

La ciencia de datos, herramienta para la política actual

Data science, a tool for actual politics

Susana Soledad Ferreira Cazeneuve¹

Universidad Nacional de Asunción

<https://orcid.org/0009-0002-0844-4709>

Resumen

Este artículo busca reflexionar sobre la ciencia de datos en el ámbito de las ciencias políticas convirtiéndola en herramienta clave para la posterior toma de decisiones y reducir las desigualdades en la sociedad. La ciencia de datos es el conjunto de procesos multidisciplinarios para extraer datos a través de la intervención de ordenadores y analistas expertos generando conocimiento relevante. No obstante, la desinformación y manipulación de datos engañosos podrían dificultar la producción de conocimiento incrementando la brecha de desigualdad.

Palabras clave

Ciencia de datos, ciencias políticas, reducción de desigualdad.

Abstract

This article seeks to reflect on data science in the field of political science, turning it into a key tool for subsequent decision-making and reducing inequalities in society. Data science is the set of multidisciplinary processes to extract data through the intervention of computers and expert analysts, generating relevant knowledge. However, misinformation and manipulation of misleading data could hinder the production of knowledge, increasing the inequality gap.

Keywords

Data science, political science, inequality reduction.

¹ Universidad Nacional de Asunción. Facultad de Derecho de Ciencias Sociales. Escuela de Ciencias Políticas y Sociales. Cuadro de honor 2023. Correo Electrónico: sferreirac7@gmail.com

Introducción

Este artículo pretende reflexionar sobre las nuevas tecnologías en el mundo de la política, en particular la ciencia de datos y de su utilización como herramienta para las políticas públicas y toma de decisiones para reducir las desigualdades. A partir de las reflexiones de esta ciencia como herramienta para las ciencias políticas, su potencialidad podría generar resultados beneficiosos y así contribuir a la justicia social.

En la actualidad, vivimos en la era de los datos. Durante los últimos años, el mundo digital ha ido avanzado a gran escala hasta convertirse en un punto indispensable en la vida cotidiana del ser humano. Mientras que los ordenadores han disminuido su tamaño y su peso, los datos y su importancia han crecido cada vez más a lo largo de los años.

Según Cao (2017) la ciencia de datos abarca múltiples definiciones; primeramente, lo define como “la ciencia de los datos” o “el estudio de los datos”. De la misma manera, en su segunda definición, desde la perspectiva disciplinaria, aclara que la ciencia de datos se define como el nuevo campo interdisciplinario que sintetiza y construye a partir de estadística, informática, comunicación, management y sociología para estudiar datos y su ambiente, para transformarlos en decisiones siguiendo una metodología estructurada. Además, este mismo autor explica el largo proceso histórico que ha tenido la ciencia de datos. La primera vez que este término fue utilizado fue en el libro *Concise Survey of Computer Methods* publicado por el autor Peter Naur en el año 1974. Fue planteado por numerosos estadísticos desde entonces y fueron ellos los primeros en nombrarla de esta manera.

La ciencia de los datos en el mundo empresarial

Hoy en día, esta ciencia es utilizada para lograr cambios significativos dentro del mundo de los negocios y las grandes empresas. En efecto, su correcto uso puede ser muy atractiva ya que genera grandes beneficios. De esta manera, se distingue la metodología de las empresas con un antes y después del uso de la ciencia de datos para optimizar las ventas o vistas, como, por ejemplo, las empresas de Netflix y Amazon. Numerosos son las industrias que ya son usuarias de esta ciencia, entre

ellos los supermercados como Walmart, empresas petroleras como Shell, en la educación con IBM Watson, parques de atracciones como Walt Disney Parks & Resorts, empresas tecnológicas como Google y Microsoft, redes sociales como LinkedIn y X (antiguamente Twitter) y, dentro del gobierno estadounidense, para inteligencia y agencias de seguridad contra el terrorismo (Marr, 2016).

En primer lugar, uno de los elementos muy importante de la ciencia de datos es el Big Data o los macrodatos, en español. El Big Data o los macrodatos se refieren a grandes cantidades de datos, de gran volumen, variedad y velocidad. Estos datos fueron creados a partir de cada clic en el ordenador, cada mensaje en WhatsApp y cada foto tomada con un smartphone. En efecto, se describe la creación de un conjunto de datos complejos a partir de fuentes cotidianas como los mencionados anteriormente, además de los monitores de tráfico, las redes sociales y las compras en internet, y se usan como base para innovaciones tecnológicas tales como la optimización de rutas de viajes, la publicidad focalizada o los sistemas de recomendaciones de compras como lo menciona el estadista Spiegelhalter (2023). Además, todos estos datos están almacenados en distintas bases de datos, lo cual difiere de algunos años atrás, cuando eran más escasos y retrasaban el sistema. Es por ello que su estudio y análisis es prioridad para las grandes empresas y necesitan de los mejores expertos con conocimientos de alta calidad.

En segundo lugar, otro elemento de la ciencia de datos es el IoT o Internet de las Cosas; se refiere a la conexión de los artefactos cotidianos a un sistema de software, sin cable y a conexión remota. También existe esta conexión a niveles masivos; un ejemplo de ello es la conectividad de un ascensor, conectado con un sensor a través del IoT y un técnico remoto, en donde puede recibir indicaciones de cuando éste necesite de un mantenimiento antes de averiarse. Numerosas son las utilidades de esta ciencia para el progreso de un desarrollo social más justo. Junto con la Inteligencia Artificial, los autores Kharlamov, A. A., & Pilgun, M. (2024), publicaron un estudio multidisciplinario en donde se analizan inmensas cantidades de datos obtenidos de las redes sociales para valorar sus opiniones en tiempo real sobre el transporte público, además de crear algoritmos para un proyecto urbano sobre el sistema de transportación. Así, estos autores demuestran que, con la contribución de la tecnología y el riguroso análisis de los datos, se puede lograr el éxito de proyectos sociales.

Sin embargo, esta herramienta puede considerarse riesgosa para la opinión pública si se informan noticias falsas con datos engañosos. En efecto, si existen datos de inmensa cantidad, variedad y volumen, y estos son incorrectamente analizados y manipulados, pueden ser alterados para distorsionar sus resultados y, a su vez, transformar la realidad. Los datos falsos o engañosos son fraudes que atacan a usuarios individuales o a organizaciones enteras con el objeto de manipular datos. Para ejemplificar este hecho, se presentan las distintas noticias en donde los datos son el protagonista. En el 2021, en España, el Secretario General del Partido Popular (PP) Teodoro García publicó en redes sociales que el número de desempleados españoles corresponde a seis millones de personas. Este dato es falso, según el *fact-checking* (verificación de datos) de la página Newtral.es (2021) que cita a las instituciones españolas encargadas de contabilizar el desempleo. Estas instituciones son el Servicio Público de Empleo Estatal (SEPE) y la Encuesta de población activa (EPA) y, como menciona la página web, el número de personas desempleadas es de tres millones, no de seis millones de personas. De esta manera, se evidencia la utilización de datos para transmitir un mensaje erróneo a la opinión pública, lo cual supone un riesgo para la sociedad. Lo que parece ser informativo e importante, al final, solo es una imagen engañosa en donde los datos representan lo que se desea expresar y está muy lejos de la realidad. Además, dentro de los procesos electorales, también se difunden informaciones falsas utilizando datos falsificados.

Riesgos en procesos electorales

Según la UNESCO, en colaboración de la comunidad Chequeado (2023), grupo argentino especializado en verificación de datos, existen diez tipos de desinformación electoral común. Entre ellos, están el número de voto de personas fallecidas, el número de voto de personas no habilitadas, falsas encuestas y manipulación de votos para evitar votación. Todos estos datos irreales afectan a la sociedad generando distintos niveles de impacto. Para ejemplificar, Olivia Sohr, de Chequeado, presenta el primer nivel como “no dañino”, al cual la información y los datos no representan un riesgo sino una tergiversación de las cosas. También agrega el segundo nivel como “alejamiento de democracia”, que genera abuso de poder y desconfianza en la democracia. En este nivel, las personas pueden opinar que el proceso electoral y el

voto carecen de valor democrático. En el tercer nivel, ella menciona el daño económico de los individuos, de las empresas y del sistema. Dentro del cuarto nivel está la “interferencia en la democracia”, en donde una posible desinformación es causada por interferencias durante el proceso electoral y termine en una falencia en el sistema social de creencias y actitudes. Finalmente, en el último nivel, Sohr añade al “riesgo de vida” como impacto de la desinformación. Este último se trata del más extremo nivel, en donde la salud pública es afectada, existe una radicalización peligrosa o hay riesgo de violencia masiva. Así es como los datos en las manos incorrectas logran beneficiar a unos y perjudicar a otros. Fomentar la desigualdad con los datos es un hecho que ocurre hoy en día, engañando a las personas y creando barreras entre ellas.

Conclusiones

En la actualidad, la ciencia de datos es una ciencia multidisciplinaria computacional de gran importancia, con innovadoras herramientas y tecnologías como la Big Data y el Internet de las Cosas que van actualizándose a cada instante y desarrollándose aún mucho mejor con la ayuda de la Inteligencia Artificial. Si bien esta ciencia ya está incorporada en las ciencias naturales como en la medicina, también tiene mucho potencial en las ciencias sociales y ciencias políticas. Creando herramientas estadísticas con el apoyo de expertos matemáticos, analistas de datos, estadísticos y polítólogos, las políticas públicas podrían cubrir las necesidades que enfrenta la sociedad y así reducir las barreras de desigualdad.

En conclusión, la ciencia de datos como herramienta clave de las políticas públicas significaría un progreso para la población en su globalidad, en donde la tecnología es el protagonista principal y posible agente de cambio para la posterior toma de decisiones a través de los datos analizados bajo la lupa de expertos.

Referencias bibliográficas

Acurio Del Pino, S. (2016, november). *Delitos Informáticos: Generalidades*.

<https://repositorio.usam.ac.cr/xmlui/bitstream/handle/11506/2374/LEC%20DE%20200019%202016.pdf?sequence=1>

Cao, L. (2017). *Data Science: A Comprehensive Overview*. ACM Computing Surveys, 50(3), 1–42. <https://doi.org/10.1145/3076253>

Casanova, H. (2022). *Estadística y Ciencia de Datos ¿Qué hay de nuevo?* Actualidad Económica, 32(107), 55–75.

Chequeado, & UNESCO. (2023, August 3). *10 ejemplos de desinformación que circula en elecciones*. YouTube.

https://www.youtube.com/results?search_query=10+ejemplos+de+DESINFOR+MACI%C3%93N+que+circula+en+Elecciones.+

Kharlamov, A. A., & Pilgun, M. (2024). *Data Analytics for Predicting Situational Developments in Smart Cities: Assessing User Perceptions*. Sensors, 24(15), NA. <http://dx.doi.org.ezproxy-cicco.conacyt.gov.py/10.3390/s24154810>

Marr, B. (2016) *Big Data in Practice. How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results*. Wiley, Hoboken.
<https://doi.org/10.1002/9781119278825>

Meneses Rocha, M. E. (2018). *Grandes datos, grandes desafíos para las ciencias sociales*. Revista Mexicana de Sociología, 80(2), 415-444.
<https://doi.org/10.22201/iis.01882503p.2018.2.57723>

Naur, P. (1974) *Concise survey of computer methods*. Petrocelli Books

Newtral. (2021, December 24). #FACT-CHECK | *Las afirmaciones falsas y engañosas más repetidas por los políticos en 2021*. <https://www.newtral.es/factcheck-mentiras-repetidas-politicos-2021/20211224/>

Spiegelhalter, D. (2023). *El arte de la estadística*. Capitán Swing Libros.

UNESCO. (2023, November 16). *El desafío de la desinformación en procesos electorales*. Unesco.org. <https://www.unesco.org/es/articles/el-desafio-de-la-desinformacion-en-procesos-electorales>