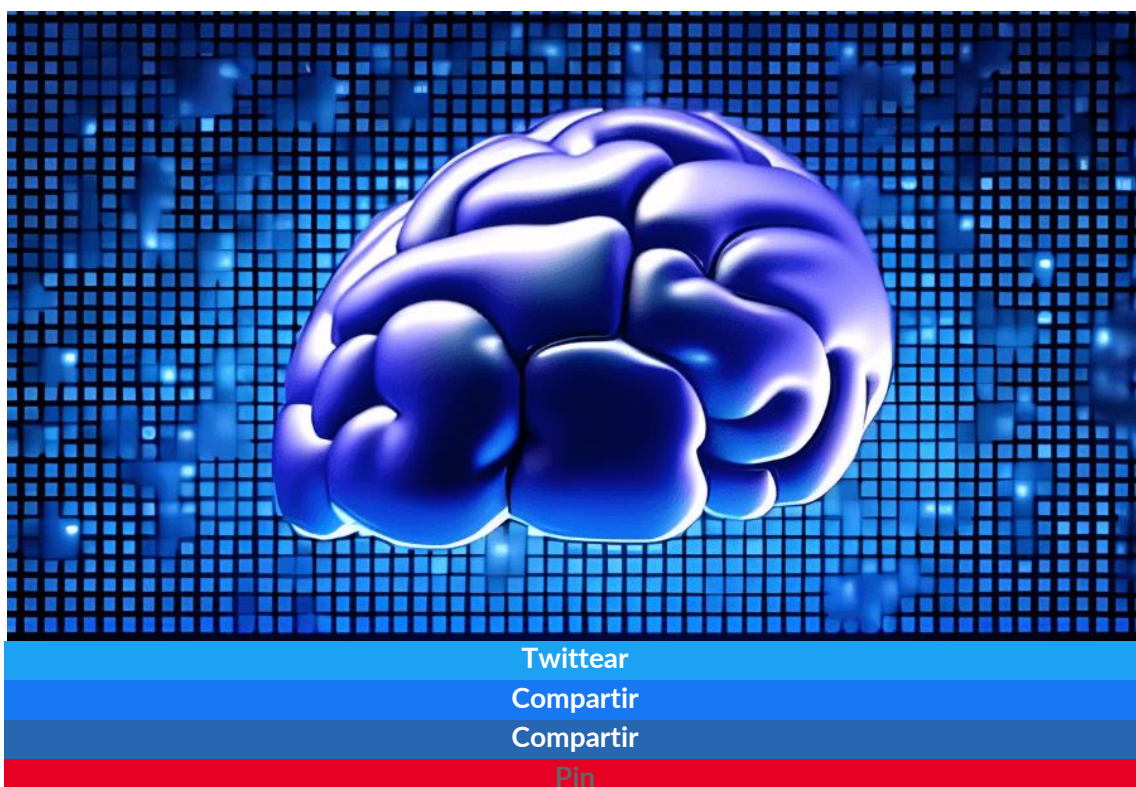


Inteligencia Artificial, datos y sesgos: notas para las administraciones tributarias

- 17 julio, 2023
- Escrito por: Antonio Seco



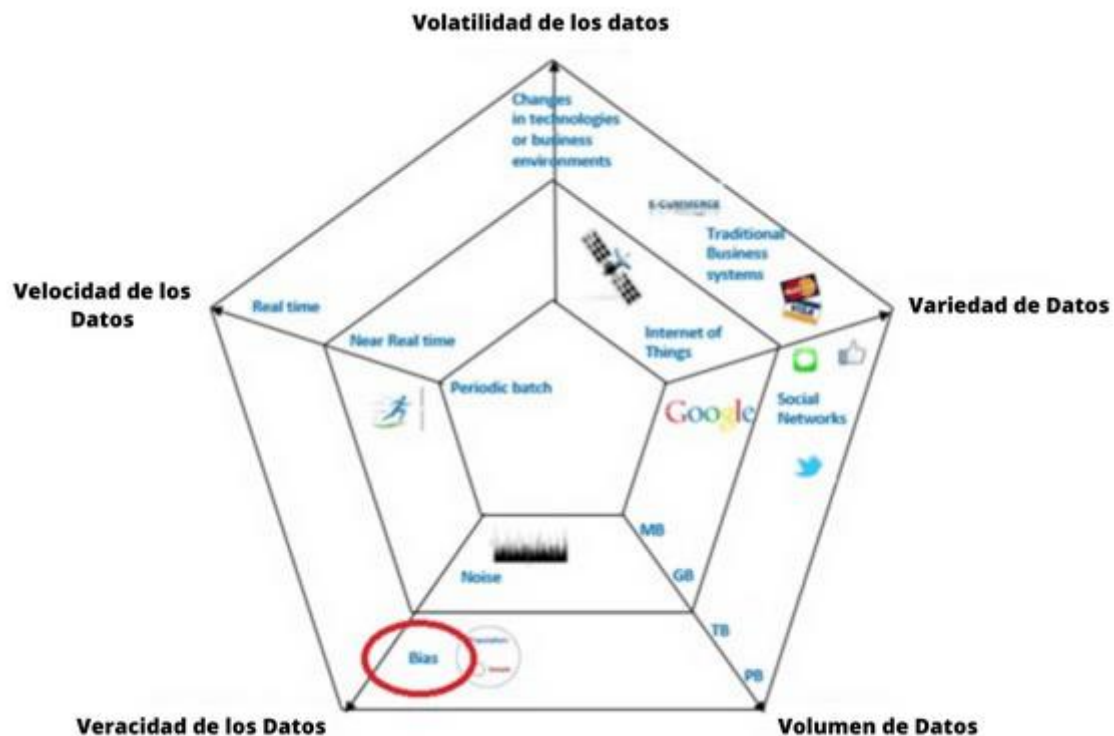
COMPARTIR

Es evidente que la Inteligencia Artificial (IA) – aquí tratada genéricamente como las tecnologías fundadas en Big Data, analítica de datos, algoritmos, aprendizaje de máquinas, redes neuronales y otras técnicas de simulación de la inteligencia humana – representa una gran oportunidad de avance para las Administraciones Tributarias (AT).

Entre los potenciales de uso de las tecnologías digitales en las AT, se verifica que las mismas brindan considerable oportunidad de transparencia, imparcialidad, precisión y eficiencia en sus procedimientos. Pueden, además, simplificar procesos administrativos, reducir tiempo y costos para los contribuyentes, fortalecer la fiscalización, mantener bajo vigilancia a los contribuyentes no conformes. De igual manera, el procesamiento y análisis de grandes cantidades de datos por las AT contribuye a la gestión del gasto público y aumenta la percepción pública del control de la autoridad tributaria., pronosticando los ingresos fiscales con precisión y transparencia [Almeida, 2020].

Todo esto está basado en datos. Muchos datos. Inmensas cantidades de datos – Big Data – caracterizadas por los 5V – variedad, volumen, volatilidad, velocidad, y

veracidad. El gráfico a continuación destaca los elementos que componen cada característica del Big Data, con su complejidad y amplitud avanzando desde el centro hacia el exterior.



Fuente: Hammer, 2017

La UNECE [1] clasifica a los datos relevantes como subproductos de los siguientes orígenes:

- **I) Redes sociales (información de origen humana):** redes sociales, blogs y comentarios, documentos personales, fotos, vídeos, búsquedas por Internet, mensajes de texto, mapas generados por usuarios, correo electrónico.
- **II) Sistemas de negocio tradicionales (datos mediados por procesos):** datos producidos por agencias públicas (registros médicos); datos producidos por negocios (transacciones comerciales, registros bancarios, registros bursátiles, comercio electrónico, tarjetas de crédito).
- **III) Internet de las cosas (IoT - datos generados por máquinas):** datos producidos por sensores fijos y móviles, datos producidos por sistemas computarizados (logs y web logs).

Tales cantidades, complejidades y fuentes requieren una estrategia y nuevos modelos de gestión para los datos de las AT, es decir, una gobernanza de datos, conforme explorado en [Zambrano, 2022].

Asimismo, regulaciones de protección de datos establecidas por distintos gobiernos – como la **General Data Protection Regulation (GDPR)** efectivada en 2018 para la Europa – implican en nuevas responsabilidades para el tratamiento de datos personales, basadas en los siguientes principios: (a) legalidad, equidad y transparencia; (b) recopilación para fines específicos, explícitos y legítimos y no procesamiento de manera incompatible con esos fines – limitación de propósito; (c) adecuación, pertinencia y limitación a lo necesario en relación con los fines para los que se tratan – minimización de datos; (d) precisión y, cuando sea necesario, actualización – exactitud; (e) conservación en formato que permita la identificación de los interesados durante no más tiempo del necesario para los fines del procesamiento de los datos personales – limitación del almacenamiento; y (f) procesamiento de una manera que garantice la seguridad adecuada de los datos personales, incluida la protección contra el manejo no autorizado o ilegal y contra la pérdida, destrucción o daño accidental – integridad y confidencialidad [European Union, 2018].

En IA, sigue válida la vieja expresión “*basura entra, basura sale*”: en cualquier sistema, la calidad de la salida está determinada por la calidad de la entrada.

La OCDE alerta que para la analítica cumplir su promesa de ayudar a las administraciones tributarias a hacer mejores predicciones y sacar inferencias más sólidas, necesita conjuntos de datos representativos y precisos que capturen todos los hechos de las características y el comportamiento de los contribuyentes [OCDE, 2016]. Hay que considerar especialmente los datos de entrenamiento que se utilizan para construir un algoritmo de aprendizaje de máquina. Así, la calidad de los datos asume importancia fundamental para la IA.

Las aplicaciones iniciales del uso de IA resaltan la importancia creciente de una dimensión de calidad de datos [III] anteriormente en la “sombra” de las demás: el **sesgo** [III] que un conjunto de datos puede ocasionar. Este ocurre cuando no se incluyen variables que capturan adecuadamente el fenómeno que se quiere estudiar y predecir. Los sesgos pueden estar en los datos o en los diseños de algoritmos. Los algoritmos, a su vez, también pueden ser inducidos a sesgos por sus datos de entrenamiento.

Información sesgada produce resultados inexactos, con serios reflejos en los usuarios y en la sociedad. Como alerta la Agencia Europea para los Derechos Fundamentales, pueden generar resultados que infringen los derechos fundamentales de las personas. A ejemplo: un algoritmo de contratación que prefiere los hombres a las mujeres; un chatbot en línea que reproduce lenguaje racista después de un par de horas de su lanzamiento; traducciones automáticas con sesgos de género; y sistemas de reconocimiento facial que tienen mejor funcionamiento para los hombres blancos en comparación con las mujeres negras [FRA Focus, 2019].

Las AT son extremadamente sensibles para sesgos en sus datos y también ya experimentaron los amplios impactos sociales y políticos negativos de sistemas de IA con base en datos y/o algoritmos sesgados, destacándose la cuestión de la asignación de beneficios para cuidados de niños, sucedido en los Países Bajos [Leahey, 2022]. En

otro caso reciente, el Servicio de Impuestos Internos de los Estados Unidos (*IRS*) admitió que las personas negras están siendo auditadas de manera desproporcionada, exponiéndose a sanciones fiscales. El problema se atribuyó parcialmente a sesgos en el algoritmo de selección de casos [*Adams, 2023 & IRS, 2023*].

Hay que contar con la posibilidad de ocurrencia de sesgo en el control de riesgos fiscales y la selección de contribuyentes a efectos de inspecciones tributarias cuando, por ejemplo, el modelo opera y aprende únicamente a partir de los datos suministrados sobre aquellos obligados que fueron objeto de una regularización o sanción. El sesgo existirá desde el momento en que no se cuente con los datos relativos a todos los supuestos de defraudación, sino sólo de los que hayan sido detectados [*López, 2022*].

La mayoría de los conjuntos de datos reales tiene sesgos ocultos. Ser capaz de detectar el impacto del sesgo en los datos del modelo y luego reparar el modelo es fundamental.

Considerando datos provenientes de la Internet, actualmente una importante fuente para las AT, hay que tener en cuenta que la propia red, su alcance y tipología de usuarios ya introducen sesgos en los datos recolectados. Como ejemplo, las indicaciones a continuación son relativas a la Unión Europea en 2018: el acceso a la Internet se distribuye por 98% de hogares de altos ingresos, 74% de hogares de bajos ingresos, con promedio general de 78%; en los Países Bajos, este promedio es de 98%, en Bulgaria, 72%; 11% de las personas dicen nunca usar Internet, predominando las situadas entre 55 y 74 años, con bajo nivel de educación y mujeres; solamente un promedio de 56% de la población utiliza redes sociales, entre ellos estudiantes y personas de altos ingresos [*FRA Focus, 2019*]. Por lo tanto, hay limitaciones para uso amplio y genérico de datos de esta fuente, dependiendo de los objetivos de la aplicación. Hay que notar que estas consideraciones son para personas naturales, no para empresas, y que, probablemente, para los individuos de interés tributario, estos sesgos sean más reducidos.

Aunque se dice que es más fácil sacarle el sesgo a una máquina que a una mente humana, el problema de identificación y resolución sigue complejo, con iniciativas de empresas, países y asociaciones múltiples trabajando sobre el tema.

De estos esfuerzos, aunque en etapas iniciales, ya resultan orientaciones de interés. En esta línea, son presentados seis pasos esenciales para tratar los sesgos, conforme indicado en [*Manyika, Silberg, Presten, 2019*]:

- 1- Los líderes institucionales deberán mantenerse actualizados en este campo de investigación, que está en rápida evolución;
-
- 2- Cuando su organización esté implementando IA, establezca procesos de responsabilidad que puedan mitigar el sesgo. Google e IBM han publicado marcos sobre el tema [*IV*];
-

- 3- Participe en conversaciones basadas en hechos sobre posibles sesgos humanos. Por ejemplo, ejecutar un algoritmo y comparar sus resultados con decisiones tomadas por humanos y utilizar "técnicas de IA explicable" [V] para comprender cómo el algoritmo llegó al resultado indicado;
-
- 4- Considere cómo humanos y máquinas pueden trabajar juntos para mitigar sesgos. Algunas técnicas de sistemas de IA permiten que humanos participen de sus decisiones e indiquen las direcciones más adecuadas, o hagan doble verificación previa de los resultados;
-
- 5- Invista más, obtenga más datos y adopte un enfoque multidisciplinario en investigaciones sobre sesgos; e
-
- 6- Invista en más diversidad en el campo de IA. Una comunidad de IA más diversificada estará mejor equipada para anticipar, rever y detectar sesgos.

Otro método de eliminación o reducción de sesgos es propuesto en [Uzzi, 2020], utilizando la conocida "prueba de sabor a ciegas" (*blind taste test*), y puede brindarnos una oportunidad de identificar y eliminar los sesgos de decisión de los algoritmos, incluso si no podemos eliminarlos por completo de nuestros propios hábitos mentales.

La búsqueda por la identificación, eliminación o control de sesgos en el uso de IA prosigue.

Además de investigaciones y experimentos de gobiernos y AT, asociaciones híbridas cooperan en el desarrollo de este campo de la IA, como el **AI Now Institute** (<https://ainowinstitute.org>), **The Alan Turing Institute** (<https://turing.ac.uk>), **AI Ethics and Governance of AI Initiative** (<https://aiethicsinitiative.org>) y **Partnership on AI** (<https://partnershiponai.org>). Organismos multilaterales de apoyo a las AT, como el CIAT, BID y FMI, tienen roles importantes, especialmente diseminando experiencias y financiando iniciativas.

Finalmente, se concluye con lo expresado por Martín López [López, 2022]: *"la inteligencia artificial no puede enjuiciarse como inherentemente buena o mala. En el ámbito del control de riesgos y de la lucha contra el fraude fiscal presenta un potencial innegable, aunque también riesgos que hay que tratar de minimizar. Mas la amenaza no es la herramienta en sí, sino un diseño o utilización inapropiados de la misma. En abstracto, la inteligencia artificial podría ayudar a la adopción de decisiones más objetivas, incluso, que las de índole humana. Pero, para ello, se requiere que la selección de los datos utilizados y la configuración y entrenamiento del modelo sean adecuados y fiscalizables, en orden a controlar y eliminar los sesgos discriminatorios que pudieran ocasionarse"*.

REFERENCIAS

Adams, C. 2023. "A Black professor has long said what the IRS now admits: The tax system is biased". May, 23. Blog NBC News. Disponible en: <https://www.nbcnews.com/news/nbcblk/irs-acknowledges-racial-bias-tax-auditing-based-professors-work-rcna85630>

Almeida, A. 2020. "The Necessary Dialogue Between Big Data and Trust in Brazil's Tax Administration". Oxford University / Blavatnik Scholl of Government, August

European Union. 2018. "General Data Protection Regulation (GDPR)". Consultado en: <https://gdpr-info.eu/>

FRA Focus. 2019. "Data Quality and Artificial Intelligence – Mitigating Bias and Error to Protect Fundamental Rights". European Agency for Fundamental Rights

Hammer, C. et al. 2017. "Big Data: Potential, Challenges and Statistical Implications". IMF Staff Discussion Note, September

IRS. 2023. Carta del Comisionado del IRS al Senado de los Estados Unidos sobre el problema de la selección para auditoría. Mayo 15, 2023. Disponible en: <https://www.irs.gov/pub/newsroom/werfel-letter-on-audit-selection.pdf>

Leahey, A. 2022. "We Can All Learn a Thing or Two From the Dutch AI Tax Scandal". Bloomberg Tax Newsletter. July, 12. Disponible en: <https://tinyurl.com/5ysh56zm>

López, J. 2022. "Inteligencia Artificial, Sesgos y No Discriminación en el Ámbito de la Inspección Tributaria". Crónica Tributaria. Número 182/2022. Disponible en: <https://dx.doi.org/10.47092/CT.22.1.2>

Manyika, J., Silberg, J., Presten, B. 2019. "What Do We Do About the Biases in AI?". Harvard Business Review, October 25

OCDE. 2016. "Advanced Analytics for a Better Tax Administration". OCDE Publications, Paris

Uzzi, B. 2020. "A Simple Tactic That Could Help Reduce Bias in AI". Harvard Business Review, November 04

Zambrano, R, et al. 2022. "Data Governance for Tax Administrations: A Practical Guide". CIAT / GIZ. Disponible en: <https://www.ciat.org/data-governance-for-tax-administrations-a-practical-guide/?lang=en>

[I] United Nations Economic Commission for Europe

[II] Dimensiones: exactitud, completitud, consistencia, integridad, razonabilidad, oportunidad, unicidad,

[III] "Error sistemático en el que se puede incurrir cuando al hacer muestreos o ensayos se seleccionan o favorecen unas respuestas frente a otras", Diccionario de la Lengua Española, Real Academia Española

[IV] <https://ai.google/responsibility/responsible-ai-practices/>

<https://www.ibm.com/opensource/open/projects/ai-fairness-360/>

[V] Requisitos para que las conclusiones de algoritmos de IA puedan ser explicadas están presentes en normas legales de algunos países. Para más informaciones, véase <https://www.ciat.org/inteligencia-artificial-explicable-xai-y-su-importancia-en-la-administracion-tributaria/>

2,035 total views, 76 views today