

EL SESGO TECNOLÓGICO EN EL ANÁLISIS DE INTELIGENCIA DE FUENTES ABIERTAS. NOTAS DE TRABAJO.

Pablo las Heras, EULEN Seguridad (plasheras.seguridad@eulen.com)

María L. Díaz Monzón, EULEN Seguridad (mldiazm.seguridad@eulen.com)

Carlos Blanco Torres, EULEN Seguridad (cblancot.seguridad@eulen.com)

NOTA BIOGRÁFICA DE LOS AUTORES

Pablo las Heras

Analista de la Unidad de Inteligencia de EULEN Seguridad. Graduado en Ciencias Políticas por la Universidad Complutense de Madrid con especialización en el análisis de conflictos armados en Europa, Latinoamérica y Oriente Medio. Experto en Gestión de Crisis y Resolución de Conflictos (Curso REACT). Experiencia profesional en el sector de la inteligencia, la mediación en conflictos y los Organismos Internacionales.

María L. Díaz Monzón

Analista de la Unidad de Inteligencia de EULEN Seguridad. Licenciada en Ciencias Políticas y de la Administración por la Universidad Complutense de Madrid, con especialización en Relaciones Internacionales. Máster en Relaciones Internacionales y Comercio Internacional (INFOREM-Colegio de Politólogos y Sociólogos) y en Analista de Inteligencia (URJC-UC3M). Experiencia profesional en el sector de la inteligencia y la seguridad.

Carlos Blanco Torres

Jefe de la Unidad de Inteligencia de EULEN Seguridad. Licenciado en Ciencias Políticas y de la Administración (UCM), máster en Relaciones Internacionales (UCM) y en Analista de Inteligencia (URJC/UC3M). Doctorando en CC. Políticas y RR.II. de la Facultad de CC. Políticas y Sociología de la UCM. Tesis: Multipolaridad contrahegemónica. El ascenso de China y la inserción internacional de América del Sur. Director: José Antonio Sanahuja Perales.

RESUMEN

Ante la creciente automatización de las labores de obtención de la información dentro del proceso de inteligencia, se hace cada vez más necesario integrar su importancia en las revisiones teóricas y prácticas que se realizan sobre el análisis de inteligencia. En este documento proponemos partir de los sesgos cognitivos que afectan al analista durante la realización de sus tareas para ofrecer una aproximación teórica a la forma en que estos se trasladan a las herramientas de uso en el análisis de inteligencia de fuentes abiertas (OSINT), utilizando el concepto de “sesgos tecnológicos”. Estos sesgos aparecen durante el proceso de desarrollo de las herramientas, y tienen un impacto similar sobre el análisis que el que tienen los sesgos cognitivos. Se presenta una serie de ejemplos prácticos con los que se pone de manifiesto que distintas herramientas devuelven resultados disímiles sobre una misma búsqueda.

PALABRAS CLAVE

Análisis de inteligencia, fuentes abiertas, proceso de inteligencia, sesgo cognitivo, sesgo tecnológico.

“El análisis de inteligencia es el procedimiento que engloba diferentes procesos o fases de actuación en el tratamiento de información recibida de observaciones planificadas, de acuerdo con las necesidades de información establecidas en cada momento; se basa en el método de análisis científico hipotético deductivo, con algunos procesos previos de carácter inductivo y deductivo con objeto de que el analista conozca al detalle la situación real, antes de aplicar el método científico y facilite al decisor su labor.”

Diccionario LID de Inteligencia y Seguridad, Pág. 37

INTRODUCCIÓN

El análisis de inteligencia vive actualmente un pronunciado proceso de automatización que afecta, de manera singular, a los procesos de recolección y procesamiento de la información. Esta automatización, provocada por el uso de herramientas de apoyo¹, es especialmente pronunciada en el análisis de inteligencia basado en la recolección de información localizada en fuentes abiertas OSINT (*Open Source INTeligence*), definida por el diccionario LID de Inteligencia y Seguridad (2013: pág. 201) como:

“[...] información elaborada por analistas mediante la utilización de fuentes abiertas para una audiencia concreta, con la finalidad de dar respuesta a un requerimiento específico de información”.

Esta automatización ha dado, a su vez, un protagonismo creciente –paralelo al del analista- a las herramientas de recolección de información, cuya operación o manejo ha pasado a formar parte sustancial del día a día de las unidades de inteligencia públicas y privadas. De este modo, pretendemos repasar el concepto ya estudiado de sesgo cognitivo y realizar un análisis preliminar sobre la afectación –similar a un sesgo sobre el analista- que atañe a la otra columna en la que, hoy por hoy, se asienta el análisis de inteligencia: las herramientas.

Este análisis preliminar se realizará primero desde una aproximación teórica al concepto de sesgo tecnológico, donde se definirá a qué nos referimos con ello y su impacto sobre el análisis; pasando después a la visión en la práctica de cómo estos sesgos afectan a las herramientas de recolección y análisis de información en fuentes abiertas, a través de tres búsquedas típicas de un análisis de inteligencia OSINT.

¹ Además de 'herramienta', aceptamos como sinónimos en este trabajo las palabras 'programa', 'software' y 'producto'.

La reflexión inicial que precipita este estudio preliminar consiste en que mientras que los ‘fallos’ soportados por los analistas son conocidos –y asumidos- en el ámbito de la inteligencia, los ‘fallos’ en las herramientas suelen ser achacados a cuestiones puramente técnicas, no siendo común su estudio desde un punto de vista más ‘humano’. La hipótesis es, pues, que igual que los analistas están influidos por su contexto y condiciones, lo que dirige su análisis impidiendo que sea completamente objetivo, las herramientas también están influidas por el contexto donde se crean, debiendo esto ser tenido en cuenta a la hora de procesar la información reportada por estos soportes.

Las fallas del analista: sesgos cognitivos

Existe una amplia bibliografía sobre las afectaciones intrínsecas que sufre el analista de inteligencia –en su condición de humano- a la hora de proceder al análisis. El análisis es un proceso que requiere por parte de quien lo aplica la discriminación, manipulación e integración de la información a través de un proceso mental que, precisamente por ser personal, está influido por sesgos y tendencias propias (Martín, 2016).

Ahondando en esta definición, el diccionario LID de Inteligencia y Seguridad define el sesgo cognitivo (2013: pág. 236) como una:

“Inclinación o predisposición a favor o en contra de algo, que genera una desviación en el procesamiento de la información y puede dar pie a una interpretación equivocada o a un juicio inexacto”.

La influencia de los sesgos cognitivos y, en general, aquellas fallas en el proceso mental del analista que afectan negativamente al análisis de inteligencia, ha sido ampliamente estudiada, presentando una literatura variada tanto en efectos como en causas, así como en posibles soluciones. Los sesgos pueden clasificarse en cuatro grandes categorías, basada en los estudios de Richards J. Heuer (Jordán, 2011):

- Aquellos que afectan a la evaluación de evidencias.
- Aquellos que inciden sobre el discernimiento de las relaciones causa/efecto.
- Aquellos que afectan al cálculo de estadísticas y probabilidades.
- Aquellos que afectan a la evaluación a posteriori de otros análisis.

Al no ser el objeto del presente trabajo una revisión al concepto de sesgo cognitivo, se aportará una breve reseña sobre los más habituales (Jordán, 2014; Manzano, 2010; Sáiz, 2010)²:

- Sobredimensión de la confianza en la información recibida de forma interpersonal o basada en experiencias propias en relación con aquella recibida a través de otros medios independientemente de su valor empírico.
- Exceso de confianza en las técnicas cuantitativas (expresadas a través de valores numéricos).
- Sobredimensión de los hechos ya conocidos en detrimento de aquello no conocido, independientemente de lo determinante que sea.
- Propensión a la asunción de conclusiones generales basadas en piezas de información reducidas.
- Asunción de la totalidad de una información como veraz tras la comprobación sólo de una parte de ella.
- Sobredimensión de las “primeras impresiones”, especialmente si están en consonancia con las hipótesis iniciales.
- Atribución de coherencia y causalidad a hechos encadenados independientemente de la relación de unos con otros.
- Asunción de una causa común para hechos sincrónicos.
- Presunción de motivaciones comunes hacia actores que comparten categorización.
- Tendencia a igualar magnitudes entre las variables causa/efecto.
- Sobredimensión de la voluntad del individuo sobre factores externos.

Todas estas circunstancias –intrínsecas a la condición humana- llevan al analista a valorar más aquellas informaciones que apunten a lo que espera encontrar, aquellas informaciones que certifiquen su impresión inicial o su juicio previo (Martín, 2016). Este hecho afecta de forma inevitable al resultado final del análisis y, con ello, a las fases siguientes del proceso de inteligencia³.

Conocida esta situación y asumida como inevitable, se han establecido numerosos protocolos destinados a corregir –o reducir- el impacto de los sesgos del analista sobre el producto final. Sin embargo, como se ha apuntado y se desarrollará en los siguientes puntos, las herramientas,

² Recogidos por Diego Martín (2016).

³ Utilizamos 'proceso de inteligencia' como concepto superador del más tradicional 'ciclo de inteligencia'.

en tanto que tecnológicas y no humanas, gozan de una presunción de ‘objetividad’ que, de no cuestionarse, puede influir en el análisis de manera similar a como influyen los sesgos cognitivos del analista. Igual que a la hora de recibir un documento de inteligencia ponderamos el posible peso que estos sesgos tendrían en los resultados, es relevante ponderar de igual manera el posible peso que tendrán los sesgos tecnológicos que residen en la/las herramienta/s utilizadas durante el proceso de inteligencia.

En definitiva, como más adelante quedará reflejado, las herramientas presentan limitaciones no sólo técnicas, sino también ‘humanas’ en tanto en cuanto su desarrollo está influido por el contexto que lo rodea, pudiendo incluso entenderse que el sesgo cognitivo que tiene el desarrollador de la herramienta queda, de alguna manera, plasmado en su desarrollo y posterior puesta en marcha de la misma.

SESGO TECNOLÓGICO: UNA APROXIMACIÓN

Como se ha señalado, la intención de este trabajo consiste en realizar una aproximación al concepto de sesgo tecnológico, entendido como una barrera que dificulta las tareas de obtención y análisis de información, centrales en los trabajos de inteligencia. Se dejan de lado por tanto otras desviaciones asociadas a la tecnología, como la que se ocupa del desigual acceso al proceso científico, o la que pone en relación el par habilidades/cambio tecnológico. En cambio, se tendrán muy en cuenta los efectos –o la influencia- de los sesgos cognitivos humanos.

Se pretende diferenciar el concepto de referencia de otros, rechazando la equiparación de sesgo tecnológico con error tecnológico. Es decir, el sesgo es una limitación inadvertida en la base de diseño de la tecnología y no un fallo de arquitectura o una limitación funcional.

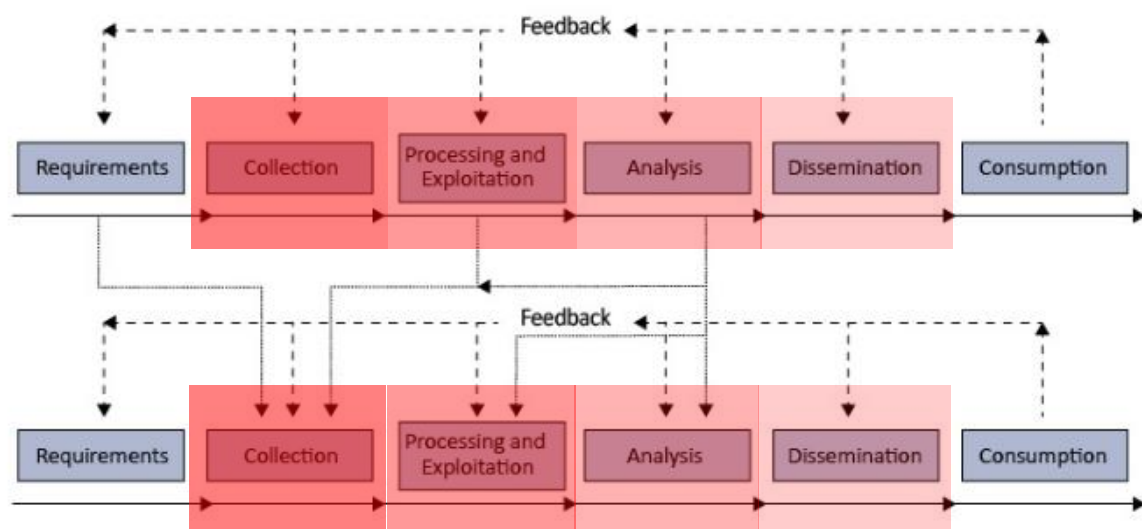
Este obstáculo técnico se suma a los sesgos cognitivos que presentan los analistas, que pueden mitigarse a partir del reconocimiento de los mismos. Así, entender los sesgos de las herramientas en uso permitirá a los trabajadores del ámbito de inteligencia conocer las debilidades del producto elaborado y establecer métodos de corrección. El impacto de este problema puede limitarse desde fases tempranas de un proyecto/servicio, simplemente mediante el estudio e investigación de las capacidades y las funcionalidades de las muchas herramientas que existen en la actualidad, o mediante el uso de múltiples programas de este tipo de modo que unos complementen a otros. En todo caso, esto no hace sino reafirmar la

importancia decisiva de la planificación –objetivos y recursos- en la realización de un trabajo de inteligencia sólido y fiable.

La tecnología es un elemento de gran relevancia –hoy y siempre- en el conjunto de acciones que conforman el proceso de inteligencia, que es el conjunto de tareas y fases de las que se compone el trabajo de las personas implicadas en los esfuerzos de inteligencia de una organización y cuya complejidad ha obligado a abandonar, poco a poco, el concepto tradicional de ciclo de inteligencia (Jordán, 2016). Defendemos aquí que las herramientas tecnológicas que se utilizan de manera rutinaria en inteligencia no son perfectas, presentan limitaciones de diverso tipo y, por tanto, ello debe tenerse en cuenta en el momento de utilizar los resultados que producen. Según el uso que se dé a la tecnología, los sesgos que esta presenta pueden resumirse en dos grandes grupos:

- Obtención y procesado de información: se refiere al hecho de que diversas herramientas generan diversos resultados ante búsquedas iguales. Además, cada una de ellas presentará también grados distintos de potencia a la hora de procesar grandes volúmenes de información.
- Análisis de información: una vez delimitado el cuerpo de información bajo análisis, diferentes programas establecerán diferentes tipos de conexiones entre los datos. Es consecuencia del sesgo que afecta a la tarea de obtención.

De este modo, una primera categoría de interés para definir qué es un sesgo tecnológico es la de reconocer a qué proceso o tarea afecta.



Esquema del proceso de inteligencia (Lowenthal, 2012) y tareas afectadas por los sesgos tecnológicos.

Estas situaciones están a su vez determinadas por la propia arquitectura del software de uso, que es evidentemente un producto de actividad humana. De esta manera, identificamos tres tipos de sesgos de interés, que proceden de diversos actores involucrados en la creación de una herramienta de obtención y análisis: el sesgo trasladado programador-programa, el trasladado usuario-programa vía retroalimentación y el producido por el aprendizaje sesgado del software, dependiendo de las fuentes que se utilicen para tal fin (Rosenblat, 2016).

- Sesgo trasladado programador-programa: consiste en la transferencia de los sesgos cognitivos humanos durante la fase de programación de una solución informática dada.
- Sesgo trasladado usuario-programa: aparece según la forma en que una herramienta evoluciona y es modificada a partir de los requerimientos, comentarios y sugerencias de los usuarios.
- Aprendizaje sesgado: muy relacionado con los procesos de *machine learning*, se refiere al hecho de que las herramientas utilizan una serie de fuentes determinadas para mejorar su entendimiento de situaciones complejas. Por ello, una mala selección de fuentes –que presentarán sus propios sesgos- vicia de partida la capacidad de aprender de un programa.

Por tanto, como segunda característica de relevancia, consideramos que un sesgo tecnológico es la traducción informática de un sesgo humano de partida. Los modelos mentales que presentamos los humanos para comprender cómo funcionan determinados aspectos de la realidad –relaciones sociales, vínculos económico-políticos, causa/efecto de eventos, etc.- pueden ser poco efectivos o, en el peor de los casos, directamente erróneos. Del mismo modo, si una herramienta ha de extraer la información de la que se nutre de fuentes que presentan problemas, como por ejemplo conjuntos de datos no representativos en fuentes relacionadas con datos sobre criminalidad, el resultado presentado ante una consulta será defectuoso (Burns, 2017).

Otro elemento de interés al aproximarnos al concepto de sesgo tecnológico aparece si asumimos que las limitaciones básicas que presentan los programas de uso habitual en el análisis de inteligencia no tienen la misma importancia. De esta manera, el impacto de un sesgo como los que aquí se delinean es superior según tenga capacidad de afectar en mayor o menor medida al resultado final del análisis. Así, una forma de observar este aspecto podría ser relacionar el impacto de un análisis errado con las limitaciones tecnológicas sufridas durante su elaboración, en términos económicos (en el caso de la inteligencia económica o de

negocio), de vidas humanas (en el caso de la inteligencia de seguridad), o lo que proceda en cada caso. No obstante, este extremo es difícil de medir, pues podría alegarse que un mal análisis sin repercusión en el mundo real, por muy desviado que se encontrase respecto al resultado final, no ofrece capacidad de ser medido.

Por último, diremos que un aspecto importante del sesgo tecnológico obedece al carácter no inocente o neutral del propio progreso/desarrollo científico. La ciencia, en cuanto a su camino de evolución, no es objetiva (Roth, 2009). Los avances obtenidos por la ciencia –se produzcan estos en el campo en el que se produzcan- proceden de los impulsos que proporcionan determinados actores a través de, por ejemplo, la financiación de proyectos o la definición de programas de investigación, que obligan a seguir un camino dado o, cuanto menos, establecen altos costes de oportunidad para aquellos investigadores que se sitúan en los márgenes de las corrientes principales de la investigación científica. En el contexto que nos ocupa, esta dimensión es pertinente cuando nos preguntamos el porqué de las funcionalidades de determinadas herramientas (tipo de relaciones que establecen entre objetos) o de incluso de su propia naturaleza (idioma principal de captación de información, por ejemplo).

Creemos que definir qué es un sesgo tecnológico y entender cómo actúa en un caso dado es de relevancia en el campo del análisis de inteligencia, por varios motivos:

- La tecnología cada vez está más presente en el trabajo de recolectores de información y analistas. Esto es un hecho positivo, pues permite obtener el músculo necesario para detectar, obtener, procesar y analizar grandes cantidades de información. Sin embargo, es importante entender que una herramienta concreta, obedeciendo a una consulta determinada, ofrecerá resultados limitados por sus capacidades y arquitectura interna.
- La tecnología ayuda, o puede ayudar, a establecer relaciones ocultas, no observables en la superficie, entre piezas de información aparentemente inconexas. Sin embargo, interesarse por conocer cómo y por qué se establecen determinadas uniones puede mejorar la práctica analítica de las personas.
- La tecnología no es infalible o perfecta. El uso crítico de la misma en procesos y tareas en el ámbito de la inteligencia debería ser un requisito básico, al igual que se establecen sistemas de revisión y validación centrados en las técnicas implementadas y las valoraciones obtenidas por los analistas.
- La tecnología nace y evoluciona a un ritmo muy elevado, pero ello no ocurre de manera casual. Generalmente, los productos destinados a la obtención y análisis de

información son desarrollados para cubrir ciertas necesidades –reales o creadas, ese es otro debate- y por tanto se orientan hacia prácticas tradicionales para obtener resultados tradicionales.

En resumen: características y definición preliminar de sesgo tecnológico

Como resumen, diremos que las características principales de los sesgos tecnológicos en el contexto del análisis de inteligencia son:

- Efecto: puede afectar a las tareas, realizadas por el propio software en uso, de obtención y procesado, de análisis, o ambas.
- Origen: es la traducción de un sesgo humano de origen.
- Impacto: alto o bajo, según afecte a uno o varios componentes de la arquitectura del programa y la relevancia de estos en el contexto de los resultados ofrecidos por una herramienta.
- Agenda: el progreso científico y tecnológico no es neutral, sino que, de manera consciente o inconsciente, avanza según los intereses de ciertos grupos de interés.

Una definición, somera y preliminar de sesgo tecnológico sería por tanto: toda debilidad tecnológica de una herramienta que produce una desviación en el procesamiento o análisis de la información básica obtenida, originada durante el diseño de la misma, con capacidad para distorsionar los resultados del trabajo en mayor o menor medida y que obedece a las limitaciones presentadas por las personas y la propia naturaleza no neutral del avance científico.

Consideramos que la relevancia de este hecho presentado tiene mucho que ver con la necesidad de que el análisis de inteligencia, en muchas ocasiones más un arte que una ciencia, avance hacia la cimentación de un cuerpo teórico sólido –o varios, en realidad-, la estandarización de sus métodos y prácticas, el acercamiento máximo al método científico y, en general, el establecimiento definitivo como disciplina cuyos resultados sean corroborables y replicables. Para tal fin, tan importante es maximizar la potencia que ofrece el análisis de inteligencia como reconocer sus defectos y debilidades. No obstante, ello no debe suponer una limitación de la parte más creativa del análisis.

A continuación se presenta un caso práctico destinado a ofrecer ejemplos reales de lo que se ha definido como sesgo tecnológico en este documento. En este punto aún temprano de la

investigación no se pretende explicar por qué se producen variaciones en los resultados obtenidos a través de varias herramientas –aunque las condiciones en las que se han realizado las peticiones de información hayan sido las mismas- sino presentar pruebas de que dichas variaciones existen y, por tanto, deben tenerse en cuenta. Los ejemplos seleccionados afectan principalmente a las labores de obtención y procesamiento de información.

CASO PRÁCTICO: USO DE HERRAMIENTAS ONLINE

La relación entre el análisis de inteligencia y las herramientas de apoyo es evidente. Entre ellas, son muy comunes en el repositorio de programas de uso intensivo de las unidades de inteligencia las de tipo online, que suelen estar creadas para aplicarse a una fuente de información concreta. Como se viene explicando a lo largo de todo el trabajo, es de vital importancia conocer los sesgos tecnológicos que presentan dichas herramientas para que, siendo conscientes de las debilidades que presenta la tecnología, se puedan establecer protocolos de actuación que mitiguen los errores producidos por estos sesgos.

Tipos de herramientas utilizadas: VPNs, buscadores (motores) web y *trackers* de redes sociales

Se presentan a continuación algunos ejemplos reales de tareas realizadas habitualmente por los analistas, en la que se evidencian algunos sesgos tecnológicos mediante las variaciones de los resultados adquiridos en las primeras fases del proceso de inteligencia –obtención y análisis-. Para asegurar la validez del trabajo, las búsquedas se han realizado manteniendo las condiciones más controladas posibles (herramientas similares, mismo ordenador, misma conexión y misma localización IP).

Los ejemplos expuestos se han dividido en tres clasificaciones correspondientes con tres tipos de herramientas diferentes y muy utilizadas por los analistas a la hora de obtener los resultados que, en una segunda fase, serán analizados. Dichos tipos son: conexiones VPN, buscadores web y *trackers* de redes sociales. Estos ejemplos no son limitativos, ya que los sesgos tecnológicos existen en todas y cada una de las herramientas utilizadas por los analistas. No obstante, con ellas pretendemos explicar tres sesgos tecnológicos bien diferenciados entre sí, demostrando que existen, deben conocerse y tenerse en cuenta a la hora de poder desarrollar un trabajo lo más objetivo posible:

- Conexiones VPN (Virtual Private Network o Red Privada Virtual): tradicionalmente, la utilización de una conexión VPN se ha relacionado con la intención, a menudo con fines ilícitos, de evitar la trazabilidad de las conexiones a Internet. Es decir, que se pueda navegar de manera privada sin que el proveedor de servicios sepa qué se busca, dónde se busca y qué información se consume. Otra característica de las conexiones VPN derivada de esta última y la que realmente interesa a los analistas de inteligencia es que evita la censura de resultados y las prohibiciones geográficas en la visualización de contenidos y es esto precisamente, el hecho de que permitan rodear las prohibiciones geográficas, lo que hace que varíen los resultados dependiendo de si la conexión se realiza desde, por ejemplo, un país u otro, ya que se salta el bloqueo que utilizan algunos *sites* para enviar contenidos selectivos. Un claro ejemplo de esto es el uso de VPN en China para acceder a contenidos censurados o a ciertas redes sociales: un usuario que se encuentre en China pero que navegue con una VPN alojada en Estados Unidos puede acceder a Facebook o Twitter, a pesar de estar prohibido, ya que a todos los efectos este usuario, que físicamente se encuentra en China, está conectado desde Estados Unidos. Cierto es que este ejemplo concreto nace de las prohibiciones establecidas por los regímenes de cada país –prohibiciones políticas- y no tanto por la actuación de un sesgo tecnológico. Así y todo, incluso sin existir una censura previa, un buscador ofrece diferentes resultados para una misma búsqueda en tanto en cuanto varíe la localización de conexión, por ejemplo, mediante el uso de una VPN, como se verá en los ejemplos adjuntos.
- Motores de búsqueda: existen numerosos buscadores globales en Internet, pero cada uno ofrece resultados diferentes por motivo de sus algoritmos de búsqueda y criterios de presentación de resultados. Una forma de mitigar el sesgo tecnológico intrínseco que sufren los buscadores es conocer para qué tipo de búsquedas ha sido desarrollado cada uno de ellos, de manera que se utilicen dependiendo de la tarea en progreso. Google y otros ofrecen opciones de búsquedas avanzadas que permiten ajustar hasta niveles considerables los resultados de las peticiones. El uso de operadores booleanos es otra manera de actuar en este sentido, pero esto requiere cierto conocimiento y tiempo. Normalmente es más rápido y fiable utilizar buscadores desarrollados específicamente para cubrir una tarea específica: así, el buscador “Pipl” creado para buscar información de personas, ofrece resultados más ajustados y específicos si el objeto de la búsqueda es una identidad personal. Por último, si el objetivo es adquirir

conocimiento sobre un tema general, 'Ask' ofrece resultados ajustados basados en respuestas a preguntas realizadas, por lo que enriquece y facilita el acceso a más información.

- *Trackers* de redes sociales: otras de las fuentes más utilizadas en la fase de obtención del análisis de inteligencia son las redes sociales –Twitter, Facebook, Instagram, LinkedIn, etc.-. Desde las mismas se puede acceder a numerosa información de notable valor en términos OSINT –para la gestión de crisis reputacionales 2.0, conteo de número de menciones de términos de interés, ratio de influencia de perfiles de interés, etc.-. Existen numerosas herramientas que, además de las ofrecidas por la propia red social, presentan automáticamente datos de especial relevancia para los análisis a los que normalmente un analista debe de hacer frente. Una vez más, en la monitorización de estas fuentes, se detectan sesgos tecnológicos que provocan que, habitualmente, no coincidan los resultados obtenidos según se utilice una u otra herramienta. De hecho, estas búsquedas ni siquiera coinciden con las realizadas manualmente desde aplicación de la red social monitorizada.

Todos estos sesgos tecnológicos quedarán demostrados a continuación, donde exponemos ejemplos reales para cada uno de los tipos anteriormente explicados, quedando evidenciada la importancia de ser conocedores como analistas de inteligencia de los sesgos sufridos y teniendo la obligación, siempre que sea posible, de mitigarlos.

Resultados obtenidos mediante uso de VPN

A continuación se compara la misma búsqueda con dos conexiones VPN diferentes. En este caso la búsqueda es 'lula da silva', realizada el día 13/07/17 coincidiendo con la publicación de la noticia por la que el ex presidente brasileño ha sido condenado a 9 años de prisión acusado de corrupción.

La primera captura de pantalla corresponde con la búsqueda 'lula da silva' en 'google.com.br' con conexión desde Madrid. La segunda corresponde a la misma búsqueda desde el mismo buscador con conexión VPN desde Brasil:

lula da silva  

Todo [Imágenes](#) [Noticias](#) [Vídeos](#) [Maps](#) [Más](#) [Preferencias](#) [Herramientas](#)

Cerca de 10,300,000 resultados (0.50 segundos)

Noticias destacadas

		
<p>Lula é condenado por Moro em primeiro processo na Lava Jato</p> <p>UOL Notícias · Hace 13 hor...</p>	<p>El expresidente de Brasil Lula da Silva, condenado a 9 años y medio de prisión por</p> <p>BBC.com · Hace 13 horas</p>	<p>Brasil: condenan a Luiz Inácio Lula da Silva a nueve años de prisión</p> <p>Clarín · Hace 7 horas</p>

→ [Más sobre lula da silva](#)

El expresidente de Brasil Lula da Silva, condenado a 9 años y medio ...
www.bbc.com/mundo/noticias-america-latina-40577510 ▼
 El expresidente de Brasil Luiz Inácio Lula da Silva fue condenado este miércoles a 9 años y medio de prisión por corrupción pasiva y lavado de dinero.

Luiz Inácio Lula da Silva - Wikipedia, la enciclopedia libre
https://es.wikipedia.org/wiki/Luiz_Inácio_Lula_da_Silva ▼
 Luiz Inácio "Lula" da Silva ([Acerca de este sonido audio](#) (?·i); Caetés, 27 de octubre de 1945) es un político brasileño, presidente de la República Federativa de ...

Brasil: condenan a Luiz Inácio Lula da Silva a nueve años de prisión
https://www.clarin.com/.../brasil-condenan-luiz-inacio-lula-da-silva-anos-prision_0_r1... ▼
 hace 8 horas - El ex presidente brasileño fue condenado por corrupción y lavado de dinero, una sentencia que podrá recurrir en libertad pero que amenaza ...

Lula é condenado por Moro em primeiro processo na Lava Jato ...
<https://noticias.uol.com.br/.../lula-e-condenado-por-moro-em-pri...> ▼ [Traducir esta página](#)
 hace 14 horas - ... condenou, nesta quarta-feira (12), o ex-presidente Luiz Inácio Lula da Silva (PT) a nove anos e seis meses de prisão em regime fechado.

Lula da Silva | Últimas Notícias | EL PAÍS Brasil
<https://brasil.elpais.com> › [Mais notícias](#) ▼ [Traducir esta página](#)
 Informações sobre Luiz Inácio Lula da Silva, sua possível participação como ministro da Casa Civil do governo Dilma e as investigações da Operação Lava ...

Lula Da Silva, auge y caída del "hijo de Brasil" - La Vanguardia
www.lavanguardia.com/internacional/20170713/.../lula-auge-caida-hijo-brasil.html ▼
 hace 3 horas - "Prueben que soy corrupto e iré caminando a la cárcel". Así defendía Luiz Inácio Lula da Silva su inocencia cuando ya estaba cercado por la ...

Por qué Lula da Silva sigue libre y otras claves de la condena de ...
www.lanacion.com.ar › [El Mundo](#) › [Crisis en Brasil](#) ▼
 hace 10 horas - El ex presidente brasileño Luiz Inácio Lula da Silva fue condenado este miércoles a nueve años y medio de cárcel por corrupción y lavado ...

Lula da Silva, condenado a prisión por corrupción y lavado de dinero ...
cnnespanol.cnn.com/.../lula-da-silva-condenado-a-prision-por-corrupcion-y-lavado-d... ▼
 hace 13 horas - Luiz Inacio Lula da Silva, expresidente de Brasil, fue condenado a 9 años y 6 meses en prisión, según la sentencia del juez Sergio Moro.

Lula da Silva, condenado a nueve años de cárcel por corrupción ...
www.elmundo.es › [Internacional](#) › [Brasil](#) ▼
 hace 14 horas - El ex presidente brasileño Luiz Inácio Lula da Silva fue condenado a nueve años y seis meses de cárcel por corrupción y blanqueo de dinero, ...

Captura de pantalla correspondiente a la búsqueda "lula da silva" con conexión VPN desde Brasil.

lula da silva

Todo Imágenes Noticias Vídeos Maps Más Preferencias Herramientas

Cerca de 9.950.000 resultados (0,38 segundos)

Noticias destacadas

		
Lula é condenado por Moro em primeiro processo na Lava Jato	El expresidente de Brasil Lula da Silva, condenado a 9 años y medio de prisión por...	Brasil: condenan a Luiz Inácio Lula da Silva a nueve años de prisión
UOL Noticias - Hace 13 hor...	BBC.com - Hace 13 horas	Clarín - Hace 7 horas

→ Más sobre lula da silva

El expresidente de Brasil Lula da Silva, condenado a 9 años y medio ...
www.bbc.com/mundo/noticias-america-latina-40577510
 El expresidente de Brasil Luiz Inácio Lula da Silva fue condenado esta miércoles a 9 años y medio de prisión por corrupción pasiva y lavado de dinero.

Brasil: condenan a Luiz Inácio Lula da Silva a nueve años de prisión
https://www.clarin.com/.../brasil-condenan-luiz-inacio-lula-da-silva-anos-prision_0_r1...
 hace 8 horas - El ex presidente brasileño fue condenado por corrupción y lavado de dinero, una sentencia que podrá recurrir en libertad pero que amenaza ...

Luiz Inácio Lula da Silva - Wikipedia, la enciclopedia libre
https://es.wikipedia.org/wiki/Luiz_Inacio_Lula_da_Silva
 Luiz Inácio "Lula" da Silva (Acerca de este sonido audio (?:) · Casés, 27 de octubre de 1945) es un político brasileño, presidente de la República Federativa de ...

Lula é condenado por Moro em primeiro processo na Lava Jato ...
<https://noticias.uol.com.br/.../lula-e-condenado-por-moro-em-pri...>
 hace 14 horas - ... condenado, nesta quarta-feira (12), o ex-presidente Luiz Inácio Lula da Silva (PT) a nove anos e seis meses de prisão em regime fechado.

Lula Da Silva, auge y caída del "hijo de Brasil" - La Vanguardia
www.lavanguardia.com/internacional/20170713/.../lula-auge-caida-hijo-brasil.html
 hace 3 horas - "Prueben que soy corrupto e iré caminando a la cárcel". Así defendía Luiz Inácio Lula da Silva su inocencia cuando ya estaba cercado por la ...

Lula da Silva | Últimas Noticias | EL PAÍS Brasil
<https://brasil.elpais.com> · Más noticias · Traducir esta página
 Informações sobre Luiz Inácio Lula da Silva, sua possível participação como ministro da Casa Civil do governo Dilma e as investigações da Operação Lava ...

Lula, condenado a nueve años de cárcel por corrupción ...
<https://internacional.elpais.com> · Internacional · Crisis política en Brasil ·
 hace 10 horas - El expresidente más popular del país, Luiz Inácio Lula da Silva, fue condenado este miércoles a nueve años de cárcel por corrupción y ...

Lula da Silva fue condenado a nueve años y medio de prisión - Infobae
www.infobae.com/.../lula-da-silva-fue-condenado-a-nueve-anos-y-medio-de-prision/
 hace 13 horas - El juez federal Sergio Moro condenó hoy al ex presidente Luiz Inácio Lula da Silva a nueve años y medio de prisión por los crímenes de ...

Lula da Silva, condenado a prisión por corrupción y lavado de dinero ...
cnnespanol.cnn.com/.../lula-da-silva-condenado-a-prision-por-corrupcion-y-lavado-d-...
 hace 13 horas - Luiz Inacio Lula da Silva, expresidente de Brasil, fue condenado a 9 años y 6 meses.

Captura de pantalla correspondiente a la búsqueda "lula da silva" con conexión VPN desde España.

En la comparación de resultados se observa que los resultados totales de la primera búsqueda son 10.300.000, mientras que en la segunda se detectan 9.950.000 resultados. Además, aunque todos los resultados corresponden a medios de comunicación –debido a que todos tratan la noticia de la condena por corrupción- ni el orden ni las fuentes de los mismos coinciden.

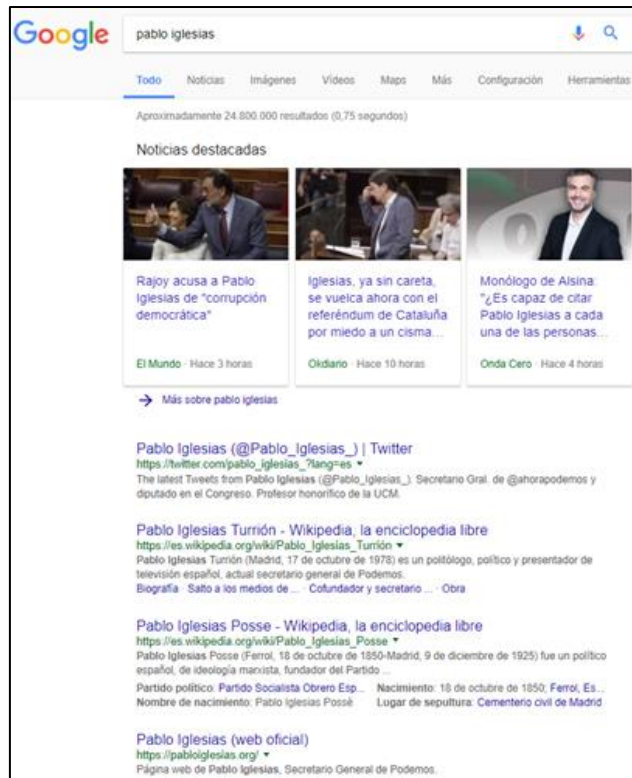
Esto responde a todo lo explicado anteriormente sobre el sesgo tecnológico existente en las conexiones VPN. Será importante para los analistas de inteligencia conocer este sesgo para valorar si, en algunas de las acciones a realizar, conviene conectarse con una VPN alojada en otro país y así tener acceso a información más completa y objetiva para realizar su análisis.

Resultados obtenidos mediante motores de búsqueda

Como se ha señalado, existen numerosos motores/buscadores generalistas que ofrecen resultados distintos ante búsquedas iguales. Ello se confirma mediante un ejemplo real de una búsqueda de una identidad, ‘Pablo Iglesias’, realizada de la misma manera con los buscadores Google y Ask, y se analizan los cinco primeros resultados obtenidos por cada uno de ellos.

Google:

Mediante la búsqueda simple de ‘Pablo Iglesias’ en Google, el buscador ofrece en primer lugar noticias destacadas, después su perfil de Twitter y, por último, dos entradas de Wikipedia seguidas de su web oficial:



Búsqueda simple de 'Pablo Iglesias' en Google.

Ask:

Realizando exactamente la misma búsqueda en Ask comprobamos como los primeros resultados que ofrece corresponden a noticias agregadas de diferentes medios de comunicación, un vídeo de YouTube y la página web de la Fundación Pablo Iglesias (sin relación con la identidad buscada):



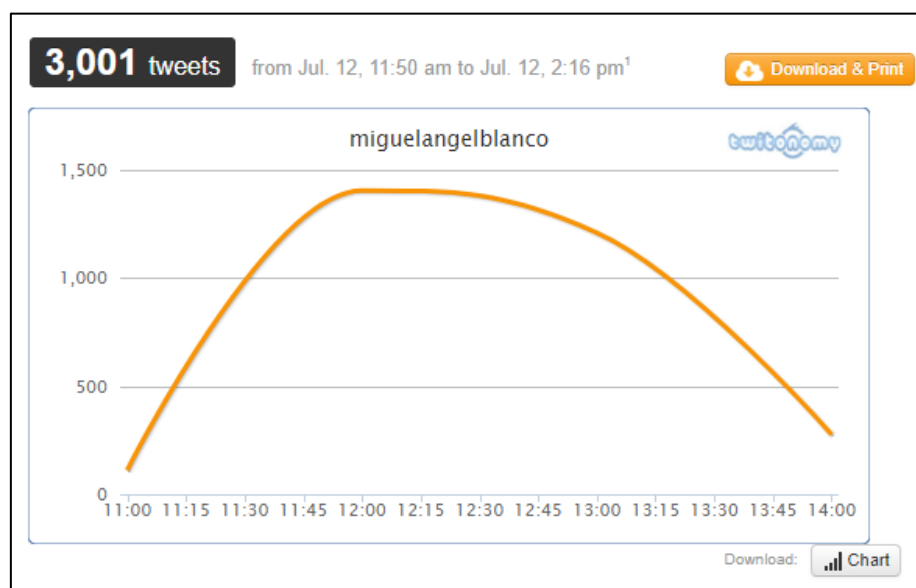
Búsqueda simple de 'Pablo Iglesias' en Ask.

Se demuestra así la diferencia, en ocasiones muy notable, entre los resultados ofrecidos por un buscador y otro. Lógicamente, la búsqueda de información sobre un personaje público, como ocurre en el caso explicado, puede cubrirse mediante una simple petición en Google, que iremos ajustando según se requiera. Sin embargo, si lo que se debe examinar es la huella digital de una persona desconocida es necesario recurrir a otros tipos de buscadores específicamente diseñados para tal fin ('Pipl', por ejemplo), que reducen el ruido generado por motores no especializados. De esta manera, y aunque todos los buscadores cuentan con sus propios sesgos derivados de la tecnología aplicada en su diseño, se pueden corregir estos 'fallos' enfocando las búsquedas a través de herramientas que nos proporcionen resultados más cercanos a los deseados.

Resultados obtenidos mediante *trackers* de redes sociales:

Para este tipo de herramientas se han implementado búsquedas sobre el *hashtag* o etiqueta '#miguelangelblanco' dos herramientas de tipo online muy populares: Twitonomy y Keyhole.

Twitonomy:



Número de publicaciones por horas

🔄 Most retweeted tweets

- 
PSOE @PSOE 12:22 PM - 12 Jul 2017 via Twitter for iPhone 🔄 85 ★ 119
 Minuto de silencio de la Comisión Ejecutiva Federal del @PSOE y de los trabajadores de Ferraz en homenaje a... twitter.com/i/web/status/8...
- 
Ministerio Interior @interiorgob 12:15 PM - 12 Jul 2017 via Twitter Web Client 🔄 53 ★ 69
 La unidad de todos, imprescindible para vencer al terrorismo #MiguelAngelBlanco #20AñosMAB <https://t.co/drDFVImx1U>
- 
CCD @CCDcentro 12:21 PM - 12 Jul 2017 via Twitter Web Client 🔄 50 ★ 40
 #MiguelAngelBlanco IN MEMORIAN a todas las Víctimas del Terrorismo, de la sinrazón y la cobardía !NUESTRA GRATITUD... twitter.com/i/web/status/8...
- 
Maite Pagazaurtundúa @maitepagaza 11:57 AM - 12 Jul 2017 via Twitter for iPhone 🔄 43 ★ 42
 Hoy en el @Europarl_ES vamos a recordar a #MiguelAngelBlanco y el Espíritu de Ermua <https://t.co/j0nMgjzAy8>
- 
Maite Pagazaurtundúa @maitepagaza 12:47 PM - 12 Jul 2017 via Twitter for iPhone 🔄 40 ★ 42
 La memoria sigue viva. #MiguelAngelBlanco <https://t.co/8n7mygeVYc>
- 
Ciudadanos Madrid C. @CsMadridCiudad 11:52 AM - 12 Jul 2017 via Media Studio 🔄 36 ★ 29
 ☐@begonavillacis "La unión hace 20 años por #MiguelAngelBlanco hay que celebrarla también con unión. Tenemos que ma... twitter.com/i/web/status/8...











Publicaciones más retuiteadas

Top hashtags

#miguelangelblanco 1789 #miguelAngelblanco 910 #20añosmab 361
 #homenajeblancoep 62 #ermua 56 #eta 46 #20anosmab 42 #estápasando 24
 #valladolid 24 #pp 22

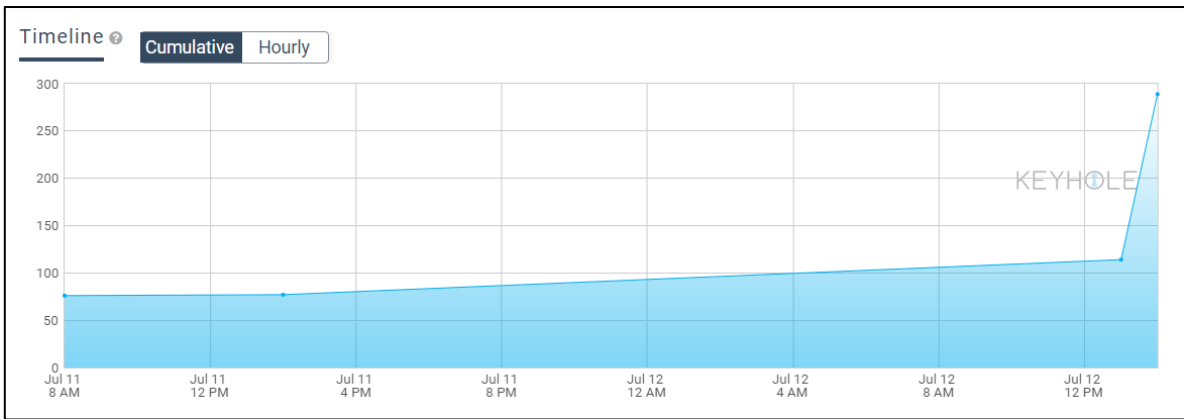
Hashtags más utilizados

★ Most engaging users

 204	 167	 122	 90	 80
 65	 64	 59	 58	 48

Usuarios más importantes

Keyhole:

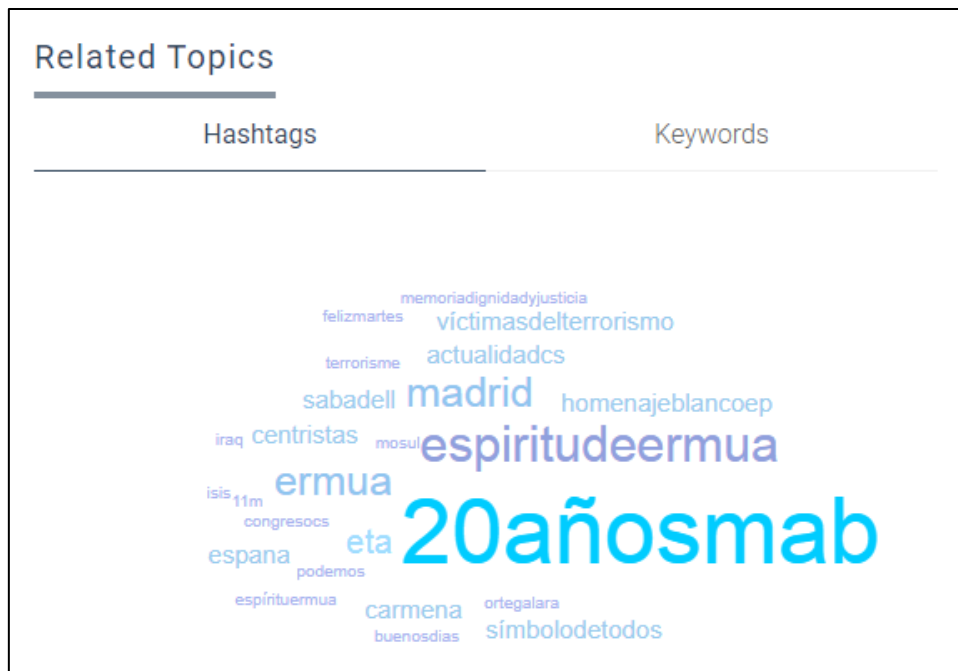


Número de publicaciones por horas

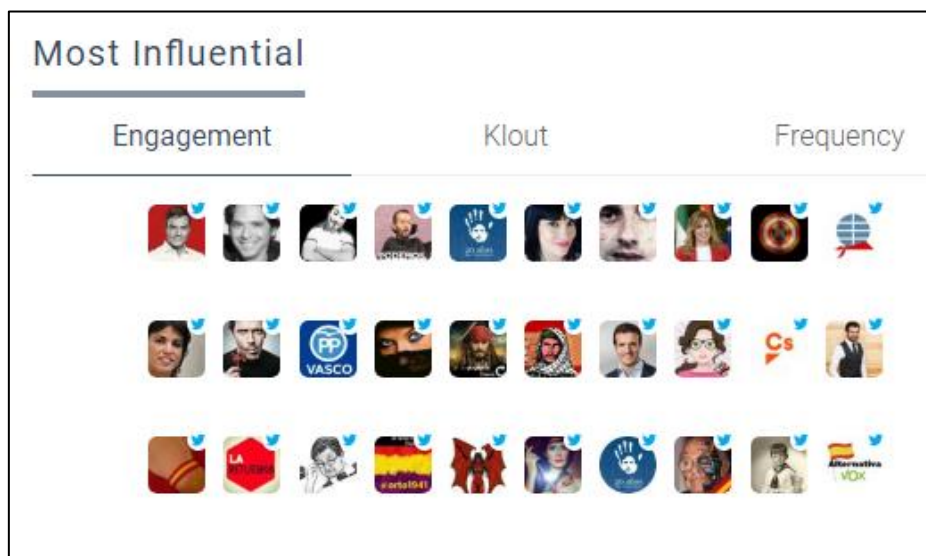
Top Posts

RT / Likes	Klout	Recent
 Percival Manglano @pmanglano Jul 10 Homenajes hechos por Carmena y sus concejales: -Hugo Chávez✔ -Fidel Castro✔ -Alfon✔ - #MiguelAngelBlanco✘		1,518 ↻
 Pedro Sánchez @sanchezcastejon Jul 10 Aquel día todo un país se unió contra el miedo y se echó a la calle gritando libertad. Siempre en nuestro recuerdo #MiguelAngelBlanco https://t.co/q7fEHRnGcS		1,100 ↻
 ★Antídoto★ @Yo_Soy_Asin Jul 10 Exto es lo que hizo Aznar 2 años después del asesinato por parte de ETA de #MiguelAngel Blanco Imaginar si esto lo hiciera podemos... https://t.co/JtMtNDmqSA		958 ↻
 Percival Manglano @pmanglano Jul 10 Y aquí tenemos a una concejala de Carmena negándose a exhibir una pancar stinquir entre víctimas, como con #MiguelAngelBlanco. https://t.co/PhCFtdov		675 ↻

Publicaciones más retuiteadas



Hashtags más utilizados



Usuarios más importantes

En el caso de este tipo de herramientas el sesgo tecnológico es mucho más visible que en los ejemplos anteriores. Cada una de ellas está diseñada para trabajar atendiendo a criterios distintos, lo que hace que ofrezcan resultados que, como se ha observado, no representan la realidad objetiva analizada.

Así, si se compara cada una de las capturas de pantalla anteriores en uno y otro *tracker* se observa que las publicaciones presentadas no coinciden debido a cómo está representada la línea temporal, ya que en el expuesto por Twitonomy las publicaciones aumentan en la zona central, correspondiendo con las 12:00h de la mañana, mientras que en el de Keyhole se mantienen más o menos estables hasta esa misma hora. El incremento de contenidos en las horas centrales de la mañana coincide pero la representación gráfica difiere notablemente, lo que puede provocar que los analistas interpreten su significado de manera desigual.

En cuanto al resto de datos representados en las capturas de pantalla incluidas –publicaciones más retuiteadas, *hashtags* más utilizados y usuarios más importantes- los contenidos no coinciden en una y otra herramienta, demostrándose de nuevo que ante una misma petición de información cada una se comporta de forma diferente. Queda de manifiesto el sesgo tecnológico existente en este tipo de tecnología, muy utilizada en el análisis de redes sociales y el ámbito más amplio de OSINT.

CONCLUSIONES

Ante los hechos descritos en este trabajo, y teniendo en cuenta que:

- i) A pesar de que las herramientas no son personas, su desarrollo está influido por sus diseñadores. Por tanto, el sesgo humano está transferido a la herramienta de una manera similar a como está presente en el analista de inteligencia.
- ii) Este sesgo afecta de manera especial a la recolección de información y su procesamiento, así como – en su caso – al análisis automático de la información dada.
- iii) Esta afectación se suma a la que posteriormente tendrá el sesgo cognitivo del analista.
- iv) Los casos prácticos estudiados demuestran que herramientas que teóricamente sirven para el mismo fin –buscadores; *trackers* de redes sociales- muestran resultados notablemente divergentes, sin que pueda achacarse esta divergencia a ‘errores técnicos’, sino a decisiones conscientes de sus desarrolladores tomadas sobre dónde buscar, cómo buscar y cómo presentar y representar los resultados.
- v) Estas decisiones no son públicas y, a menudo, podrían no tenerse en cuenta asumiendo que una herramienta –por el hecho de ser una tecnología- actúa de manera neutral presentando información objetiva y exenta de juicios apriorísticos.

Podemos concluir que el análisis de inteligencia requiere atender a la existencia de sesgos tecnológicos en las herramientas que se utilizan –especialmente las que se refieren a fuentes OSINT, que por su extensión son las más proclives a requerir automatización para ser explotadas- para ser consciente de las limitaciones que hereda un análisis y para conseguir paliar en la medida de lo posible, sus efectos sobre la calidad del producto de inteligencia.

Las herramientas cumplen una función vital en la inteligencia actual, función que es probable que se vea aumentada con el paso de los años. El desarrollo tecnológico ha dotado a estos programas de la capacidad de discernir, cribar, descartar –habilidades del razonamiento humano- y es ahí precisamente donde están expuestas a sesgos similares a los de los analistas.

Como decimos, su uso es necesario pero, igual que los productos elaborados por analistas deben ser revisados y validados para identificar y mitigar los posibles sesgos que puedan haberlo afectado, los productos emitidos por herramienta pueden –y deben- soportar una exploración similar.

BIBLIOGRAFÍA

- Burns, Edward. 2017. “Developing AI apps free from bias crucial to avoid analytics errors”, *Search Business Analytics*. Disponible en web: <http://searchbusinessanalytics.techtarget.com/news/450422298/Developing-AI-apps-free-from-bias-crucial-to-avoid-analytics-errors> [Consulta: 10 de julio de 2017].
- Jordán, J. 2011. “Introducción al análisis de inteligencia”, *Grupo de estudios en Seguridad Internacional (GESI)*. Disponible en web: <http://www.ugr.es/~gesi/analisis/2-2011.pdf> [Consulta: 06 de junio de 2017]
- Jordán, J. (2014), “Análisis de Inteligencia estratégica”, en Jordán, J. (coord.) *Manual de Estudios Estratégicos y Seguridad Internacional*, Madrid: Plaza y Valdés. [Tomada de Martin, Diego. 2016. “El análisis de Inteligencia: técnicas de análisis y fuentes de error. Una aproximación desde la teoría argumentativa”, *Revista de Estudios en Seguridad Internacional*, 2 (1): 103-123].
- Jordán, J. 2016. “Una revisión del Ciclo de Inteligencia”, *Grupo de estudios en Seguridad Internacional (GESI)*. Disponible en web: <http://www.seguridadinternacional.es/?q=es/content/una-revisi%C3%B3n-del-ciclo-de-inteligencia> [Consulta: 29 de junio de 2017].
- Lowenthal, Mark M. 2012. *Intelligence: From Secrets to Policy*. Washington DC: CQ Press.
- Manzano, Fernando. 2010. “Analistas, el riesgo de ser humanos: consecuencias derivadas de los esquemas mentales”, *Inteligencia y Seguridad: Revista de Análisis y Prospectiva*,(7): 117-130. [Tomada de Martin Ortega, Diego. 2016. “El análisis de Inteligencia: técnicas de análisis y fuentes de error. Una aproximación desde la teoría argumentativa”, *Revista de Estudios en Seguridad Internacional*, (2) 1: 103-123].
- Martín Ortega, Diego, 2016. “El análisis de Inteligencia: técnicas de análisis y fuentes de error. Una aproximación desde la teoría argumentativa”, *Revista de Estudios en Seguridad Internacional*, (2) 1: 103-123. Disponible en web: <http://dx.doi.org/10.18847/1.3.6> [Consulta el 31 de mayo de 2017].
- Rosenblatt, Gideon. 2016. “Machined Prejudice: Three Sources of Technology Bias”, *The Vital Edge*. Disponible en web: <http://www.the-vital-edge.com/technology-bias/> [Consulta: 9 de julio de 2017].
- Roth, Lorna. 2009. “Looking at Shirley, the Ultimate Norm: Colour Balance, Image Technologies, and Cognitive Equity”, *Canadian Journal of Communication*, 34 (1):

111-136.Sáiz, Enrique. 2010. "Unas primeras aproximaciones a los sesgos cognitivos en el análisis de Inteligencia", *Inteligencia y Seguridad: Revista de Análisis y Prospectiva*, (7): 197-236. [**Tomada de** Martín Ortega, Diego, 2016. "El análisis de Inteligencia: técnicas de análisis y fuentes de error. Una aproximación desde la teoría argumentativa", *Revista de Estudios en Seguridad Internacional*, (2) 1: 103-123].