

BIG DATA EN LAS CIENCIAS SOCIALES. UN ACERCAMIENTO NECESARIO

BIG DATA IN THE SOCIAL SCIENCES. A NECESSARY APPROACH

Eduardo O' Bourke Díaz

<https://orcid.org/0000-0002-2801-0409>
CEDEM, Universidad de La Habana, Cuba
obourke96@gmail.com

Rachel Lambert Correoso

<https://orcid.org/0000-0003-2976-8542>
CEDEM, Universidad de La Habana, Cuba
rachelambert99@gmail.com

Resumen

Las ciencias sociales constituyen un cuerpo de conocimiento dentro de sistemas sociales complejos en constante evolución. Es por ello que hoy existe la necesidad de integrar orgánicamente los saberes metodológicos, epistemológicos y prácticos en función de producir oportunamente lo necesario para el desarrollo de las diversas colectividades.

El Big Data específicamente es uno de estos términos a los cuales los medios de comunicación nos tienen acostumbrados; pero desde visiones científicas de actualidad se hace imprescindible una comprensión más completa de esta técnica. Efectivamente, Big Data no se refiere a una tecnología, sino a un conjunto de técnicas que se sirven de tecnologías para el trabajo con grandes bases de datos.

Desde la demografía se observan aportes notables que podrán suceder tras el vínculo de esta técnica puesta en función de estudios en dicho campo. Sin embargo, por el momento Latinoamérica se encuentra recién incursionando en este tipo de estudios.

Palabras clave: tecnologías, datos, conocimiento, estudio, aportes.



Artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Atribución-NoComercial 4.0 Internacional.

Abstract

The social sciences are a body of knowledge within complex social systems. That's why, there is a need to organically integrate what we know in terms of methodology, epistemology and practice in order to produce what is necessary to make real the development of communities.

"Big Data" is a term to which the media have accustomed us, but from scientific perspectives, a more complete understanding of this technique is essential. This is not a technology; it is a set of techniques the use technologies to work with large databases.

The demography is one of these sciences that can clearly make use of this. Now, Latin America is beginning to venture into studies that integrate the elements of Big Data and social sciences.

Keywords: *technologies, data, knowledge, study, contributions.*

INTRODUCCIÓN

El tiempo es cada vez más importante en términos científicos. La velocidad con la que se necesita definir patrones y situaciones en contextos sociales complejos en pos de tomas de decisiones diversas es de suma importancia hoy día. Se observan referencias constantes al empleo de las nuevas tecnologías casi en abstracto. Los métodos y técnicas al alcance de las comunidades productoras de conocimiento son de especial relevancia en un contexto de globalización y altos flujos de información (Wajcman, 2017).

Temas cruciales como la rentabilidad, la fiabilidad, la epistemología, los referentes y las metodologías o la viabilidad misma de las ciencias y los proyectos sociales se ven directamente afectados por los factores de acceso a bienes y servicios tecnológicos y el empleo que se logre de ello. Entonces, se impone conocer, al menos, las generalidades y potenciales aportes del Big Data como parte del ambiente científico técnico de la actualidad. Es vital que los investigadores en el área de las ciencias sociales y humanísticas en general presten atención a las posibilidades, fortalezas y debilidades de estas tecnologías. Ser críticos, realistas, oportunos y eficientes en cuanto a las capacidades que exigen estos tiempos es fundamental para mantener la vigencia y la eficiencia en múltiples áreas sociales.

Bien se entiende que cada aporte del presente contiene sus propias complejidades, costos y beneficios. Las tecnologías digitales proporcionan mayores velocidades de procesamiento de datos y análisis de información. Esto no indica que crean conocimiento. Hay distancias entre datos, información y conocimiento; pero esto no es un punto a profundizar en el presente artículo. Poder adoptar posturas científicas críticas en función de sacar partido a las oportunidades que ofrecen las tecnologías, necesariamente, parte de estudiarlas. Así, la siguiente revisión bibliográfica busca arrojar luces en estos sentidos. Una visión desde las ciencias sociales y en particular la demografía en su rol transversal de la actualidad.

1. Los estudios demográficos en la era digital. Latinoamérica como contexto

La demografía es una ciencia compleja y multidisciplinaria que se sirve de múltiples áreas del conocimiento en pos de generar nuevos saberes. Las miradas cuantitativas son fundamentales e históricamente típicas en los estudios de este campo. La razón fundamental yace en sus propias fuentes de información y objeto de estudio.

Es un acuerdo ampliamente aceptado que la demografía es *“una ciencia que tiene como finalidad el estudio de la población humana y que se ocupa de su dimensión, estructura, evolución y caracteres generales considerados fundamentalmente desde un punto de vista cuantitativo”* (International Union for the Scientific Study of Population, 2004, pág. 17). Por tanto, se sirve de los datos de diversas fuentes y bases de datos, a través de técnicas como: cuestionarios, inventarios, entrevistas, registros, censos, etc. Con ello impacta disímiles áreas dentro del devenir social de los territorios como: política, economía, salud pública, planificación urbana, etc. (Consejo Mexicano de las Ciencias Sociales, 2022).

Los paradigmas epistemológicos para el desarrollo de las ciencias guardan una relación que determina la práctica que se demanda de la especialidad, los tipos de datos considerados válidos y los criterios de verdad ponderados por los autores (Pérez, 1998). En la actualidad se observa una notable diversidad de modos y metodologías. Las nuevas técnicas propician alcances, mixturas y procedimientos impensables para los científicos de hace 20 años.

Estos impactos tecnológicos, metodológicos, procesales y sistémicos han llegado a todas las áreas del conocimiento. Las muestras de ello están a la mano. Por solamente mencionar del área geográfica latinoamericana, se observan trabajos de actualización técnica para abordar la demografía, actualizar procedimientos e infraestructuras científicas (Ruiz, 2019; Miró, 2009; Gomis-Balestreri, 2017; Peláez, 2014). En la actualidad la mayoría de estudios publicados emplean de alguna forma modos de gestión del conocimiento con mediación de tecnologías modernas.

Específicamente en Cuba constan las colaboraciones del Centro de Estudios Demográficos con el Centro de Estudios de Sistemas Complejos de la Facultad de Física de la Universidad de la Habana (Aja, Orbea, & Rodríguez, 2021; Aja & Rodríguez A. y Orbea, 2020). Una oportunidad de aprovechamiento de las técnicas de Big Data en función del análisis de múltiples factores dinámicos en un período de crisis epidemiológica por COVID-19.

Al revisar parte de la bibliografía se evidencia entonces, que la calidad, veracidad, velocidad, variedad y volumen de los datos que se procesan hoy son mayores que los que se podrían manejar hace 20 o 30 años (Escobar & Mercado, 2019). Es sabido que las tecnologías iniciaron sus impactos en la demogra-

fía desde los propios inicios de la informática como área del conocimiento aplicado. Esto se ha visto reforzado con la aparición de softwares como: ATLAS.ti, MAXqda, Mortpack, RUPEX, PRODEM, PRODEX, POPGROUP, SPSS entre otros tantos (Peláez, 2014) y la universalización del uso de ordenadores e internet.

El avance informático ha alcanzado a todas las áreas del conocimiento científico de alguna manera. Es sabido que la digitalización de los registros de identidad, natalidad, estado civil, etc. ha supuesto adelantos destacables en las formas de gestión jurídica, política y económica de los sistemas sociales del mundo. Variados trabajos dentro del contexto regional versan en torno a las formas de sacar provecho de ello (Ruiz, 2019; Gomis-Balestreri, 2017; Méndez, Sánchez, & Atencia, 2022).

Los países en vías de desarrollo enfrentan problemas de acceso tecnológico y dependencia en relación con otras naciones o grandes empresas tecnológicas. Es en los últimos años, que se produce impacto de elementos como el márketing digital sobre modos de gestión en esta zona geográfica (Méndez, Sánchez, & Atencia, 2022). Los servicios necesarios para establecer y regular las infraestructuras de progreso informático son imprescindibles para el ajuste funcional de cada nación a las tantas dinámicas de la actualidad (Cristina, 2020). Hoy dependen del acceso a internet, desde los sistemas internacionales de compra/venta, hasta la divulgación científica y el desenvolvimiento cotidiano de los sujetos comunes.

Es así que se entienden algunas de las complejidades de un contexto regional de importantes retos para las ciencias demográficas.

1.1. La Movilidad

Para estudiar la movilidad en particular es importante entender algunos elementos esenciales. El movimiento describe la cualidad de desplazamiento de un elemento en un lapso de tiempo determinado. Sin embargo, para la demografía hay movilidad social, que responde a la recomposición o cambios de atributos socio demográficos de los sujetos (soltero-casado/bachiller-universitario/ empleado-desempleado/etc.). Por su parte, el movimiento natural o vegetativo, responde a los nacimientos y defunciones en una población (CEDEM, 2022).

Los fines, las temporalidades, sentidos y motivaciones de los movimientos, pueden llevar a definir especificidades nominales al momento de conceptualizar cada evento dado. Es así que, un desplazamiento sobre territorio físico puede ser considerado: inmigración, emigración, turismo, desplazamiento laboral, movimiento cotidiano, viaje de corto plazo, cambio de residencia, etcétera (CEDEM, 2022). Es relevante definir con detalle el movimiento específico a estudiar en cada caso. Así, el objeto de estudio y el lenguaje a emplear quedan adecuadamente definidos.

Para el caso que ocupa al presente estudio es preciso hablar en términos de movimiento mecánico o espacial. Este es el que se refiere al tipo de movimiento físico relativo al traslado de individuos entre territorios. Más específicamente a la movilidad internacional (movimientos de individuos que implica traspasar fronteras entre países). Podría parecer una obviedad; pero es oportuno llamar la atención respecto a que no todo movimiento de este tipo implica migración. Más allá de las distinciones al respecto, se estima que, los viajeros son personas que traspasan límites geográficos.

2. Big Data

Plantear las potenciales relaciones entre el Big Data y la demografía debe partir de entender inicialmente los fundamentos de lo que “Big Data significa”. Empezar por las 3 “uves” de la programación es oportuno. Según la bibliografía son: variedad, volumen y velocidad (Finger, 2014). Antes de existir las condiciones de desarrollo actuales, los usuarios de las tecnologías debían elegir entre trabajar con alta variedad de datos, gran volumen o velocidad de procesamiento (Gil, 2016). Cada una de estas cualidades en el proceso excluiría a las otras dos con su existencia. Hoy es posible ser beneficiarios de sistemas en que estas tres características de procesamiento estén presentes.

Entonces, el big data precisamente es un sistema de procesamiento de datos a través de las tecnologías de la información, que se caracteriza por su variabilidad, velocidad y volumen. Especialmente, se ha empleado en el área del mercado y la economía. Ha tenido fuertes impactos en los modos de gestión empresarial a nivel global (Escobar & Mercado, 2019). Los sectores sociales se ven directamente afectados por los impactos de las diversas acciones que tienen lugar a través de estos procedimientos.

Dicho de otro modo, el big data es el contiguo tecnológico que permite tratar cantidades masivas de datos derivados de fuentes heterogéneas, con el objetivo de otorgarles alguna utilidad determinada (Gil, 2016). Es posible descubrir patrones de comportamiento de los clientes de una organización para crear publicidad dirigida y más efectiva (conocidos por todos, el escándalo de Cambridge Analítica y WikiLeaks¹), predecir tendencias económicas o descubrir relaciones antes desconocidas entre variables que puedan abrir las puertas a la innovación.

¹ Ambos ejemplos citados tuvieron repercusiones políticas de alta relevancia para el panorama internacional. Se utilizaron datos privados de miles de usuarios en pos de dirigir campañas personalizadas en función de guiar las intenciones de votos en el primer caso. Mientras, el segundo ejemplo tuvo que ver con filtraciones de espionaje cibernético a grandes cantidades de personas con fines estrictamente militares.

Algo relevante es que no se refiere a una tecnología en sí misma, sino un planteamiento metodológico, un paradigma funcional de trabajo dirigido a manipular eficientemente volúmenes muy grandes de datos (Marr, 2022). Esto es en esencia una herramienta extrapolable a las ciencias en general y ha arrojado evidencias del impacto y alcance que puede llegar a alcanzar en tiempos relativamente cortos.

La vida diaria se ve atravesada por el uso de las redes sociales y los conocidos algoritmos (muchos de los cuales poseen bases en el Big Data). Los modos de consumo, el acceso a contenidos diversos, lo que se compra y se vende son partes del entramado tecnológico que se ve presente entre la cotidianeidad. Es así que el impacto de este avance es posible de encontrar en casi todos lados.

Para el año 2000, un cuarto de toda la información mundial era almacenada en formato digital; el resto, en medios analógicos (papel, cintas, tarjetas perforadas, vinilos, etc.). A día de hoy más del 98% de toda nuestra información es digital (Kenneth & Mayer-Schöenberger, 2013).

Se observan ventajas para el procesamiento de la información, dificultades éticas y políticas, altos índices de dependencia y por tanto de riesgos sistémicos. Evidentemente, existen debates abiertos al respecto de cada una de dichas temáticas y estudiosos de muchas esferas se han adentrado en estos campos para realizar aportes y brindar soluciones y alternativas a estas cuestiones.

El llamado fundamental es a incorporar los conocimientos adquiridos al área de las políticas públicas y las ciencias aplicadas, en pos de la eficiencia y la toma de decisiones. Como reflexión, los sectores sociales deben invertir en tecnologías para el aprovechamiento de las técnicas actuales de manejo de datos (Escobar & Mercado, 2019).

Hoy se habla de tecnologías disruptivas (Ortiz, 2015; Marr, 2022; Gil, 2016). El término hace alusión al impacto directo y la fuerza con que modifican los comportamientos en sociedad. A pesar del carácter aparentemente negativo del vocablo, las conceptualizaciones negacionistas o contrarias al desarrollo no han alcanzado popularidad entre las comunidades científicas. Los aportes son innegables y las evidencias abundan al respecto. Así, por ejemplo, el fenómeno del internet de las cosas, el Cloud Computing o el Big Data son, en efecto, disruptivas y no por ello negativas. Sus alcances revolucionan la forma en que funciona el mundo.

Con estas tecnologías, los datos han devenido en un activo fuertemente cotizado. Al decir de Elena Gil (pág. 16): "Hoy en día se crean más datos que nunca antes en la historia"; recogerlos, almacenarlos y tratarlos es posible de formas eficientes y relativamente sencillas. Cada ciudadano activo socialmente aporta a esta realidad digital con las redes sociales, las transacciones bancarias, llamadas telefónicas y muchos gestos cotidianos que generan datos. El estudio de todo ello constituye fuentes de valor descomunal.

Las empresas tecnológicas ofrecen servicios de acceso gratuito a cambio de poder acceder a nuestros datos. Desde los juegos y las plataformas de pago hasta las más conocidas redes sociales y las agencias telefónicas (Marr, 2022). Tener los medios para utilizarlos abre las puertas a innumerables cantidades de fines.

Es importante partir de que en la actualidad hay más información almacenable (y que en muchos casos está siendo almacenada) a nuestro alrededor de lo que ha habido nunca en la historia. Está siendo utilizada para nuevos usos y de modos automatizados, manuales y mixtos todo el tiempo. Para dar una idea precisa, se evidencia que desde el inicio de la historia moderna hasta 2003 los humanos habíamos creado 5 exabytes (5 mil millones de gigabytes) de información. En contraste, para el año 2011 ya creábamos esa misma cantidad de información cada dos días al decir de Dave Turek (responsable superior de desarrollo de superordenadores de IBM) citado por Gil (2016).

Con el big data es viable transformar en información muchos aspectos de la vida que antes no eran sensibles de ser cuantificados, lo que dificultaba su estudio. Es el caso de los datos no estructurados (datos no-texto como fotografías, imágenes y ficheros de audio). Este fenómeno ha sido bautizado como dataficación (datafication en inglés) por la comunidad científica (Kenneth & Mayer-Schöenberger, 2013).

Esto indica que la localización geográfica ya es dataficada, primero con la invención de la longitud y la latitud, y en la actualidad con los sistemas de GPS vía satélite. Del mismo modo, las palabras son datos analizados por ordenadores mediante minería de datos. También controlados y constantemente supervisados por sistemas automatizados (Gil, 2016). Para dar una imagen más completa del alcance de estas tecnologías es posible observar que las amistades y gustos son codificados en forma de datos, a través de gráficos de relaciones de redes sociales, cantidad de interacciones y tiempo en pantalla de contenidos diversos.

Todo ello, se emplea hoy a través de un paradigma de funcionamiento ampliamente aceptado para describir gráficamente el Big Data. Consiste en un objeto similar a un embudo a través del cual pasan los datos para ser filtrados por "algoritmos" y desembocar en una nueva forma denominada "conocimiento".

2.1 La fortaleza y debilidad

Las tecnologías no actúan por sí solas sobre la realidad, tampoco están exentas de errores. Es por eso que, los usuarios, deben poseer capacidad crítica y de análisis suficiente para sacar de ellas el conocimiento. Al respecto, es posible dilucidar reflexiones oportunas. A pesar de que hay muchas aristas a explorar y debatir el análisis se centrará en una aparente fortaleza sin discusiones del Big Data: el volumen y las correlaciones entre variables.

Una variable es *“una característica, cualidad o propiedad observada que puede adquirir diferentes valores y es susceptible de ser cuantificada o medida en una investigación”* (Oyola-García, 2021, pág. 93). Históricamente los paradigmas cuantitativos han definido el trabajo investigativo a través del estudio de variables como célula fundamental para los análisis y la construcción del conocimiento. De esta forma se busca mayor acercamiento a la objetividad científica para propiciar acercamientos a la exactitud y validar la veracidad de los resultados.

Los datos en sí mismos no son conocimiento, tampoco verdades. Los conjuntos sensibles de ser ordenados e interpretados según sistemas cognoscitivos diversos no son ciencia, pero son elementales para construirla. Las relaciones entre elementos de la realidad son suficientemente complejas como para llevar a los investigadores a conclusiones equívocas. Es así que: entre mayor volumen de variables más relaciones ilegítimas son posibles de encontrar.

Una reflexión oportuna se puede extraer del texto de 2004 publicado por la revista *Pediatric and Perinatal Epidemiology* (Höfera, Przyrembel, & Verleger, 2004). Existe correlación entre las poblaciones de cigüeñas y la natalidad en zonas rurales del mundo. Este hecho resulta interesante y posee valor significativo, en tanto, se ha replicado en varias ocasiones con resultados similares.

Esto sustenta la *“Teoría de la Cigüeña”* Theory of the Stork (Höfera, Przyrembel, & Verleger, 2004). Lo cual inició con un estudio minucioso realizado por Paul Lazarsfeld (precursor acuñado entre los padres de la sociología cuantitativa) que dio forma matemática a una observación popular en la Inglaterra postindustrial. Cuantas más cigüeñas se avistaban en Londres, más nacimientos habría (Apellaniz, 2012).

Esto indica que es posible la existencia de causalidad entre un evento y otro. Sin embargo, hay una correlación positiva que en sí misma no expresa verdad alguna. ¿Dónde yace la explicación entonces?

Se está en presencia de una relación falsa y observable entre variables. El debate relacionado con la relación oculta es de utilidad científica. Se encontró que la agricultura y más específicamente las cosechas eran causales determinantes de los comportamientos encontrados entre las variables población de cigüeñas y natalidad (Gil, 2016).

Entonces, la criticidad es fundamental en contextos de análisis destinados a hacer ciencia con las técnicas de Big Data. Los resultados numéricos y relacionales poseen la cualidad de expresar eventos de la realidad que están interconectados de modos no visibles desde una misma perspectiva. La mirada del dato mismo no es suficiente.

Es imprescindible indagar respecto a las causas y mecanismos a través de los cuales se establecen relaciones entre los sucesos. El rigor científico es una

herramienta fundamental para estos casos. Es necesario entender que con Big Data los investigadores potencialmente se exponen a trabajar con cientos de variables y millones de interacciones y fluctuaciones en lapsos determinados de tiempo. Evidentemente las probabilidades de ruidos y correlaciones espurias son muy altas.

3. Demografía y Big Data en pos de estudiar la movilidad internacional en Cuba

Una vez visto todo lo anterior es relevante exponer que: el Centro de Estudios de Sistemas Complejos de La Facultad de Física de la Universidad de la Habana posee bases de datos proporcionadas por ETECSA. En estas constan registros de Location Area Updates (LAU)² de más de 6 millones de usuarios de telefonía móvil en Cuba. Solamente con esto es posible estimar la movilidad de grandes masas poblacionales.

Teniendo en cuenta que el único proveedor de servicios de telefonía móvil del país es la entidad que proporciona la base de datos con la que se propone trabajar; se está en presencia de la totalidad de usuarios activos dentro de territorio nacional. La población cubana actual es de 11 113 215 personas, según el Anuario Demográfico de 2021 publicado por la Oficina Nacional de Estadísticas e Información. En relación con esto, la cantidad de usuarios en la base de datos con la que trabaja el Centro de Estudios de Sistemas Complejos es superior a 6 millones, lo cual representa más del 50% de la población del país. Es evidente el carácter significativo de los potenciales resultados de investigaciones que empleen estos datos como base. Se prevén análisis generalizables de cara a las diversas dinámicas a estudiar.

Todo lo señalado son ventajas y oportunidades a explotar. Es posible articular miradas diversas y aplicar conocimientos de las ciencias sociales y exactas a estas grandes masas de datos. Ejemplos sobrados de la eficacia de estos enfoques existen en la historia reciente (CambridgeAnalytic y las controversias de Facebook, Google y las ventas de inmensas bases de datos, el propio funcionamiento de los algoritmos que personalizan la experiencia de navegación etc.). De aquí nace la inquietud de emplear las nuevas tecnologías y las habilidades al alcance en pos de la construcción de cuerpos de conocimiento.

El Centro de Estudios Demográficos de la Universidad de la Habana posee líneas de investigación vinculadas a la movilidad poblacional. De los desafíos

² Es el rastro de localización geográfica que deja un usuario de telefonía móvil al mantener su dispositivo conectado al servicio celular. Se obtiene como resultado del anclaje del teléfono a una de las antenas del proveedor del servicio. No es un método de localización preciso, dado que las distancias entre las antenas y los usuarios puede oscilar entre 1 metro y más de 1 000 metros de radio, al tomar la antena como punto central de una circunferencia imaginaria.

que resultan naturalmente complejos en el panorama actual están: la modelación de escenarios futuros y la urgencia temporal para los análisis y la producción de conocimientos. Con las tecnologías y su incorporación oportuna a los modos de hacer de las ciencias sociales es racional asegurar que las brechas entre las demandas sociales y la producción científica pueden ser estrechadas.

Aquí hablamos de que el alcance de las técnicas tradicionales de recopilación de datos, de cara a estudios demográficos, cada vez resulta más dilatado respecto a la velocidad de desarrollo de la sociedad. Dicho de otro modo, crear, aplicar y analizar encuestas, censos, entrevistas, formularios etc. demanda cierta cantidad de tiempo y recursos. Se entiende que los volúmenes de información a manejar en caso de buscar representatividad para poblaciones grandes requieren del trabajo coordinado de equipos de investigación y en función de las capacidades y recursos disponibles mayores o menores lapsos relativos. Para cuando los resultados están ya la realidad suele ser diferente.

Con estas problemáticas a la vista se propone una investigación de corte multidisciplinario. Se persigue entrelazar varias herramientas en función de propiciar técnicas potencialmente novedosas. Así se busca impactar en la obtención de resultados científicos de carácter generalizable en tiempos relativamente cortos. Específicamente se habla de la población cubana que viaja al extranjero.

Ya se ha visto que es posible medir, cuantificar, explicar, observar, predecir y actuar sobre comportamientos poblacionales con el empleo de sistemas automatizados de análisis de datos. Ahora queda abierta la posibilidad de aplicar esto al entorno nacional, tomando un contexto específico para probar y validar un algoritmo de acción concreto.

CONCLUSIONES

El Big Data no es una tecnología en sí misma. Es una técnica que echa mano de diversos elementos tecnológicos para servirse del análisis y procesamiento de grandes bases de datos. El desarrollo de los sistemas complejos sociales hace que esta herramienta sea de una importancia fundamental para las ciencias sociales. Existen diversas problemáticas al momento de encarar su empleo, entre las que destaca el volumen de variables con las que se suelen trabajar y las diversas dinámicas de correlación entre estas.

Las ciencias sociales y humanísticas deben servirse de las posibilidades que brinda el empleo de estos avances tecnológicos. Han sido poco empleados en Latinoamérica con fines de desarrollo científico; sin embargo, es sabido que grandes empresas tecnológicas hacen uso de estas técnicas con el fin de predecir, guiar o manipular comportamientos de grandes grupos poblacionales. Entonces, se plantea la existencia de todo un campo de investigación y evolu-

ción a la vista, que pasa por: la legalidad, el derecho, la regulación, el empleo, la ética, la oportunidad, la eficiencia etc.

El presente documento también demostró la voluntad de realizar un estudio en el contexto cubano del año 2022, que tiene a la vista el empleo de dichas tecnologías y pretende un acercamiento necesario y provechoso a las herramientas que se ofrecen desde la colaboración entre dos centros de estudios de la Universidad de La Habana.

REFERENCIAS BIBLIOGRÁFICAS

- Aja, A., & Rodríguez A. y Orbea, M. (2020). (2020). COVID-19, migración externa y desplazamientos territoriales en Cuba, una mirada deferente a la población residente en la capital del país. . *Novedades en Población*, 33-69.
- Aja, A., Orbea, M., & Rodríguez, A. (2021). COVID-19, migración internacional, movilidad territorial y fenómenos económicos y sociales asociados. El caso particular de La Habana. *Novedades en Población*, 154-184.
- Apellaniz, B. (3 de diciembre de 2012). *Naukas*. Obtenido de La confusión de las cigüeñas: <https://naukas.com/2012/12/03/la-confusion-de-las-ciguenas/>
- Burrieza-Galán, J. (2022). Estudio de la movilidad con la tecnología big data: posibilidades por explorar. *Papeles de economía española "infraestructuras terrestres, transporte y movilidad de personas"*(171).
- CEDEM. (2022). *Migración*. La Habana: CEDEM.
- Consejo Mexicano de las Ciencias Sociales. (20 de 12 de 2022). *Observatorio de las ciencias sociales*. Obtenido de <https://www.comesco.com/disciplinas/demografia#:~:text=Seg%C3%BAAn%20el%20diccionario%20demogr%C3%A1fico%20multiling%C3%BCe,desde%20un%20punto%20de%20vista>
- Cristina, I. R. (2020). *INNOVACION EN LA INVESTIGACIONES SOCIALES DE POSGRADO*. Obtenido de <https://www.researchgate.net/publication/341385246>
- Escobar, M., & Mercado, M. (2019). Big Data_ Un Analisis Documental de Su Uso y Aplicacion en el Contexto de la Era Digital. *28 Revista la Propiedad Inmaterial*.
- Finger, L. (2014). Recommendation Engines: The Reason Why We Love Big Data. *Forbes Tech*.
- Gil, E. (2016). *Big Data. Privacidad y protección de datos*. Madrid: Agencia Española de Protección de Datos.
- Gomis-Balestreri, M. (2017). Del gobierno electrónico al big data: la digitalización de la gestión pública en Colombia frente al control territorial. *Opera*, 25-53.

- Guadalupe, F., Vázquez, C., & Velázquez, M. (2019). STEM and Gender Gap in Latin America. *Revista de El Colegio de San Luis • Nueva época*, 137-158.
- Höfera, T., Przyrembel, H., & Verleger, S. (2004). New evidence for the Theory of the Stork. *Paediatric and Perinatal Epidemiology*, 88–92.
- International Union for the Scientific Study of Population. (2004). *Diccionario Demográfico Multilingüe (Segunda Edición)*. (G. A. Macció, Ed.) Liena, Bélgica: Ordina.
- Kenneth, N., & Mayer-Schöenberger, V. (2013). The Rise of Big data. How It's Changing the Way We Think About the World. *Foreign Affairs*, 92(3).
- López-Sala, A. (marzo de 2020). Avances en los estudios migratorios. Nuevos enfoques, nuevos instrumentos en el estudio de la movilidad internacional. *EMPIRIA. Revista de Metodología de Ciencias Sociales.*, 15-21. Recuperado el 21 de diciembre de 2022, de <https://www.redalyc.org/articulo.oa?id=297168989001>
- Marr, B. (2022). *Data Strategy (2nd edition). How to Profit from a World of Big Data, Analytics and Artificial Intelligence*. Milano: FrancoAngeli.
- Méndez, M., Sánchez, Z., & Atencia, J. (2022). *El marketing digital y el social media como estrategia competitiva del crecimiento empresarial. Caso latinoamericano*. Montería: Universidad Cooperativa de Colombia.
- Miró, C. (2009). La demografía en el siglo XXI en América Latina. En B. G. Castillo, & CLACSO (Ed.), *América Latina, población y desarrollo*. Siglo del Hombre .
- ONEI. (2021). *Anuario Demográfico de 2021*. La Habana: Oficina Nacional de Estadísticas e Información.
- Ortiz, D. (2015). *Metodología de análisis social en la era Big Data el retorno de la inversión (ROI) en redes sociales*. Salamanca: Universidad Pontificia de Salamanca.
- Oyola-García, E. (2021). La variable. *Revista del Cuerpo Médico del HNAAA*, 12(1), 90-93.
- Peláez, E. e. (2014). *Programas Informáticos para Proyecciones Demográficas*. La Habana: ALAP. Recuperado el 21 de diciembre de 2022, de <https://www.researchgate.net/publication/265594995>
- Pérez, C. (1998). *Sobre un concepto histórico de ciencia. De la epistemología actual dialéctica*. Santiago de Chile: LOM ediciones Ltda.
- Ruiz, Z. (2019). *Técnicas de aprendizaje automático aplicadas al procesamiento de información demográfica*. Alicante: Universidad de Alicante.
- Wajcman, J. (2017). *Esclavos del tiempo. Vidas aceleradas en la era del capitalismo digital*. Barcelona: Paidós.

CONFLICTO DE INTERESES

Los autores declaran originalidad de la presente investigación y ausencia de plagios en la misma. Además, aseguran que existe ausencia de conflicto de intereses en este artículo, por lo que se exime a la revista Ekotemas de cualquier reclamación al respecto.

CONTRIBUCIÓN AUTORAL

Eduardo O´Bourke Díaz: Conceptualización (idea, formulación y establecimiento de las metas generales de la investigación). Realización de la investigación, redacción del borrador inicial.

Rachel Lambert Correoso: Visualización (preparación de la visualización y presentación de los datos). Redacción, revisión y edición.