de Brier Estandarizado

56 Reproducido a partir de Mellers et al., 2015

Agregación

GJP probó varios métodos de agregación a lo largo del torneo, incluyendo mercados de predicción (Atanasov *et al*., 2016) y modelos complejos de simulación. Un problema clave de la agregación de juicios probabilísticos fue el cómo ponderar adecuadamente los pronósticos y los analistas en los modelos. El Modelo Ponderado de Contribución (MPC) propone un nuevo método para identificar el nivel de conocimiento de cada analista basándose en la contribución relativa que hace al rendimiento colectivo (agregado) del grupo (Budescu y Chen, 2015). EL MPC demuestra que utilizar una medida relativa del conocimiento para ponderar diferentemente a los distintos analistas es más eficaz que otros métodos más tradicionales de ponderar juicios en el proceso de agregación a grupos, se mantiene robusto ante intentos de sabotaje por agentes hostiles, y es un método más económico (Chen *et al.* 2016).

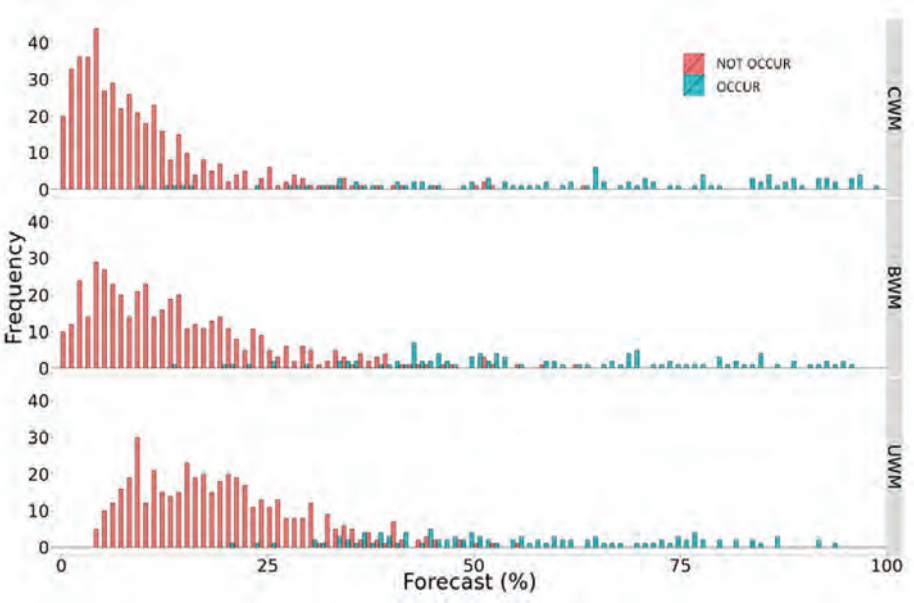
EL MPC nos enseña una manera estratégica de utilizar la sabiduría en grupo. El objetivo es utilizar las evaluaciones del rendimiento de los expertos comparado al del grupo (es decir, su contribución individual) com método para ponderar sus pronósticos probabilísticos. La medida de contribución se utiliza para dar más peso a aquellos expertos que hagan las mayores contribuciones o dar menos peso/eliminar a aquellos cuyas contribuciones sean menores, creando así un grupo más pequeño y más selecto. El modelo se ajusta a la información que proporciona cada analista, al tipo de suceso, y a la duración del periodo de formulación de los pronósticos. Es más eficaz que cualquier otro método de agregación y se puede actualizar de forma dinámica para reflejar cualquier cambio en las contribuciones individuales de cada experto; es decir, el modelo puede detectar el nivel de conocimiento de cada experto instantáneamente y utilizar esa información para mejorar los pronósticos sobre nuevos sucesos.

La Tabla 1 y la Figura 3 ilustran nuevos aspectos acerca de la superioridad de los pronósticos que utilizan el MPC frente a aquellos que utilizan otros métodos. El método UWM es la media sin ponderar del grupo. El método BWM pondera al grupo utilizando su rendimiento total (es decir, el Índice de Brier). La figura muestra la distribución de pronósticos probabilísticos sobre todos los sucesos formulados utilizando los tres modelos en los años 2 y 3. Las dos distribuciones (rojas y azules) representan sucesos que se pronosticó que no ocurrirían (81%) o que sí ocurrirían (19%). La distribución de los pronósticos del MPC sobre sucesos que no sucedieron (sucedieron) se acerca más al 0 (1), lo cual significa que este modelo distingue mejor entre estos sucesos. La DI' del MPC (Wallsten *et al.*, 1997), que muestra la diferencia entre las estimaciones medias de la distancia estandarizada entre las medias de las distribuciones rojas y azules, es la más grande.

Tabla 1. Resumen de las distribuciones de los pronósticos de los 3 modelos sobre sucesos que resultaron verdaderos o falsos

Clase de Sucesos	Número de sucesos (n=142)	UWM	BWM	CWM
		Media (y Desviación Típica) de Pronósticos	Media (y Desviación Típica) de Pronósticos	Media (y Desviación Típica) de Pronósticos
Falsos	115	19.80 (9.53)	14.99 (11.76)	10.47 (9.63)
Verdaderos	27	55.95 (17.28)	58.73 (20.29)	63.82 (23.85)
DI'= (Mverdaderos – Mfalsos) / Desviación Típica Agrupada		3.18	3.18	3.96

Figura 3. Análisis Discriminante de los 3 modelos[57]



57 Reproducido a partir de Chen y Budescu, 2016.

Entre la comunidad de inteligencia existe la preocupación de que un agente hostil quiera sabotear deliberadamente su trabajo. Si un espía consiguiese infiltrarse en tu grupo y hacer pronósticos erróneos, ¿conseguiría afectar negativamente a los pronósticos agregados? Este modelo se distingue de otros métodos de agregación en que se mantiene robusto ante cualquier intento de desinformación deliberada, que podría producirse en el caso de que un agente hostil intentase sabotear los pronósticos haciendo predicciones erróneas.

El ejercicio de recoger pronósticos e información de varios expertos puede ser muy caro, por lo cual se utilizan las ponderaciones de contribución para formar un subconjunto óptimo de expertos y saber cuándo descartar a ciertos expertos. Por lo general, hay que elegir entre tener un mayor número de expertos, cuyos conocimientos forman parte del agregado final, o entre que la operación sea más económica. El modelo proporciona una medida más eficaz que la del rendimiento individual para elegir los expertos a descartar y el número de preguntas de formación requeridas. Puede haber dos expertos muy eficaces que, sin embargo, proporcionen la misma información, por lo que puede no compensar retener a los dos.

El modelo es fácil de implementar, robusto, reactivo a cambios en el entorno, y económico. Las ponderaciones son proporcionales a las notas de contribución de cada experto, pero sólo si se utilizan a expertos que hagan contribuciones positivas. De este modo, en un momento dado, sólo se está utilizando a la mitad de los expertos a la vez pero, como hemos descrito antes, la identidad de los expertos a los cuales se les atribuye una ponderación positiva puede ir cambiando de un momento a otro. De esta forma, los pronósticos agregados son verdaderos pronósticos de grupo que utilizan al mejor subconjunto de expertos en cada momento.

Conclusión

Al final del segundo año, el GJP no sólo superó el *benchmark* de precisión establecido por la IARPA por un 50%, sino que lideraba la competición. Al tercer año, IARPA eliminó al otro grupo de la competición y dedicó todos sus recursos al GJP para que expandiesen su investigación de la formulación de pronósticos a áreas tales como la inteligencia artificial, la teoría de respuesta al ítem, y las redes bayesianas, entre otros. Cuando a los investigadores principales del GJP se les pidió que describiesen el factor clave que explicaba el éxito de sus investigaciones, apuntaron al hecho de que la estructura de un torneo introduce un elemento de responsabilidad en las previsiones de las organizaciones (Tetlock *et al*. 2014). Durante un torneo, los

analistas formulan previsiones que luego son puntuadas y medidas para determinar su nivel de precisión. Los torneos motivan a los analistas a que hagan pronósticos exactos, en vez de proporcionar incentivos perversos que fomenten las intrigas internas, las jerarquías, o el que los analistas se dediquen a culparse unos a otros por haber malinterpretado ciertas señales. De este modo, si lo único que se mide es tu nivel de precisión, ganas a base de ser preciso.

Los descubrimientos sobre los pronósticos precisos de este proyecto están siendo comercializados por la empresa Good Judgement (www. goodjudgement.com).

Bibliografía

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological bulletin*, 110(3), 486.

Atanasov, P., Rescober, P., Stone, E., Swift, S.A., Servan-Schreiber, E., Tetlock, P., & Ungar, L., Mellers, B. (2016). Distilling the Wisdom of Crowds: Prediction Markets vs. Prediction Polls. *Management Science*, 63(3) 691-706.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly weather review, 78*(1), 1-3.

Budescu ,D.V., & Chen, E. (2015). Identifying expertise to extract the wisdom of crowds. *Management Science, 61*(2), 267–280.

Chang, W., Chen, E., Mellers, B. A., & Tetlock, P. E. (2016). Developing expert political judgment: The impact of training and practice on judgmental accuracy in geopo- litical forecasting tournaments. *Judgment and Decision, 11*(5), 509–526.

Chen, E., Budescu, D.V., Lakshmikanth, S.K., Mellers, B.A., & Tetlock, P.E. (2016) Validating the contribution-weighted model: Robustness and cost-benefit analyses. *Decision Analysis, 13*(2), 128–152.

Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 2: impediments to and strategies for change. *BMJ Quality & Safety*, bmjqs-2012-001713.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668-1674.

Jolls, C., & Sunstein, C. R. (2004). Debiasing through law. Retrieved from https:// chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1157&context=law_and_ economics, January 31, 2019.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological review*, 80(4), 237.

Kahneman, D., & Tversky, A. (1977). *Intuitive prediction: Biases and corrective procedures*. Retrieved from

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. *cognition, 11*(2), 123-141.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39(4), 341.

Mellers, B., Stone, E., Murray, T., Minster, A., Rohrbaugh, N., Bishop, M., . . . Horowitz, M. (2015). Identifying and Cultivating Superforecasters as a Method of Improving Probabilistic Predictions. *Perspectives on psychological science*, 10(3), 267-281.

Mellers, B. A., Baker, J. D., Chen, E., Mandel, D. R., & Tetlock, P. E. (2017). How generalizable is good judgment? A multi-task, multi-benchmark study. *Judgment and Decision Making*, 12, 369–381.

Morewedge, C. K., & Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends in cognitive sciences*, 14(10), 435-440.

Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance, 3*(4), 544.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Tetlock, P.E., & Kim, J. I. (1987). Accountability and judgment processes in a personality prediction task. *Journal of Personality and Social Psychology, 52*(4):700–709.

Tetlock, P. E., & Mellers, B. A. (2002). The Great Rationality Debate. *Psychological Science, 13*(1), 94–99.

Tetlock, P.E., Mellers, B., Rohrbaugh, N. & Chen, E. (2014). "Forecasting tournaments: Tools for increasing transparency and the quality of debate". *Current Directions in Psychological Scienc,. 23*(4): 290–295.

Wallsten, T.S., Budescu, D.V., Erev, I. & Diederich, A. (1997). Evaluating and combining subjective probability estimates. *Journal of Behavior and decision Making, 10*, 243-268.

Eva Chen es científica jefe de la startup Good Judgment. Su investigación se centra en la identificación de técnicas de vanguardia para la predicción de eventos geopolíticos y el crowdsourcing para generar una amplia gama de productos analíticos para los servicios de inteligencia en EEUU. La Dra. Chen obtuvo su doctorado en Ciencias de la Decisión y Ciencias Económicas en la Universidad de Concordia (Montreal, Canadá). Recientemente la Dra. Chen ha recibido el Premio Exeter de Investigación en Economía Experimental, Teoría de la Decisión y Economía del Comportamiento en el año.

Italia: la aplicación del programa de análisis predictivo *Delia*® en la inteligencia policial

Virginia Cinelli

Introducción

El oficial de inteligencia estadounidense Richard J. Heuer plantea en el libro *Psychology of Intelligence* que "por definición, los problemas de inteligencia tienen un cierto grado de incertidumbre" y la tarea del analista "es trascender los límites de la información incompleta aplicando su juicio analítico" (Heuer, 1999). Para hacer frente a esta incertidumbre y al aumento del volumen de información disponible, desde finales de los años 90 la tecnología y la estadística han sido piezas clave para la investigación en el ámbito de la seguridad, ayudando a potenciar los límites humanos en el proceso de obtención de información y mejorando exponencialmente su capacidad de análisis.

En el sector de la seguridad, la tecnología es ampliamente utilizada en el marco preventivo con el objetivo de calcular la probabilidad de que ocurra, o no, un suceso o crimen. A este respecto, se entiende por análisis predictivo "el uso de técnicas de análisis, en particular técnicas cuantitativas, para identificar objetivos potenciales que requieren la intervención policial, además de prevenir o resolver crímenes pasados mediante pronósticos estadísticos" (Perry *et.al*., 2013). El precursor de esta técnica es el *software Compstat*, el cual fue desarrollado en 1994 por el Departamento de la Policía de Nueva York. Posteriormente, muchos otros sistemas fueron implementados a nivel mundial, incluyendo *i2 Analyst Notebook* (Francia), *PredPol* (Estados Unidos y Reino Unido), *Crime Anticipation System* (Países Bajos), *Precobs* (Alemania) o *Policy Cloud System* (China).

En el caso de Italia, el primer *software* de análisis predictivo, denominado *Delia*®, fue desarrollado en 2007 en el cuartel general de la policía de Milán por el entonces asistente de la policía estatal, Mario Venturi.[58] Dicho programa de inteligencia

58 *Delia*® no es el único *software* de análisis predictivo presente en Italia. Se tiene constancia de por lo menos otros dos modelos que se testaron y/o desarrollaron en territorio italiano. En 2014, tras una colaboración entre el centro de investigación Transcrime de la Universidad Cattolica del Sacro Cuore de Milán, la Universidad de Trento y el Ministerio de Interior de Italia, se aplicó el modelo predictivo *Risk Terrain Modeling* (RTM) a las ciudades de Milán, Roma y Bari con el objetivo de aumentar la capacidad de predecir robos en vivienda (Dugato *et al.*, 2015). De igual manera, en diciembre de 2018, la policía de Nápoles testó por primera vez *X Law*, un programa para predecir el lugar de comisión de robos (La Stampa, 2019).

integra una actividad de cálculo de *big data* que sería capaz de detectar crímenes en serie y predecir *dónde, cuándo* y *cómo* podría tener lugar el próximo acto delictivo.

El objetivo de este artículo es examinar el caso innovador de *Delia*® en el marco italiano, analizando su potencial y planteando la hipótesis de su aplicabilidad a nivel nacional para combatir la criminalidad transnacional. Para ello, el presente capítulo constará de tres secciones diferentes. En la primera parte se señalarán las fortalezas del *software*, prestando especial atención al proceso analítico adoptado para la prevención y persecución de casos reales. En segunda instancia se explicará cómo el programa, tras ser testado en un estudio de caso a nivel micro (la ciudad de Milán), es capaz de mejora de manera significativa el rendimiento de la investigación y posterior actuación de los agentes. En última instancia, se analizará la posibilidad de utilizar *Delia*® en un nivel más complejo como un sistema con capacidad para pronosticar y combatir los delitos transnacionales.

Delia®: inteligencia entre el analista y la máquina

Aunque el proceso de aplicación de *Delia*® tiene similitudes al de un *software* de análisis predictivo común –cuyo funcionamiento podría dividirse en las fases de recopilación de datos, análisis, fase operativa y respuesta criminal (es decir, cómo cambia el comportamiento y las decisiones del criminal tras la operación policial)– su relevancia reside en sus particularidades (Perry *et al*., 2013).

En primer lugar, el programa se caracteriza por su complejidad analítica. Al igual que el cerebro humano, *Delia*® trabaja con datos previamente almacenados, aplicando durante su procesamiento prácticas y herramientas extraídas de disciplinas como las matemáticas, la estadística, la psicología del comportamiento o el análisis geoespacial. El *software* es capaz de procesar hasta un millón y medio de variables, focalizándose en dos tipos de datos; simples y genéricos (por ejemplo, hora del acto criminal, lugar, día, etc.) o sofisticados y detallados (características y comportamiento del delincuente). A su vez, elabora las premisas de la investigación en función de cuatro elementos fundamentales de cada crimen: su tipología, el objetivo, el *modus operandi* (incluyendo objetos, armas y medios de transporte empleados) y las características psico-físicas del autor (como ropa, tatuajes, piercings, cicatrices, o cualquier objeto visible que pueda identificarlo). El enfoque analítico de la investigación permite trabajar con la información utilizando un proceso de

análisis complejo que cuenta con dos fases fundamentales: una primera inductiva,[59] en la que se analiza con detalle un crimen concreto para identificar los elementos comunes con otros sucesos de características similares y relacionarlos con un único delincuente; y una segunda deductiva,[60] que permitiría tras observar los elementos clave identificados en la serie criminal, predecir *cuándo, dónde* y *cómo* se cometerá el delito futuro.

El segundo elemento al que se hace referencia es el hecho de que el *software* funciona únicamente si está en relación continua con el analista, factor que reduce el margen de error del análisis y de sus resultados. En su publicación, Heuer identifica algunos de los elementos que suelen presentar más dificultades durante el análisis de inteligencia. Entre ellos, considera que la mayoría de los fallos en el proceso analítico están causados por tres aspectos: la limitada capacidad de la memoria humana, la dificultad del cerebro humano de evaluar más de siete hipótesis múltiples y competitivas a la vez, y la posibilidad de subestimar, malinterpretar, ignorar o excluir una información relevante. En este sentido, el procesador de la máquina permite acumular cantidades de datos *a priori* inabarcables para el ser humano y trabajar con ellos de manera simultánea. Dicho proceso evita la perdida de información, limitando la posibilidad de que ninguna variable insertada en el sistema sea excluida y comparándolas todas de manera homogénea.

Así como Heuer identifica las debilidades del analista a la hora de elaborar inteligencia, también destaca la imaginación y la creatividad como dos elementos fundamentales para el propio éxito del análisis. En el caso de *Delia*®, el analista que interactúa con la máquina no perdería tales características. A modo de ejemplo, durante la fase de identificación de la serie criminal y después de una comparación de todos los crímenes archivados llevada a cabo por el programa, es el analista quien se ocupa de 'aceptar', 'descartar' o señalar como 'probable' la similitud entre dos crímenes, rebajando de esa manera el margen de error de análisis de la máquina al introducir un elemento de subjetividad y pensamiento crítico en el proceso de decisión.

Más aún, dicha interacción permite a su vez aplicar tres de las técnicas más utilizadas en inteligencia en un mismo análisis, como son la lógica situacional, la comparación y el uso de creencias y conocimientos del analista (también llamadas teorías). Raramente un analista aplica dichas técnicas a la vez, ya sea porque la persona no

59 Se define como 'análisis inductivo' al método científico que obtiene conclusiones generales a partir de premisas particulares (Moore, 2010).

60 Se define como 'análisis deductivo' al método científico que obtiene conclusiones particulares a partir de premisas generales (Moore, 2010).

es partidaria de alguna de ellas o porque carece del tiempo necesario para hacerlo. En el caso de *Delia*®, el programa procesa la información de manera muy similar a las técnicas de lógica situacional y comparación, mientras que es el analista, con su conocimiento del fenómeno y su capacidad analítica, quien llega a las conclusiones mediante la aplicación de las teorías.

El tercer elemento que define la complejidad del *software* es el hecho de que vaya más allá del elemento predictivo, influyendo tras su aplicación en las tres áreas del delito: la prevención, la persecución del delito y la parte legal-procesal posterior al crimen. Si bien ya hemos profundizado en el rol de *Delia*® a nivel de prevención, faltan por examinar los otros dos aspectos. Con respecto a la persecución del delito, el programa actuaría sobre dos factores. Por un lado, haciendo un uso operativo de los detalles probabilísticos identificados por el *software*, los agentes de campo tendrán más posibilidades de arrestar al sospechoso, poniendo fin a su actividad criminal. Por otro lado, tal y como se demostrará en el caso de estudio, *Delia*® permitiría influir directamente en el número de delitos que componen la serie criminal, reduciéndola drásticamente (Mastrobuoni, 2017). Por último, a nivel legal *Delia*® influye sobre un factor relevante. Desde un punto de vista procesal, ya que la identificación de un crimen en serie permite abordar los casos no de manera individual, sino agrupándolos en una misma causa penal. Este hecho permite que, a la hora de procesar al perpetrador, la acción penal sea más efectiva y eficiente, pudiendo enjuiciar al autor por un mayor número de delitos y con una condena más extensa.

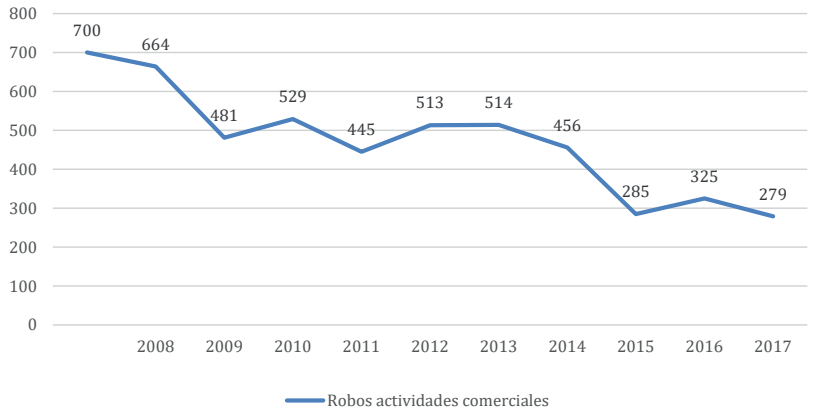
Caso de estudio: *Delia*® aplicado a los robos de farmacias en Milán

Más allá de las fortalezas que el *software* pueda presentar, es necesario comprender la funcionalidad efectiva del mismo para evaluar su potencial de implementación a nivel nacional y para hacer frente a crímenes transnacionales. Desde 2008, *Delia*® ha estado sujeto a dos fases de experimentación distintas en el cuartel general de la policía de Milán, además de pasar por la evaluación empírica de personal externo al desarrollo del programa. En un primer momento, se puso en marcha en la propia ciudad[61] para combatir el fenómeno de los robos contra establecimientos comerciales. Esta decisión vino fundamentada por el gran número de negocios

61 En 2017 Milán era la segunda ciudad italiana por número de habitantes (1.380.873), con una superficie de 181,67 km2.

afectados por este ilícito, alrededor de 700 casos al año (Venturi, 2014; Taurisiano, 2016) y del carácter dual del delito de atraco, que lo hizo una tipología delictiva particularmente interesante de analizar, al tratarse al mismo tiempo de una acción contra la persona y contra la propiedad. Desde su primer año de experimentación y hasta 2017, se registró una disminución casi constante en el número de robos, así como un aumento de los casos resueltos por *Delia*® (Fig. 1 y 2). Entre todos los datos, destaca el resultado del primer año de aplicación del *software*, cuando se registró un descenso del 28% en el número de robos (de 664 a 481). Debido al éxito de la primera prueba, se decidió extender el uso del programa a la totalidad de la provincia de Milán (3.234.658 habitantes)[62] y expandirlo para analizar los robos perpetrados contra entidades bancarias. Desde su aplicación en este escenario, los resultados obtenidos mejoraron de manera significativa, situando ya en el primer año, la tasa de casos resueltos en un 37%.

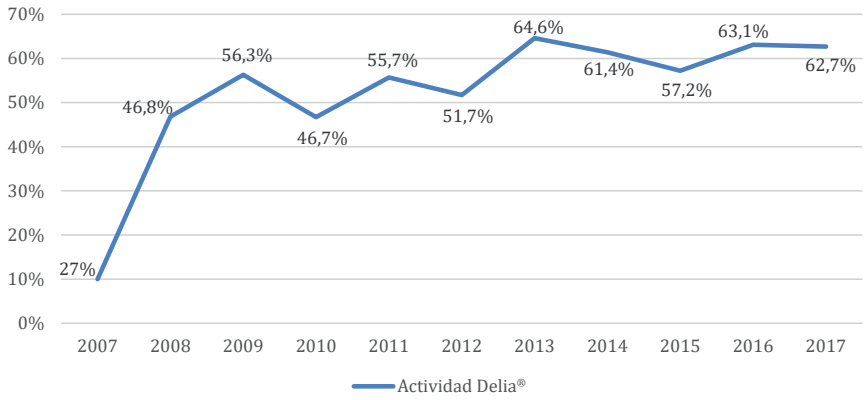
Figura 1. Robos a establecimientos comerciales perpetrados entre 2007 y 2017



Fuente: keycrime.com

62 Fuente ISTAT, fecha de actualización 01/01/2019, Disponible en: http://dati.istat.it/Index.aspx?DataSetCode=DCIS_POPRES1

Figura 2. Robos resueltos por *Delia*® entre 2007-2017[63]



Fuente: keycrime.com

Tal y como se ha mencionado anteriormente, *Delia*® es uno de los pocos programas de análisis predictivo que ha sido evaluado por personal externo al desarrollo del mismo. El economista experto en seguridad Mastrobuoni (2014) estudió la aplicación del *software* al caso de Milán, corroborando los datos obtenidos y la razón de la eficacia de *Delia*®. Tras un análisis comparativo de los dos cuerpos policiales, Carabinieri y Policía, actuando en una misma zona, uno de ellos utilizando el programa y el otro aplicando técnicas de investigación tradicionales, Mastrobuoni demostró que *Delia*® aumentó en un 8% la probabilidad de resolver un crimen en serie, disminuyendo el número de robos que cada (grupo) criminal es capaz de perpetrar antes de su detención, pasando de una media de 17,8 delitos a 6,4 y con un máximo de 11 robos por serie. En el caso de los robos a bancos, el dato es aún más interesante ya que se registró un cambio de tendencia criminal a nivel local (Mastrobuoni, 2017). Tras analizar el número de robos a entidades bancarias perpetrados mensualmente entre los años 2004 y 2015 en las diez mayores ciudades de Italia (Milán, Turín, Génova, Bolonia, Verona, Venecia, Florencia, Roma, Nápoles y Bari), Milán resultó ser la ciudad con un mayor número de casos de este delito. Sin embargo, durante los años de aplicación del programa (a partir de 2008), esta tendencia se invirtió logrando pasar de una tasa de robos de 1,4 a 0,5 por cada 100.000 habitantes.

63 Ibid, p.91.

Mastrobuoni argumenta que la razón detrás de los resultados obtenidos por *Delia*® se fundamenta en la capacidad del *software* para identificar algunos de los factores clave que determinan la "persistencia" de un crimen. Según su análisis, algunos elementos son más relevantes que otros a la hora de explicar que un crimen permanezca inmutable en el tiempo y en el espacio, ya que "los grupos criminales suelen seleccionar el mismo tipo de objetivo, a la misma hora del día y en la misma zona, sobre todo si el primer crimen ha sido exitoso". Así, toda información relativa a la hora, el área de acción, la distancia entre el primer crimen y el segundo y el objetivo del crimen, son elementos que *Delia*® toma en consideración desde su primera fase de análisis, integrándola en la totalidad del proceso analítico.

Otro dato relevante del análisis de Mastrobuoni es el beneficio económico del uso del programa. Según el autor, aplicado en los tres niveles del delito (prevención, persecución del delito y la parte procesal-legal) *Delia*® lleva a una reducción del coste y de los considerables esfuerzos operacionales, además de agilizar la fase procesal y penal del caso. A este respecto, se ha calculado que en Italia la aplicación del *software* permitiría ahorrar alrededor de 1,2 millones de euros por año, incluyendo los costes de personal y de gestión del programa, además del gasto que conlleva el encarcelamiento de un mayor número de criminales (Mastrobuoni, 2014). De hecho, la aplicación de *Delia*® no implica una inversión adicional en personal, ya que a nivel operativo el programa permite una acción más precisa de las patrullas, las cuales serán enviadas a las áreas de mayor riesgo seleccionadas previamente. Asimismo, menos personal legal estará involucrado en la resolución de un caso, haciendo los procesos más agiles y más precisos, acabando en penas más largas y ayudando a la reducción de la tasa de criminalidad.

Delia® en perspectiva transnacional

Una de las principales ideas planteadas en el capítulo es el hecho de que los datos presentados hasta la fecha hacen alusión a una fase de desarrollo del *software* en la cual no había sido aplicada ninguna tecnología avanzada. En 2017, con el objetivo de hacer que el programa fuera extrapolable a nivel global, el creador de *Delia*®, Mario Venturi, alcanzó un acuerdo con IBM para aumentar el potencial de la herramienta mediante la implementación de inteligencia artificial.[64] Dicho avance ha permitido evolucionar la fase de recopilación, mejorando la calidad y cantidad de datos que

64 En febrero de 2018, el inventor de *Delia*® Mario Venturi decidió fundar una *start up* para desarrollar el programa, llamándola "Keycrime S.r.l".

pueden analizarse, además de crear un modelo para la gestión de la información y expandir el uso del *software* en la investigación de todo tipo de crímenes en serie. [65]

Sobre esta base, la aplicación del programa a delitos más complejos es ciertamente prometedora. Más allá del desarrollo tecnológico, se considera que, por naturaleza, *Delia*® explotaría aún más sus capacidades frente a fenómenos como la criminalidad transnacional. La argumentación de esta hipótesis se apoya en tres elementos principales (Guida, 2018). Primero, el programa, aplicado a una zona geográfica más extensa, permitiría aumentar la precisión del resultado predictivo dilatando las dinámicas criminales y representando de manera más clara tanto la forma como la localización específica de los flujos criminales. Segundo, cuanto más grande sea el espacio geográfico, más elevada será la incidencia criminal en el mismo; en otras palabras, un mayor número de delitos cometidos proporciona más información acerca de las características operacionales y de la composición de la organización criminal. Por último, la complejidad de un grupo organizado facilita el realizar un perfilado preciso del delincuente, ya que cuantos más actores participan en un hecho delictivo, más extensa y reconocible será la "firma" criminal (Mastrobuoni, 2017). Teniendo en cuenta lo expuesto hasta ahora, en un contexto de lucha contra fenómenos transnacionales, será más fácil identificar una organización que tenga sus bases logísticas en zonas geográficas extensas (como por ejemplo Oriente Medio o África Subsahariana) y, por tanto, con dinámicas organizativas complejas.

A estos tres elementos se añade una reflexión acerca del tipo de fuentes a utilizar para recopilar la información e insertarla en *Delia*®, y de cómo esto puede afectar al proceso de análisis. Si bien a nivel municipal en los casos de robo sólo se toman en consideración tres tipos de fuentes (informes policiales, entrevistas a las víctimas y grabaciones de las cámaras de seguridad), en lo relativo a crímenes transnacionales serán necesarias otras fuentes adicionales. En primer lugar, suponiendo que las fuentes mencionadas sean suficientes para el análisis de un crimen transnacional, éstas no están disponibles de igual manera en todos los países. Por ejemplo, respecto a los informes policiales y a las denuncias por parte de la víctima, es necesario que exista un determinado nivel de confianza entre la comunidad y la policía, o bien que ésta no sea de ninguna manera corrupta y que la comunidad esté educada en denunciar los hechos ilícitos. En cuanto a las imágenes de las cámaras de seguridad, las mismas no están presentes en todas las ciudades ni en todos los países en igual medida. En caso de ausencia de alguna de estas fuentes, sería oportuno adaptar el

65 Entrevista con Mario Venturi, 14 de enero 2019.

espacio de actuación (es decir, una ciudad, un pueblo o un estado) a las necesidades de *Delia*®, lo que requiere una inversión de capital humano y económico que puede no estar disponible. En segundo lugar, si consideramos la fase de compartir y recopilar la información, se necesitarían datos que no son posibles de recopilar sin la ayuda de oficiales de inteligencia, policía y fuerzas armadas sobre el terreno, además de agentes locales en todos los países involucrados. Por último, sería necesario el intercambio de inteligencia entre la policía local y estatal, siendo más fácil que tenga lugar este flujo de información si ya hay organizaciones que se ocupan de este tipo de actividad.

Si bien el propósito de su creador es aplicar el programa a nivel mundial, las variables expuestas nos hacen pensar que, al menos a corto plazo, su aplicabilidad sólo es factible si se respetan ciertas condiciones, incluyendo que el país o ciudad donde *Delia*® se implemente cuente con un nivel suficiente de organización estructural (administración estatal, fuerzas de seguridad, infraestructuras y composición social) y disponga de las capacidades técnicas para llevar a cabo la recopilación de la información necesaria. A raíz de este razonamiento, se concluye que el programa tiene potencial para ayudar en la lucha contra crímenes transnacionales siempre y cuando al utilizarlo se haga referencia a dos análisis: uno a nivel local, realizado con técnicas tradicionales y que se enfoque sobre los crímenes perpetrados en esa zona concreta, extrayendo así información sobre el *modus operandi* y la composición de un grupo criminal; y un segundo análisis a nivel internacional y transnacional que se ocupe, utilizando *Delia*®, de procesar toda la información e identificar la dinámica serial.

Partiendo de estas premisas, desde una perspectiva nacional, en Italia se plantea una primera implementación de *Delia*® en la lucha contra el terrorismo. Este fenómeno es suficientemente complejo como para contrarrestarlo utilizando un programa de inteligencia de estas características y cuenta con una base legal propicia a la colaboración institucional en materia de inteligencia.[66] Desde una perspectiva internacional, el caso de la UE es prometedor, pues los países miembros cuentan con organización estructural suficiente para conseguir las fuentes de información necesarias, existiendo ya una fuerte red de compartición de datos para contrarrestar crímenes transnacionales mediante las agencias de Europol, Interpol y Frontex.

66 La ley 124/2007 establece que, para permitir una actividad de prevención efectiva, es necesaria la colaboración entre los servicios de inteligencia (DIS, AISE, AISI), las Fuerzas Armadas y las fuerzas y cuerpos de seguridad.

Conclusión

El sector de la seguridad está tendiendo hacia un cambio debido a la continua evolución de nuevas tecnologías y la aplicación de la inteligencia artificial en los procesos de gestión de la información. El desarrollo e implementación de programas de análisis predictivo a nivel mundial es una buena muestra de dicha tendencia. Si bien es cierto que aún se plantean dudas sobre algunos aspectos de esta tecnología emergente, el presente capítulo ha dado visibilidad a las oportunidades asociadas a su desarrollo.

El caso de *Delia*® en Italia es un ejemplo particular de cómo la tecnología puede potenciar las capacidades humanas de análisis e investigación criminal. Más allá del potencial que el programa ha demostrado tener a nivel operacional y económico a escala local y, posiblemente, a nivel nacional para contrarrestar crímenes transnacionales, se destacan tres factores que resultan ser clave en los programas de análisis predictivo: el enfoque analítico, la relación continua existente entre analista y máquina y la capacidad de influir en las tres esferas del delito. Estos elementos ejemplifican un paradigma que el mundo de la inteligencia está asumiendo, cada vez más, de compaginación entre el hombre y la tecnología, permitiendo reducir el margen de error del análisis y de sus resultados.

Bibliografía

Dugato, M., Caneppele, S., Favarin, S. & Rotondi, M. (2015), Prevedere i furti in abitazione, Transcrime Research in Brief – Serie Italia - N.1/2015, Trento (IT): Transcrime – Università degli Studi di Trento

Guida, Emilio. 2018. "Predictive Analysis, L'Approccio Logico Investigativo Nei Modellipredittivi E Il Ruolo Dell'Ampiezza Geogr Afica Nelle Dinamiche Regionali". *Academia.Edu.* http://www.academia.edu/37889709/articolo_2018_EG.pdf.

Heuer, Richards J. 1999. *Psychology Of Intelligence Analysis.* Center for the Study of Intelligence.

La Stampa (2019), "X Law il software che prevede i reati: a Napoli giá un arresto in flagranza", enero 2019. https://www.lastampa.it/2019/01/18/italia/x-law-il-software-che-prevede-i-reati-attivo-in-italia-ecco-come-funziona-mRNATzZsdbmNnqcsb7EGoK/pagina.html

Lim, Kevjn. 2015. "Big Data And Strategic Intelligence". *Intelligence And National Security 31* (4): 619-635. doi:10.1080/02684527.2015.1062321.

Mastrobuoni, Giovanni. 2017. *Crime Is Terribly Revealing: Information Technology And Police Productivity.* University of Essex. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2989914.

Mastrobuoni, Giovanni. 2014. *Crime Is Terribly Revealing: Information Technology And Police Productivity.* University of Essex. http://www.tinbergen.nl/wp-content/uploads/2015/02/Crime-is-Terribly-Revealing_-Information-Technology.pdf.

Moore, David T. 2010. *Critical Thinking And Intelligence Analysis.* Washington, DC: Center for Strategic Intelligence Research, National Defense Intelligence College.

OSSIF, 2018. *Rapporto Intersettoriale sulla Criminalità Predatoria.* Roma: ossif.it. http://www.ossif.it/public/pagine/96/400/18-rapporto-intersettoriale-interno-vers2.pdf.

Ley del 3 de agosto 2007, n° 124, *Sistema Di Informazione Per La Sicurezza Della Repubblica E Nuova Disciplina Del Segreto.*

Perry, W. L., B. McInnis, C. C. Price, S. C. Smith, and J. S. Holliwood. 2013. *Predictive Policing, The Role Of Crime Forecasting In Law Enforcement Operations.* Santa Monica: RAND Institute. https://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233

Taurisano, Glicerio. 2019. *Delia®: Il "Conclusive Reasoning" Nell'attività Anticrimine Della Polizia Di Stato.* PROFILING - I profili dell'abuso. http://eprints.bice.rm.cnr.it/15156/1/Delia®%20il%20conclusive%20reasoning.pdf.

Taurisano, Glicerio. 2016. "*Delia®*: Scienza E Progresso Per La Predictive Policing Italiana.". *Crimelist.It.* http://www.crimelist.it/index.php/analisi/criminologia/868--Delia®-scienza-e-progresso-per-la-predictive-policing-italiana.

Venturi, Mario. 2014. "*Delia®©* – La Chiave Del Crimine". *Profiling - I Profili Dell'abuso.* https://www.onap-profiling.org/Delia®-la-chiave-del-crimine/.

Sitografía

http://www.comune.milano.it/wps/portal/ist/it/amministrazione/datistatistici/Popolazione_residente_a_Milano

https://www.emmeviemme.com

http://dati.istat.it/Index.aspx?DataSetCode=DCIS_POPRES1

https://www.keycrime.com/

Entrevistas

Entrevista con Mario Venturi, 14 de enero de 2019.

Entrevista con Emilio Guida, 11 de enero de 2019.

Virginia Cinelli ha trabajado como gestora de proyectos del Real Instituto Elcano hasta febrero de 2021, coordinando la gestión del proyecto MINDb4ACT sobre radicalización en Europa, del proyecto J-SAFE sobre radicalización en prisiones, y del proyecto BRIDGE sobre polarización a nivel local. Anteriormente trabajó en el Gabinete de comunicación de la ONG italiana Cospe Onlus (Florencia). Grado en Relaciones Internacionales y en Ciencias Políticas por la Universidad LUISS Guido Carli de Roma (Italia), Máster en Estudios de Terrorismo y Seguridad Internacional por la Universidad King's College de Londres (Reino Unido), y Diplomado de Postgrado de Inteligencia y Estudios de Seguridad Nacional por la Universidad de Florencia (Italia).

El sistema de evaluación de riesgos en los Países Bajos

José Pablo Martínez

Introducción

Desde finales de la década pasada estamos siendo testigos del interés por parte de cada vez más países en implementar metodologías con base científica que ayuden a detectar y clasificar las diversas amenazas a las que tienen que hacer frente, con el objetivo de maximizar la eficiencia de los recursos destinados a atenuar los escenarios de riesgo más plausibles. Reino Unido, Estados Unidos, Canadá, Alemania, Suiza, Suecia, Noruega, Australia o Nueva Zelanda son algunos de los países cuyos gobiernos ya llevan años diseñando herramientas de esta naturaleza (Vlek, 2013), si bien un sistema de evaluación de riesgos especialmente destacable por su carácter pionero, integral y científicamente sólido es el de los Países Bajos.

Ante una realidad en la que las amenazas estaban cada vez más globalizadas e interconectadas, y por tanto requerían de un enfoque holístico que rompiera los compartimentos estancos desde los que la seguridad nacional había afrontado tradicionalmente estos desafíos, en 2007 el gobierno neerlandés estableció una Estrategia de Seguridad Nacional (ESN) estructurada en tres fases: una primera en la que se elabora un Perfil Nacional de Riesgos (PNR)[67] para identificar y analizar comparativamente todas aquellas amenazas relevantes, una segunda en la que se evalúan las capacidades necesarias para evitar tales amenazas o al menos limitar sus consecuencias, y una tercera en la que se ejecutan las medidas consideradas para reforzar dichas capacidades (*The National Network of Safety and Security Analysts*, 2016). La primera de estas fases, en la cual se centra el presente caso de estudio, es llevada a cabo por la Red Nacional de Analistas en Seguridad (RNAS) bajo las directrices marcadas por el Comité de Seguridad Nacional (CSN), estructura interministerial consejera del gobierno. Para favorecer la requerida visión global y multidisciplinar, la RNAS está formada por seis organizaciones de distinta naturaleza como son el Centro de Documentación e Investigación del Ministerio de Seguridad y Justicia, el Instituto Nacional para la Salud Pública y el Medio Ambiente, el Servicio General de Inteligencia y Seguridad, la Organización Neerlandesa para la Investigación Científica Aplicada, el Instituto Neerlandés de Relaciones Internacionales *Cligendael*, y el

67 En los primeros años de la ESN a esta metodología se la denominaba Evaluación Nacional de Riesgos (ENR), cambiando de denominación en el marco de la reforma realizada en 2014 y pasando a ser su elaboración cada cuatro años.

Instituto Internacional de Estudios Sociales de la Universidad Erasmus de Rotterdam; los cuales son asimismo apoyados ocasionalmente por expertos procedentes de otros centros de estudios, consultoras, compañías de seguros, servicios de emergencias, etc.

Junto a la ESN, un elemento adicional que revela la importancia dada por los Países Bajos a la evaluación de los riesgos a los que se enfrenta el país es la obligación por parte de las 25 denominadas "regiones de seguridad" de diseñar Perfiles Regionales de Riesgos, en cuya elaboración se tiene en cuenta lo establecido en el PNR para garantizar cierta coherencia en los riesgos percibidos a lo largo de los niveles nacional, provincial y local.

Metodología

Tal y como se adelantaba en el anterior epígrafe, la finalidad del PNR es doble. Por un lado, seleccionar todas las amenazas que ponen en serio peligro los intereses nacionales de Países Bajos, y clasificarlas en una de las ocho categorías de riesgo (*The National Network of Safety and Security Analysts*, 2016):

- Desastres naturales
- Salud pública y medioambiente
- Accidentes
- Infraestructuras críticas
- Ciberamenazas
- Subversión, extremismos y terrorismos
- Geopolítica
- Economía y finanzas

Por otra parte, el segundo y más importante propósito del PNR es realizar un análisis comparativo del riesgo que supone cada una de ellas para en las siguientes etapas de la ESN facilitar la acción política. Y dado que tanto la selección de amenazas como el cálculo de los riesgos de éstas giran en torno a los impactos y probabilidades, la obtención de ambas variables es el principal objetivo de los grupos de trabajo de la RNAS. Para ello cada grupo, encabezado por su correspondiente *Project Leader*, lleva a cabo la evaluación de una amenaza específica mediante la descripción de un abanico de escenarios que recogen la variabilidad que pueden

experimentar elementos tales como el contexto, los detonantes, las consecuencias o las vulnerabilidades.

El criterio del PNR para la selección de amenazas se basa en el impacto que puedan llegar a alcanzar. A este respecto, una vez establecidas cinco intensidades que van de "Limitado" a "Catastrófico", pasando por los intermedios "Significativo", "Serio" y "Muy serio"; tan sólo se recogen aquellas amenazas que incluyan como mínimo un escenario con un impacto potencial "Serio" o superior en al menos uno de los criterios incluidos en la metodología.

Tabla 1. Intereses de Seguridad Nacional y criterios de impacto incluidos en el PNR

Intereses de Seguridad Nacional	Criterios de impacto
Seguridad territorial	1.1 Pérdida de soberanía territorial y digital
	1.2 Perjuicio de la posición internacional
Seguridad personal	2.1 Fallecidos
	2.2 Heridos graves
	2.3 Ausencia de necesidades básicas que conlleven sufrimiento físico en la población
Seguridad económica	3.1 Costes
	3.2 Perjuicio de la vitalidad económica del país
Seguridad ecológica	4.1 Perjuicio en la naturaleza y el medioambiente
Estabilidad política y social	5.1 Disrupción en la vida diaria de los ciudadanos
	5.2 Violación del sistema democrático constitucional
	5.3 Impacto psicológico en la sociedad

Fuente: *National Risk Profile 2016*

En cuanto a la medición de los daños caben destacar dos cuestiones. Por un lado, la escala por la que se mide el impacto de los escenarios según cada uno de los criterios no es lineal sino logarítmica, con el objetivo de que cada escalón de intensidad suponga un perjuicio entre 5 (en términos temporales) y 10 (en términos cuantitativos) veces mayor que el anterior. Matemáticamente, esto se

plasma aplicando una transformación exponencial de base 3, de tal manera que al impacto "Limitado" le corresponde un valor de 1/81, al "Significativo" un 3/81, al "Serio" un 9/81, al "Muy serio" un 27/81, y al "Catastrófico" un 81/81. El segundo aspecto a subrayar es que en el PNR de 2016 el peso otorgado a cada uno de los criterios en el cálculo del impacto total del escenario es siempre el mismo (1/11), si bien en ediciones anteriores también se incluían valoraciones de impacto total en base a sensibilidades políticas no neutras que implicaban una ponderación desigual de los criterios, y de la que por tanto surgían evaluaciones de riesgo alternativas. Dichas sensibilidades eran la "individualista", visión liberal que prioriza el buen funcionamiento de la economía de mercado; la "igualitaria", visión solidaria según la cual lo primordial es el bienestar general; la fatalista, cuyo principal objetivo es el mantenimiento de la seguridad y los valores sociales; y la "jerárquica", que concede especial importancia a lo acaecido en el entorno más cercano (*Method Group for National Safety and Security*, 2014).

Junto al impacto, el segundo elemento que determina en el PNR el riesgo de un escenario es la probabilidad de que pueda consumarse a lo largo de los siguientes cinco años. Y al igual que sucede con los impactos, éstas también se dividen en cinco categorías dentro de una escala no lineal: "Muy improbable" cuando la probabilidad es inferior a 0,05%, "Improbable" entre el 0,05% y el 0,5%, "Algo probable" entre el 0,5% y el 5%, "Probable" entre el 5% y el 50%, y "Muy probable" cuando la probabilidad es superior al 50%. En la medida de lo posible, los grupos de trabajo intentan calibrar las probabilidades e impactos a partir de criterios objetivos como bases de datos de desastres similares o modelos probabilísticos, sin bien ante la ausencia de herramientas de esta naturaleza en bastantes ocasiones el cálculo tiene que basarse principalmente en la opinión subjetiva de los expertos, donde una detallada y consistente descripción del escenario (lo que incluye discernir si la amenaza es de naturaleza maliciosa o no intencionada, así como una minuciosa evaluación de las vulnerabilidades) cobra una especial importancia.

Una vez calculados los impactos y las probabilidades que determinan el riesgo de cada escenario, estos se plasman en un diagrama que permite su comparativa, principal objetivo del PNR tal y como se indicaba al comienzo de este epígrafe. A este respecto cabe remarcar que, para clarificar el análisis, se incluyen como máximo dos escenarios representativos de cada amenaza seleccionada, los cuales habitualmente (aunque no siempre) ilustran el peor escenario y el estándar. En la Figura 1 puede observarse el diagrama correspondiente al PNR de 2016, cuya lectura es la habitual de los mapas de riesgo: en el cuadrante superior izquierdo se sitúan aquellas amenazas

muy improbables pero con consecuencias extremadamente disruptivas en caso de materializarse, mientras que en el extremo opuesto se localizan las amenazas más probables cuyos efectos serían por el contrario más limitados. Por lo general, en estos diagramas el cuadrante superior derecho está vacío pues no suelen existir amenazas cuyos riesgos sean tan elevados.

Por último, el PNR también incluye un análisis de las principales tendencias ecológicas, demográficas, geopolíticas, económicas y tecnológicas, acaecidas generalmente a nivel global, con potencial para empeorar la evolución de cada una de las amenazas en el medio y largo plazo (Tabla 2). Entre las mismas figuran el cambio climático, los crecientes flujos migratorios, las guerras híbridas, el aumento de las barreras comerciales, o el perfeccionamiento de la Inteligencia Artificial.

Tabla 2. Tendencias con potencial para incrementar las amenazas a largo plazo

Tendencias ecológicas	Cambio climático	Aumenta los desastres naturales, el deterioro de la salud ciudadana, y la amenaza sobre infraestructuras críticas.
	Pérdida de biodiversidad	Supone una disrupción del ciclo ecológico.
	Presión medioambiental	Implica contaminación y sobrexplotación de los recursos naturales.
Tendencias sociales y demográficas	Envejecimiento	Conlleva mayor gasto sanitario y menor población activa.
	Polarización	Deteriora la cohesión social y la confianza en las instituciones.
	Migraciones	Si superan la capacidad de absorción pueden desencadenar sentimientos xenófobos.
	Urbanización	Áreas muy densamente pobladas son más vulnerables a los desastres naturales.

Tendencias geopolíticas	Desplazamiento del poder mundial	Ascenso de países revisionistas del orden liberal internacional.
	Incremento de las tensiones entre potencias	Mayor dificultad de una gobernanza mundial mediante instituciones multilaterales.
	Inestabilidades regionales	Puede provocar flujos migratorios masivos y amenazas terroristas.
	Vínculos entre seguridad nacional e internacional	Impacto creciente de los conflictos externos en el territorio nacional.
	Amenazas híbridas	Pueden provocar desconfianza hacia las instituciones democráticas y deteriorar la cohesión social.
Tendencias económicas	Desigualdad	Daña la cohesión social y el crecimiento macroeconómico del país.
	Automatización y robotización	Aumento del desempleo entre los trabajadores menos cualificados con el consiguiente deterioro social.
	Bajos tipos de interés	Aumentan las tensiones financieras.
	Desplazamiento del poder económico hacia áreas emergentes	El menor peso económico de Países Bajos disminuye la influencia internacional del país y debilita su Estado del Bienestar.
	Sanciones económicas	Aunque no le afecten directamente, suponen un grave perjuicio a economías especialmente abiertas como la neerlandesa.
	Proteccionismo comercial	Ídem.
	Globalización	Implica procesos de concentración productiva, y por tanto mayor dependencia en unos pocos actores.

Tendencias tecnológicas	Control de las innovaciones	La amenaza que supone sobre sectores tradicionales puede provocar conflictos sociales.
	Tecnologías de uso dual	Suponen una amenaza en ausencia de regulación adecuada.
	Accesibilidad	La facilidad de acceso a ciertas tecnologías supone una amenaza.
	Internet de las Cosas	El incremento de la conectividad y dependencia entre sistemas aumenta la vulnerabilidad ante ciberataques.
	Big Data	Permite recopilar cantidades ingentes de información que puede usarse con fines perniciosos.
	Inteligencia Artificial	Otorgar una mayor responsabilidad a estos sistemas dificulta la respuesta ante pérdidas de control.

Fuente: National Risk Profile 2016

Caso de análisis de amenaza: inundaciones

No debe sorprender que, para un país con gran parte de su territorio por debajo del nivel del mar, las inundaciones sean una de las amenazas más temidas. De hecho, uno de los desastres todavía más recordados de su historia reciente es *De Watersnood*, inundación marina que anegó el sudoeste del país la noche del 31 de enero de 1953 en la que fallecieron 1.835 personas y alrededor de 70.000 tuvieron que ser evacuadas. A raíz de este suceso los Países Bajos pusieron en marcha el Plan Delta, consistente en la construcción a lo largo de más de cuatro décadas de 13 diques y una esclusa para evitar que un episodio de esta naturaleza pudiera volver a repetirse. Adicionalmente, para reducir el riesgo en la actualidad están activos el Programa Delta 2019 y el Plan Nacional Hídrico 2016-2021. [68]

68 Más información de ambos programas en https://english.deltacommissaris.nl/delta-programme/delta-programme-2019 y https://www.government.nl/documents/policy-notes/2015/12/14/national-water-plan-2016-2021

Con respecto a esta amenaza el último PNR incluye dos escenarios: uno estándar que hace referencia a una inundación fluvial, y otro catastrófico en el que una inundación marina originada por un huracán o tormenta extrema anegaría toda la conurbación de Randstad, en la cual vive en torno al 40% de la población neerlandesa al incluir ciudades de la importancia de Ámsterdam, Roterdam, La Haya y Utrecht. En la Tabla 3 puede observarse un resumen de la evaluación efectuada por el grupo de trabajo correspondiente a un escenario de estas características.

La suma de todos los impactos determina que el impacto global de este escenario sería catastrófico:

$$[1+(3/81)+1+1+1+1+(27/81)+1+1+(3/81)+(9/81)]/11= 0,683$$

En el diagrama puede observarse como este es el único escenario incluido en el PNR 2016 cuyo impacto alcanza esta categoría, siendo la inundación fluvial menos dañina (aunque más probable). No obstante, el riesgo de una pandemia severa, en función del criterio político aplicado,[69] podría ser superior en la medida que conjuga un impacto muy serio (más de 14.000 fallecidos y alrededor de 50.000 hospitalizados) con una probabilidad del 20% en los siguientes cinco años. Por otra parte, a ninguno de los escenarios incluidos se le otorga una probabilidad de acaecimiento superior al 50%.

69 A la hora de priorizar las amenazas a las que hacer frente se pueden adoptar tres posturas (*Method Group for National Safety and Security*, 2014). Según la primera, el impacto y la probabilidad tienen la misma importancia para calibrar el riesgo; en base a la segunda, los escenarios prioritarios son aquellos que conllevan un mayor impacto, siendo la probabilidad un factor secundario; mientras que la tercera visión prioriza aquellos escenarios en los que es más factible reducir significativamente la probabilidad de acaecimiento, los cuales suelen ser aquellos con una mayor probabilidad de partida.

Tabla 3. Análisis de escenario: inundación marina

Evaluación de la probabilidad							
		Muy improbable	**Improbable**	**Algo probable**	**Probable**	**Muy probable**	**Explicación**
Probabilidad de que el escenario se materialice en los siguientes cinco años.	X						La probabilidad de que se produzca un huracán o tormenta extrema y se produzca un fallo múltiple de los diques es menor a 1/100.000 al año.

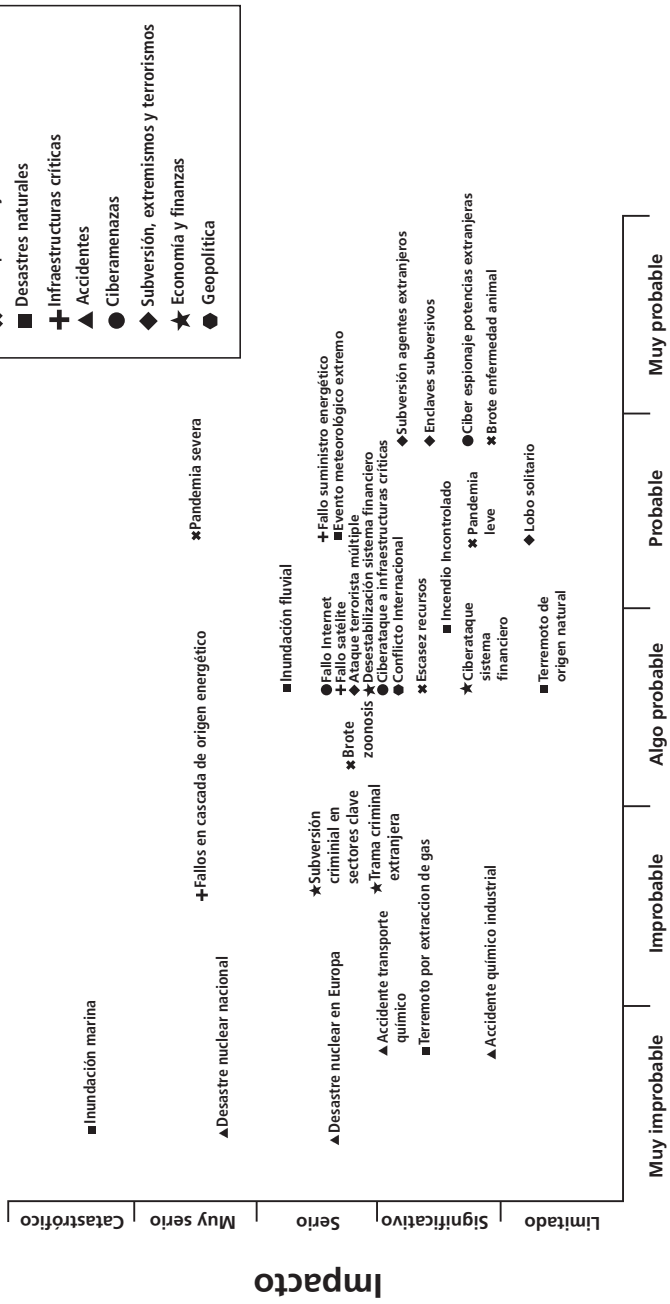
Evaluación del impacto							
Intereses de Seguridad Nacional	**Criterios de impacto**	**Limitado**	**Significativo**	**Serio**	**Muy serio**	**Catastrófico**	**Explicación**
Territorial	Soberanía territorial y digital					**X**	4.340 km2 de exclusión durante un largo periodo de tiempo.
	Posición internacional		**X**				La actividad de las embajadas y las visitas internacionales se verían obstaculizadas.
Personal	Fallecidos					**X**	Más de 10.000 fallecidos.
	Heridos graves					**X**	
	Necesidades básicas					**X**	Más de un millón de personas afectadas y fallos en el suministro de agua, que podría llevar hasta un año restaurar.

Evaluación del impacto							
Intereses de Seguridad Nacional	**Criterios de impacto**	**Limitado**	**Significativo**	**Serio**	**Muy serio**	**Catastrófico**	**Explicación**
Territorial	Soberanía territorial y digital					X	4.340 km2 de exclusión durante un largo periodo de tiempo.
	Posición internacional		X				La actividad de las embajadas y las visitas internacionales se verían obstaculizadas.
Personal	Fallecidos					X	Más de 10.000 fallecidos.
	Heridos graves					X	
	Necesidades básicas					X	Más de un millón de personas afectadas y fallos en el suministro de agua, que podría llevar hasta un año restaurar.
Económico	Costes					X	Daños estimados en 121.000 millones de euros.
	Economía del país				X		Incremento de la deuda nacional por el daño sufrido por el comercio internacional, la interrupción de la producción, el coste de las medidas implementadas, etc. Además habría consecuencias sobre el empleo, tanto a corto como a medio plazo.
Ecológico	Naturaleza y medioambiente					X	Varias reservas naturales se verían afectadas por un largo periodo de tiempo.

Sociopolítico	Vida diaria de los ciudadanos					X	Disrupción en la vida diaria de la gente afectada durante un largo periodo de tiempo.
	Sistema democrático		X				El orden público y la seguridad se verían especialmente afectadas.
	Impacto psicológico			X			Incremento de la ansiedad en la sociedad.

Fuente: *National Risk Profile 2016*

Figura 1. Diagrama de riesgo



Fuente: *National Risk Profile 2016*

Complementando esta evaluación, el grupo de trabajo señala dos tendencias que incrementan a largo plazo la vulnerabilidad ante la amenaza de una inundación de origen marino: un cambio climático que derrite los polos incrementando paulatinamente el nivel del mar, y la creciente densidad de población en ciertas áreas del país entre las que se incluye la conurbación de Randstad.

Conclusión

La elaboración del Perfil Nacional de Riesgos de Países Bajos es un procedimiento transparente que, mediante una metodología integral y científicamente robusta,[70] permite a expertos de distintas disciplinas analizar las principales amenazas a las que tiene que hacer frente el país en un horizonte temporal de cinco años, así como las tendencias generales que pueden amplificar el riesgo a más largo plazo. El principal resultado de este proceso es la elaboración de un mapa de riesgos, cuya consistencia matemática supera al de las tradicionales matrices,[71] a partir del cual los poderes públicos evalúan las capacidades del Estado y toman las medidas oportunas para mejorarlas, acciones que completan la Estrategia de Seguridad Nacional.

Bibliografía

Cox L. A. (2008) What's Wrong with Risk Matrices? *Risk Analysis*, Vol. 28, No. 2:497-512.

Method Group for National Safety and Security (2014) *Working with scenarios, risk assessment and capabilities in the National Safety and Security Strategy of the Netherlands.* Ministerio de Seguridad y Justicia, La Haya (Países Bajos).

Pruyt, E., Wijnmalen, D., y Marc Bökkerink (2013) What Can We Learn from the Evaluation of the Dutch National Risk Assessment? *Risk Analysis*, Vol. 33, No. 8: 1385-1388.

The National Network of Safety and Security Analysts (2016) *National Risk Profile 2016: An All Hazard overview of potential disasters and threats in the Netherlands.* Instituto Nacional para la Salud Pública y el Medio Ambiente, Bilthoven (Países Bajos).

Vlek, C. (2013) How Solid Is the Dutch (and the British) National Risk Assessment? Overview and Decision-Theoretic Evaluation. *Risk Analysis*, Vol. 33, No. 6: 948-9

70 Para un análisis crítico de la primera versión de la ENR (antecedente del PNR) ver Vlek (2013), rebatido a su vez por Pruyt *et.al.* (2013).

71 Para un análisis comparativo entre matrices de riesgo y mapas de riesgo ver Cox (2008).

José Pablo Martínez es ayudante de investigación del Real Instituto Elcano. Licenciado en Economía por la Universidad de Murcia y Máster en Economía Internacional por la Universidad Autónoma de Madrid, ha sido becado para realizar diversos cursos sobre Relaciones Internacionales, Economía y Empleo en la Escuela Diplomática, en la Universidad Complutense de Madrid y en la Universidad Internacional Menéndez Pelayo. Ha trabajado en el Área de Estudios y Análisis del Consejo Económico y Social de España (CES), e igualmente ha participado en proyectos de investigación económico-social en la Universidad de Murcia y en la Universidad Autónoma de Madrid.


mirror_mod.use_y = True
mirror_mod.use_z = False
elif operation == "MIRROR_Z":
mirror_mod.use_x = False
mirror_mod.use_y = False
mirror_mod.use_z = True
#selection at the end -add bac
mirror_ob.select= 1
modifier_ob.select=1
bpy.context.scene.objects.active =
print("Selected" + str(modifier_ob