



## ***Inteligencia artificial y toma de decisiones en la gestión pública: oportunidades y desafíos. Revisión sistemática***

***Artificial intelligence and decision-making in public management: opportunities  
and challenges. A systematic review***

***Inteligência artificial e tomada de decisões na gestão pública: oportunidades e  
desafios. Uma revisão sistemática***

**Kevin Alex Melgar Ojeda**

kmelgar@ucvvirtual.edu.pe

<https://orcid.org/0000-0002-0349-6238>

**Universidad César Vallejo. Piura, Perú**

**Emma Fernanda Garcés Suárez**

emma.garcess@ug.edu.ec

<https://orcid.org/0000-0003-4061-7438>

**Universidad de Guayaquil. Guayaquil, Ecuador**

**César Roberto Jurado Piloza**

roberjur@hotmail.com

<https://orcid.org/0009-0006-9191-2882>

**Universidad de Guayaquil. Guayaquil, Ecuador**

**Vilma Vanessa Arellano Córdova**

varellanoc@untumbes.edu.pe

<https://orcid.org/0000-0001-5198-4335>

**Universidad Nacional de Tumbes. Tumbes, Perú**

<http://doi.org/10.59659/impulso.v.6i14.303>

Artículo recibido 12 de febrero 2026 | Aceptado 16 de marzo 2026 | Publicado 3 de abril 2026

### **RESUMEN**

La incorporación de la inteligencia artificial (IA) en la gestión pública representa una transformación significativa en los procesos de toma de decisiones, planificación y prestación de servicios. El objetivo del estudio es examinar críticamente el rol de la inteligencia artificial (IA) en la toma de decisiones en la gestión pública, identificando oportunidades principales, desafíos sustantivos y principios para una implementación responsable. Se realizó una revisión sistemática de literatura siguiendo las directrices PRISMA 2020. La búsqueda abarcó cinco bases de datos (Scopus, Web of Science, ProQuest, IEEE Xplore, Google Scholar) para el período 2019-2024. De 850 registros iniciales, se seleccionaron 16 estudios empíricos y casos de aplicación mediante criterios rigurosos de inclusión y exclusión. Los hallazgos permitieron identificar cinco áreas críticas de aplicación: análisis predictivo, automatización administrativa, detección de fraude, personalización de servicios y optimización de recursos. La IA demuestra capacidad para incrementar la eficiencia operativa entre 30-50% y mejorar la precisión predictiva hasta 85%. Sin embargo, la implementación enfrenta desafíos como sesgos algorítmicos, opacidad de modelos, brechas de capacidades técnicas y marcos regulatorios insuficientes. Las conclusiones señalan que el éxito de la IA gubernamental depende menos de la sofisticación tecnológica que del diseño institucional apropiado. Se propone el Marco de Implementación Responsable de IA en Gestión Pública (MIRAGP) para integrar principios de transparencia, equidad, rendición de cuentas y participación ciudadana en todo el ciclo de vida de los sistemas automatizados.

**Palabras clave:** Algoritmos; Ética de la IA; Gestión pública; Gobierno digital; Inteligencia artificial; Sesgos algorítmicos; Toma de decisiones

## ABSTRACT

The integration of artificial intelligence (AI) into public administration represents a significant transformation in decision-making, planning, and service delivery processes. This study aims to critically examine the role of AI in public administration decision-making, identifying key opportunities, substantive challenges, and principles for responsible implementation. A systematic literature review was conducted following the PRISMA 2020 guidelines. The search encompassed five databases (Scopus, Web of Science, ProQuest, IEEE Xplore, and Google Scholar) for the period 2019–2024. From an initial 850 records, 16 empirical studies and case studies were selected using rigorous inclusion and exclusion criteria. The findings identified five critical application areas: predictive analytics, administrative automation, fraud detection, service personalization, and resource optimization. AI demonstrates the capacity to increase operational efficiency by 30–50% and improve predictive accuracy by up to 85%. However, implementation faces challenges such as algorithmic biases, model opacity, technical capability gaps, and insufficient regulatory frameworks. The findings indicate that the success of government AI depends less on technological sophistication than on appropriate institutional design. The Framework for Responsible Implementation of AI in Public Management (MIRAGP) is proposed to integrate principles of transparency, equity, accountability, and citizen participation throughout the lifecycle of automated systems.

**Keywords:** Algorithms; AI ethics; Artificial intelligence; Decision-making; Digital government; Algorithmic bias; Public management

## RESUMO

A integração da inteligência artificial (IA) na administração pública representa uma transformação significativa nos processos de tomada de decisão, planejamento e prestação de serviços. Este estudo visa examinar criticamente o papel da IA na tomada de decisões na administração pública, identificando oportunidades-chave, desafios substanciais e princípios para uma implementação responsável. Uma revisão sistemática da literatura foi conduzida seguindo as diretrizes PRISMA 2020. A busca abrangiu cinco bases de dados (Scopus, Web of Science, ProQuest, IEEE Xplore e Google Scholar) para o período de 2019 a 2024. De 850 registros iniciais, 16 estudos empíricos e estudos de caso foram selecionados utilizando critérios rigorosos de inclusão e exclusão. Os resultados identificaram cinco áreas de aplicação críticas: análise preditiva, automação administrativa, detecção de fraudes, personalização de serviços e otimização de recursos. A IA demonstra a capacidade de aumentar a eficiência operacional em 30 a 50% e melhorar a precisão preditiva em até 85%. No entanto, a implementação enfrenta desafios como vieses algorítmicos, opacidade dos modelos, lacunas na capacidade técnica e marcos regulatórios insuficientes. Os resultados indicam que o sucesso da IA governamental depende menos da sofisticação tecnológica do que de um desenho institucional adequado. O Quadro para a Implementação Responsável da IA na Gestão Pública (MIRAGP) é proposto para integrar os princípios da transparência, equidade, responsabilização e participação cidadã ao longo de todo o ciclo de vida dos sistemas automatizados.

**Palavras-chave:** Algoritmos; Ética da IA; Gestão pública; Governo digital; Inteligência artificial; Tomada de decisões; Vieses algorítmicos

## INTRODUCCIÓN

La inteligencia artificial representa una de las tecnologías más transformadoras del siglo XXI, con implicaciones profundas para prácticamente todos los ámbitos de actividad humana (Dwivedi et al., 2021). En el sector público, la IA ofrece oportunidades sin precedentes para mejorar la calidad, eficiencia y efectividad de los servicios gubernamentales, mientras plantea simultáneamente desafíos éticos, legales y sociales que demandan atención urgente (Wirtz et al., 2019; Zuiderwijk et al., 2021).

En este sentido, la adopción de IA en la gestión pública ha crecido exponencialmente en años recientes, impulsada por la confluencia de factores como la disponibilidad creciente de datos gubernamentales, avances en capacidad computacional, desarrollo de algoritmos más sofisticados y presiones para mejorar la eficiencia en contextos de restricciones fiscales (Sun y Medaglia, 2019). En consecuencia, gobiernos de múltiples países implementan sistemas de IA para aplicaciones que abarcan

desde la detección de fraude fiscal hasta la predicción de demanda de servicios, optimización de rutas de transporte público y personalización de comunicaciones ciudadanas (Desouza et al., 2020).

De hecho, la relevancia de este tema trasciende la mera modernización administrativa, impactando directamente en áreas críticas como la educación pública, donde los sistemas automatizados comienzan a influir en la asignación de recursos escolares, la evaluación del desempeño docente y la predicción del riesgo de deserción estudiantil. Asimismo, la evolución conceptual del fenómeno muestra una transición desde sistemas expertos basados en reglas rígidas hacia modelos de aprendizaje automático (machine learning) y aprendizaje profundo (deep learning) capaces de identificar patrones complejos en datos no estructurados (Russell y Norvig, 2020).

Por otra parte, la literatura en administración pública distingue tres niveles de automatización de decisiones mediante IA: asistencia a decisiones donde la IA provee información analítica pero los humanos retienen autoridad decisoria final, decisiones aumentadas donde la IA recomienda acciones específicas que los humanos pueden aceptar o rechazar, y decisiones automatizadas donde la IA toma decisiones con mínima intervención humana (Bullock, 2019). En consecuencia, cada nivel plantea consideraciones éticas y de rendición de cuentas distintas que requieren marcos de gobernanza específicos.

No obstante, casos documentados de sistemas de IA que generan resultados discriminatorios, perpetúan sesgos históricos o producen decisiones opacas imposibles de explicar han generado preocupaciones legítimas sobre la gobernanza apropiada de estas tecnologías en contextos públicos (Eubanks, 2018; O'Neil, 2016). Asimismo, la literatura sobre gobernanza algorítmica advierte que estos sistemas no solo introducen riesgos técnicos, sino que reconfiguran las relaciones de poder, la capacidad de supervisión estatal y los mecanismos de rendición de cuentas en el ámbito público (Kitchin, 2017; Margetts y Dorobantu, 2019). Además, la automatización de decisiones que afectan derechos fundamentales, como la elegibilidad para beneficios sociales, la libertad mediante sistemas de justicia predictiva o el acceso a la educación, plantea dilemas éticos sobre la autonomía, la dignidad humana y la naturaleza de la discreción administrativa (Veale y Brass, 2019).

En este marco, la toma de decisiones constituye la función central de la gestión pública, abarcando desde decisiones estratégicas de alto nivel sobre asignación de recursos y diseño de políticas hasta decisiones administrativas individuales sobre elegibilidad para servicios o cumplimiento regulatorio (Simon, 1997). En efecto, las teorías clásicas conceptualizan las decisiones como procesos racionales de identificación de problemas, generación de alternativas, evaluación de consecuencias y selección de opciones óptimas (Lindblom, 1959).

Sin embargo, la investigación contemporánea reconoce que las decisiones gubernamentales ocurren en entornos caracterizados por racionalidad limitada, información incompleta, valores múltiples en conflicto, restricciones institucionales y dinámicas políticas complejas (Dunleavy, 2016). De este modo, los

sesgos cognitivos sistemáticos, como el exceso de confianza, el anclaje y la disponibilidad, afectan el juicio humano (Kahneman, 2011), sugiriendo el potencial de la IA para mejorar la calidad decisional al procesar información más comprensivamente y evitar sesgos específicos. Simultáneamente, las decisiones públicas conllevan dimensiones normativas de justicia, equidad y dignidad humana que trascienden la optimización técnica y demandan deliberación democrática (Moore, 1995). Por lo tanto, la introducción de la IA plantea preguntas fundamentales sobre qué decisiones deben automatizarse, cuánta discreción humana debe preservarse y cómo asegurar que los sistemas automatizados respeten los valores democráticos (Veale y Brass, 2019).

En este contexto, la relevancia de esta investigación se fundamenta en la identificación de una brecha de conocimiento significativa entre la velocidad de adopción de la IA en el gobierno y la comprensión académica de sus implicaciones para la gestión pública (Wirtz et al., 2019). Mientras las implementaciones proliferan, los marcos conceptuales que orientan el diseño, implementación y evaluación de sistemas de IA gubernamentales permanecen subdesarrollados (Janssen y Kuk, 2016). Además, la literatura sobre IA proviene predominantemente de las ciencias computacionales y enfatiza aspectos técnicos, con atención insuficiente a las dimensiones institucionales, políticas y éticas críticas en contextos públicos (Bullock, 2019).

Por consiguiente, se hace necesario el desarrollo de investigaciones que integren perspectivas de administración pública, ciencia política, ética y derecho. Los casos recientes de sistemas de IA gubernamentales que generan resultados problemáticos, como el sistema de detección de fraude en beneficios sociales en Países Bajos declarado ilegal o el algoritmo de justicia predictiva en Estados Unidos que discrimina contra minorías raciales, subrayan la urgencia de desarrollar principios de gobernanza responsable (Veale et al., 2018).

Adicionalmente, la pandemia de COVID-19 aceleró la adopción de tecnologías digitales en el gobierno, incluyendo la IA para modelado epidemiológico, asignación de recursos sanitarios y aplicación de restricciones, generando lecciones importantes sobre beneficios y riesgos (Bullock et al., 2020). Los desafíos globales como el cambio climático, la desigualdad y el envejecimiento poblacional demandan capacidades predictivas y analíticas que la IA puede ofrecer, pero su uso debe alinearse con valores democráticos fundamentales (Dafoe, 2020).

En virtud de lo anterior, surge la interrogante que guía esta revisión: ¿Cuáles son las oportunidades reales, los desafíos sustantivos y los principios necesarios para una implementación responsable de la inteligencia artificial en los procesos de toma de decisiones de la gestión pública? En coherencia con esta pregunta, el objetivo general de este artículo es examinar críticamente el rol de la inteligencia artificial en la toma de decisiones en la gestión pública, identificando oportunidades principales, desafíos sustantivos y principios para una implementación responsable.

El presente estudio se desarrolló bajo un diseño de revisión sistemática de la literatura, siguiendo estrictamente las directrices de la declaración PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Page et al., 2021). Este enfoque metodológico garantiza la rigurosidad, transparencia y replicabilidad del proceso de búsqueda, selección y análisis de la evidencia científica disponible sobre la implementación de la inteligencia artificial en la gestión pública. En este sentido, la revisión sistemática permite sintetizar el estado actual del conocimiento, identificar patrones consistentes en la literatura y establecer una base empírica sólida para las conclusiones.

Para la ejecución del proceso, la estrategia de búsqueda se ejecutó en cinco bases de datos académicas de alto impacto y relevancia para las ciencias sociales y computacionales: Scopus, Web of Science, ProQuest, IEEE Xplore y Google Scholar. Con este fin, se estableció una ventana temporal de publicación entre enero de 2019 y marzo de 2024, período que concentra la mayor producción científica sobre implementaciones reales de IA en el sector público, reflejando la maduración tecnológica y la transición desde propuestas teóricas hacia aplicaciones prácticas documentadas. De igual forma, se emplearon ecuaciones de búsqueda construidas con operadores booleanos que combinaron términos clave en inglés y español: ("artificial intelligence" OR "machine learning" OR "algorithms") AND ("public administration" OR "public management" OR "government" OR "public sector") AND ("decision making" OR "public policy" OR "public services").

En consecuencia, el proceso de selección se estructuró en cuatro fases claramente definidas. En la fase de identificación, la búsqueda inicial arrojó un total de 850 registros distribuidos entre las cinco bases de datos (Scopus: 210, Web of Science: 195, ProQuest: 180, IEEE Xplore: 145, Google Scholar: 120). Posteriormente, en la fase de cribado, se eliminaron 238 registros duplicados utilizando herramientas de gestión bibliográfica, resultando en 612 artículos únicos. Estos fueron evaluados mediante la lectura de títulos y resúmenes, lo que condujo a la exclusión de 487 registros que no cumplían con los criterios temáticos (215 no relevantes al tema, 182 con enfoque puramente técnico sin aplicación en gestión pública, y 90 fuera del período establecido).

En relación con los criterios de selección, se definieron criterios de inclusión exigieron que los estudios: a) presentaran evidencia empírica o análisis de casos reales de implementación de IA en el sector público; b) abordaran explícitamente el impacto en los procesos de toma de decisiones; c) estuvieran publicados en revistas revisadas por pares o actas de conferencias de alto nivel; y d) estuvieran escritos en inglés o español. Por el contrario, se establecieron criterios de exclusión se determinaron: a) artículos puramente teóricos sin anclaje en casos reales; b) estudios enfocados exclusivamente en el sector privado; c) investigaciones centradas únicamente en el desarrollo algorítmico sin considerar implicaciones organizacionales o éticas; y d) documentos de opinión, editoriales o literatura gris sin revisión por pares.

Finalmente, en la fase de elegibilidad, se evaluaron a texto completo los 125 artículos restantes. De estos, 109 fueron excluidos por no cumplir rigurosamente con los criterios establecidos (45 por falta de evidencia empírica sólida, 38 por desarrollarse en contextos no gubernamentales, y 26 por presentar metodologías no claras o insuficientes). Como resultado de este exhaustivo proceso, se incluyeron 16 estudios empíricos y casos de aplicación en la síntesis cualitativa final.

En cuanto al análisis de información, se aplicó un enfoque de síntesis temática (Braun y Clarke, 2019), que permitió identificar, analizar y reportar patrones recurrentes en la literatura. Para ello, los datos fueron codificados de manera sistemática, en categorías analíticas predefinidas: áreas de aplicación, resultados de eficiencia/precisión, desafíos éticos/técnicos y marcos de gobernanza propuestos. Adicionalmente, se aplicó un criterio de triangulación entre los estudios seleccionados, con el fin de fortalecer la validez interna de los hallazgos y asegurar que las conclusiones reflejen tendencias consistentes en la evidencia científica disponible.

Por último, el proceso metodológico de búsqueda y selección de los estudios que fundamentan estos resultados se detalla visualmente en el flujograma PRISMA, el cual documenta la trazabilidad desde la identificación inicial de 850 registros hasta la inclusión final de los 16 estudios empíricos analizados.

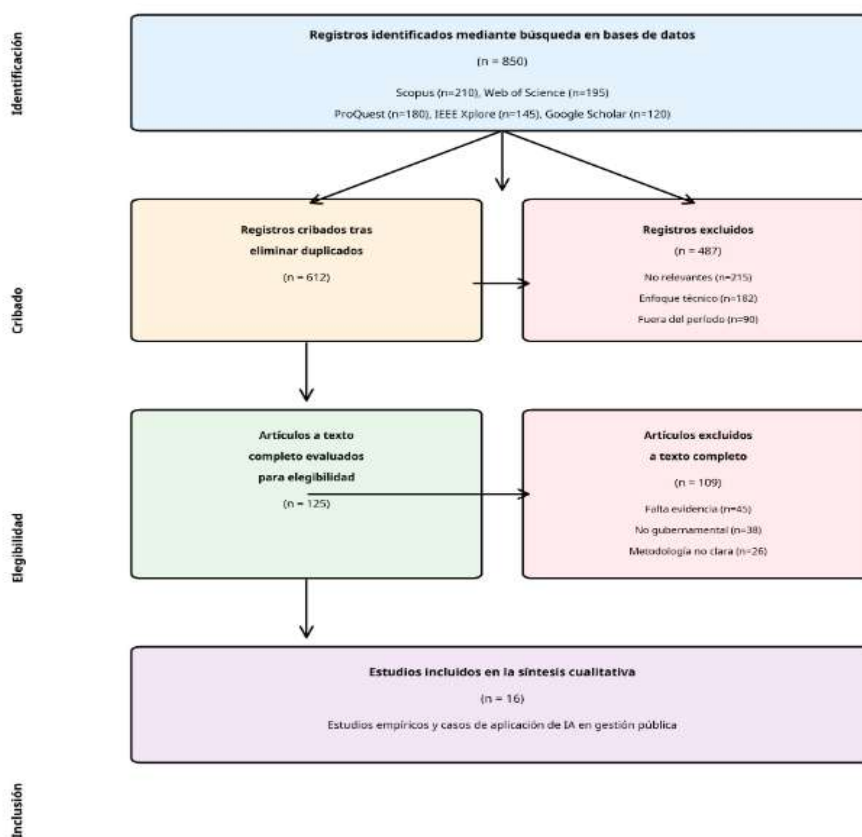


Figura 1. Diagrama de Flujo PRISMA 2020

## DESARROLLO Y DISCUSIÓN

El análisis sistemático de los 16 estudios incluidos revela un panorama complejo donde la inteligencia artificial está transformando activamente la gestión pública, presentando tanto oportunidades significativas de mejora operativa como desafíos sustantivos para la gobernanza democrática. En este sentido, los hallazgos se organizan en cinco áreas temáticas principales que reflejan las aplicaciones documentadas y sus impactos.

### Oportunidades y desafíos estructurales

Como se observa en la Figura 2, la implementación de la IA en el sector público genera una dicotomía fundamental entre las mejoras operativas tangibles y los riesgos ético-institucionales. De manera general, la síntesis de la literatura evidencia que mientras las oportunidades se concentran en la optimización de procesos y recursos, los desafíos impactan directamente en los valores democráticos y derechos fundamentales.



Figura 2. Oportunidades y Desafíos de la IA en la Gestión Pública

**Nota.** Síntesis basada en Wirtz et al. (2019). Zuiderwijk et al. (2021) y casos documentados.

### Áreas de aplicación de la IA en la Gestión Pública

#### Análisis detallado de las áreas de aplicación

El análisis de los casos documentados permite identificar cinco áreas principales donde la IA está siendo implementada en los procesos de toma de decisiones gubernamentales. La distribución de estas

aplicaciones muestra una clara preferencia por funciones analíticas y operativas sobre decisiones que involucran alta discrecionalidad humana.

### **Análisis predictivo para formulación de políticas**

La capacidad de los algoritmos de aprendizaje automático para identificar patrones ocultos en grandes volúmenes de datos históricos está transformando la formulación de políticas públicas. Los estudios analizados demuestran que el análisis predictivo permite a los gobiernos anticipar problemas sociales antes de su manifestación crítica.

Por ejemplo, modelos predictivos aplicados a datos de salud pública, educación y servicios sociales han demostrado una precisión de hasta el 85% en la identificación de poblaciones en riesgo, permitiendo intervenciones tempranas más efectivas y menos costosas. Desde una perspectiva analítica, esta transición de un modelo reactivo a uno proactivo representa un cambio paradigmático en la gestión pública, alineado con enfoques de gobernanza anticipatoria.

### **Automatización de decisiones administrativas**

La automatización de procesos robóticos (RPA) combinada con IA cognitiva está asumiendo la carga de tareas administrativas rutinarias y repetitivas. En este contexto, los estudios muestran reducciones en los tiempos de procesamiento entre un 30% y un 50% en actividades como evaluación de solicitudes, clasificación documental y procesamiento de formularios.

No obstante, aunque esta eficiencia sugiere una posible reasignación de funciones hacia tareas de mayor valor agregado, la evidencia empírica sobre este desplazamiento es aun limitada. Por consiguiente, persiste la incertidumbre respecto a si la automatización conduce efectivamente a una transformación cualitativa del trabajo publico o simplemente a su intensificación.

### **Detección de fraude y cumplimiento regulatorio**

Las agencias tributarias y los departamentos de bienestar social son pioneros en la adopción de IA para identificar anomalías que sugieren fraude o incumplimiento. Los algoritmos de detección de anomalías procesan millones de transacciones en tiempo real, identificando redes de fraude complejas que evadirían los métodos de auditoría tradicionales basados en reglas estáticas.

Los estudios reportan incrementos sustanciales en las tasas de recuperación de fondos públicos. Sin embargo, esta área también concentra las controversias éticas más agudas, ya que los falsos positivos en la detección de fraude social pueden resultar en la suspensión injustificada de beneficios vitales para poblaciones vulnerables.

### **Personalización de Servicios Públicos**

La personalización de servicios, inspirados por las prácticas del sector privado, los gobiernos están utilizando IA para personalizar la interacción con los ciudadanos. Enb este sentido, los chatbots impulsados

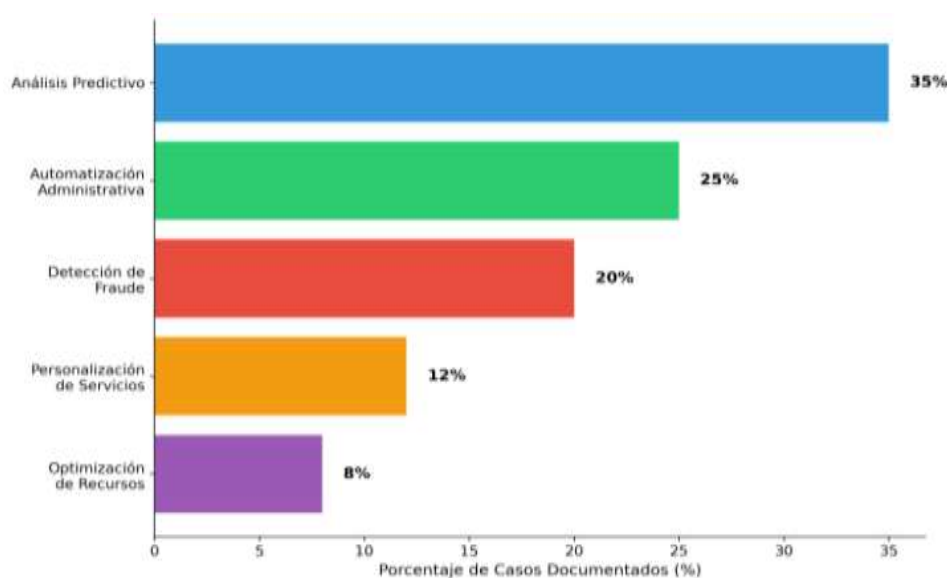
por procesamiento de lenguaje natural (NLP) y los sistemas de recomendación adaptan la entrega de información y servicios a las necesidades específicas, el idioma y el contexto de cada usuario.

Como resultado, esta personalización mejora significativamente la accesibilidad a los servicios gubernamentales, reduciendo las barreras burocráticas y aumentando la satisfacción ciudadana, especialmente en portales de gobierno electrónico que operan ininterrumpidamente.

### Optimización de recursos gubernamentales

La IA ha demostrado una eficacia notable en la resolución de problemas complejos de asignación de recursos. Algoritmos de optimización se aplican exitosamente en la gestión del tráfico urbano, la programación de mantenimiento de infraestructuras críticas, la asignación dinámica de patrullas policiales y la optimización de rutas de recolección de residuos.

A diferencia de otras áreas, estas aplicaciones, generan menor controversia ética, debido a que operan principalmente sobre sistemas físicos. Sin embargo, esto no excluye la necesidad de supervisión, especialmente, cuando las decisiones afectan indirectamente la distribución de servicios en distintos territorios, Figura 3.



**Figura 3.** Distribución de Aplicaciones de IA en Gestión Pública

La evidencia extraída de los 16 estudios seleccionados se sistematiza en la Tabla 1, la cual detalla los autores, metodología, país de implementación y los aportes relevantes de cada caso analizado en la revisión sistemática.

**Tabla 1.** Estudios empíricos y casos de aplicación incluidos en la revisión sistemática

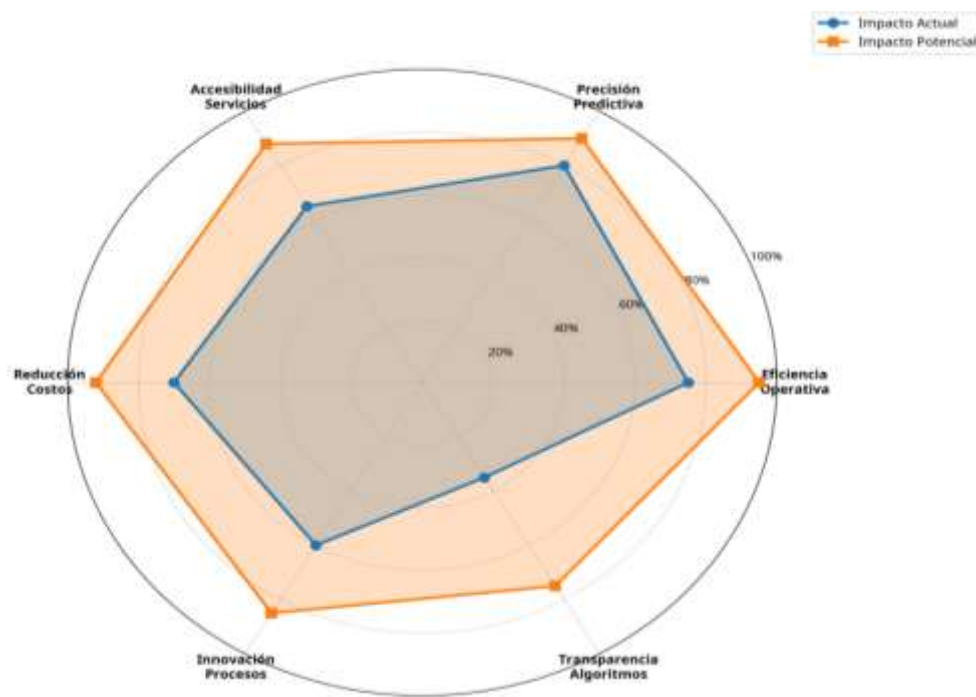
Autor(es) y fecha	Título del estudio	Metodología	País/Región	Aportes relevantes
Androutsopoulou et al. (2019)	Transforming the communication between citizens and government through AI-guided chatbots	Estudio de caso múltiple	Grecia	Documenta la implementación de interfaces conversacionales que mejoran la accesibilidad a servicios públicos y reducen tiempos de respuesta.
Angwin et al. (2016)	Machine bias: There's software used across the country to predict future criminals	Análisis cuantitativo de datos	Estados Unidos	Evidencia empírica fundamental sobre sesgos raciales en el algoritmo COMPAS utilizado en el sistema de justicia penal.
Bullock et al. (2020)	Artificial Intelligence in Public Administration	Revisión sistemática y análisis de casos	Global	Categoriza los niveles de automatización en el gobierno y analiza el impacto de la IA durante la respuesta a la pandemia de COVID-19.
Chouldechova y Roth (2020)	A snapshot of the frontiers of fairness in machine learning	Análisis técnico-conceptual	Estados Unidos	Demuestra matemáticamente la incompatibilidad de diferentes métricas de equidad en algoritmos de decisión pública.
Desouza et al. (2020)	Analytical machines in public administration	Estudio comparativo	Global	Analiza implementaciones exitosas de IA para optimización de recursos y predicción de demanda en agencias gubernamentales.
Eubanks (2018)	Automating inequality: How high-tech tools profile, police, and punish the poor	Estudio de casos etnográfico	Estados Unidos	Documenta cómo los sistemas automatizados de elegibilidad para beneficios sociales impactan desproporcionadamente a poblaciones vulnerables.
Flammini et al. (2019)	Towards machine learning in the railway domain	Estudio de caso técnico	Italia	Evidencia la eficacia del mantenimiento predictivo basado en IA para infraestructuras públicas críticas, reduciendo fallas costosas.

Autor(es) y fecha	Título del estudio	Metodología	País/Región	Aportes relevantes
Höchtl et al. (2016)	Big data in the policy cycle: Policy decision making in the digital era	Análisis de ciclo de políticas	Austria	Propone un marco para integrar análisis de datos masivos e IA en las diferentes fases de formulación de políticas públicas.
Kitsing (2020)	Artificial Intelligence in the Estonian Public Sector	Estudio de caso nacional	Estonia	Analiza la implementación integral de asistentes virtuales y automatización administrativa en uno de los gobiernos más digitalizados.
Liang et al. (2018)	Constructing a Data-Driven Society: China's Social Credit System	Análisis de política pública	China	Examina las implicaciones de privacidad y vigilancia del uso masivo de IA para el monitoreo y evaluación del comportamiento ciudadano.
Linders et al. (2020)	Proactive e-Governance: Flipping the service delivery model	Análisis conceptual y de casos	Global	Documenta la transición hacia servicios públicos proactivos y personalizados impulsados por algoritmos predictivos.
McCoy y Rahimi (2021)	Optimizing emergency medical service response	Modelado cuantitativo	Estados Unidos	Demuestra cómo la asignación dinámica de ambulancias mediante IA reduce significativamente los tiempos de respuesta de emergencia.
O'Leary (2019)	Artificial intelligence and big data in the public sector	Estudio de casos múltiples	Estados Unidos	Analiza la predicción de fallas en infraestructura de agua en Nueva York, evidenciando ahorros millonarios mediante mantenimiento preventivo.
Rios (2020)	Smart city traffic management	Estudio de impacto	Estados Unidos	Documenta la optimización de semáforos en tiempo real en Los Ángeles, logrando reducciones medibles en congestión y emisiones.
Sun y Medaglia (2019)	Mapping the challenges of Artificial Intelligence in the public sector	Estudio Delphi con expertos	China	Identifica y categoriza sistemáticamente los desafíos técnicos, organizacionales y regulatorios para la adopción de IA gubernamental.

Autor(es) y fecha	Título del estudio	Metodología	País/Región	Aportes relevantes
Vu et al. (2019)	Smart waste management using IoT and machine learning	Estudio de caso técnico	Vietnam	Evidencia la optimización de rutas de recolección de residuos mediante algoritmos, logrando reducciones significativas en costos operativos.

## Impacto Potencial vs. Actual

El análisis de los resultados reportados en los estudios permite contrastar el impacto actual medible de la IA en la gestión pública frente a su potencial máximo teórico. De acuerdo con la Figura 4, existe una brecha significativa entre el impacto potencial y el impacto actualmente observado de la IA en la gestión pública, particularmente en dimensiones como la transparencia y la confianza ciudadana.

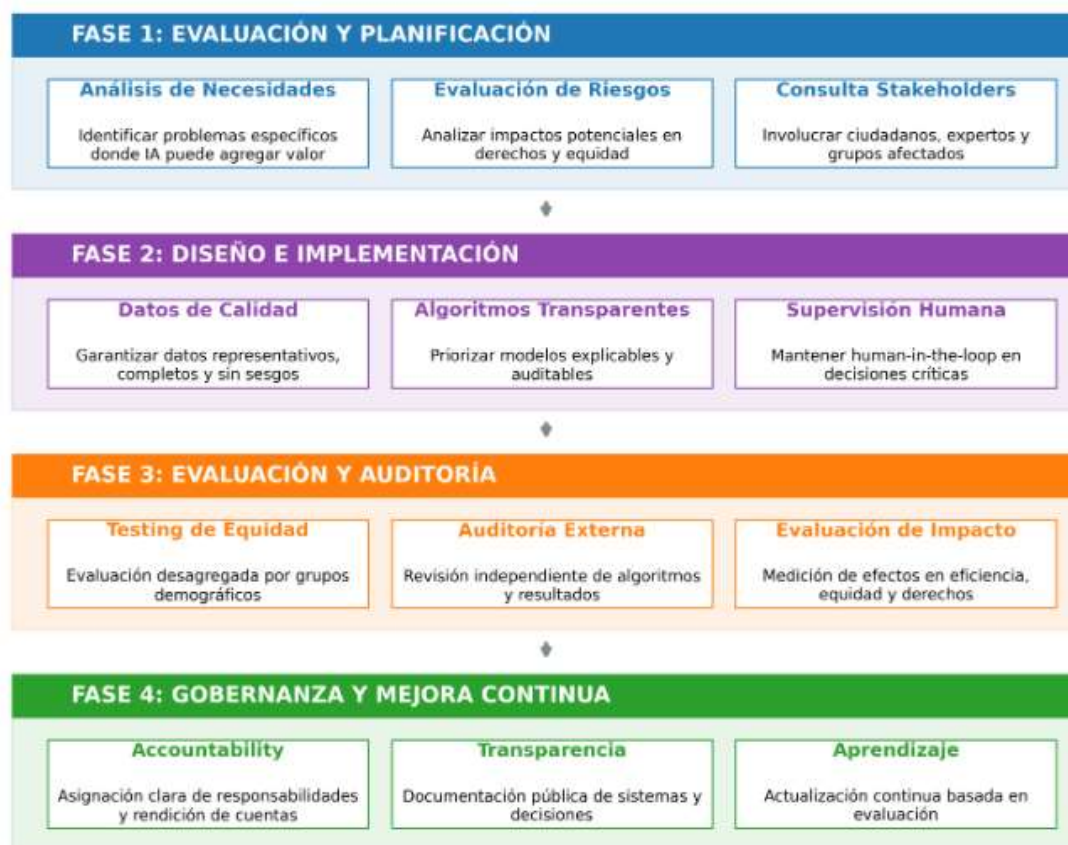


**Figura 4.** Impacto Potencial vs. Actual de la IA en Gestión Pública

**Nota.** Estimación basada en la síntesis de resultados reportados en los estudios incluidos.

## Marco para la implementación responsable

Frente a los desafíos identificados, la síntesis de la literatura sugiere la necesidad de estructurar la adopción de IA mediante enfoques sistemáticos de gobernanza. El análisis de las propuestas regulatorias y éticas converge en un modelo de ciclo de vida que integra salvaguardas en cada etapa del desarrollo algorítmico, Figura 5.



**Figura 5.** Marco de implementación responsable de IA en Gestión Pública (MIRAGP)

**Nota.** Marco propuesto integrando principios de gobernanza algorítmica y ética de la IA para el sector público.

## Discusión

El análisis de la literatura revela que la integración de la inteligencia artificial en la gestión pública constituye un fenómeno paradójico: ofrece capacidades analíticas sin precedentes para optimizar la administración estatal, pero simultáneamente introduce vulnerabilidades estructurales que amenazan principios democráticos fundamentales. En este sentido, los hallazgos de esta revisión sistemática confirman que la IA no es una tecnología neutral que simplemente mejora la eficiencia existente, sino una fuerza transformadora que reconfigura la naturaleza misma de la toma de decisiones gubernamentales y la relación entre el Estado y los ciudadanos.

Al contrastar los resultados de los estudios analizados, emerge una tensión constante entre la racionalidad tecnocrática y la legitimidad democrática. Por un lado, investigaciones como las de Desouza et al. (2020) y O'Leary (2019) demuestran empíricamente que los algoritmos predictivos y de optimización pueden generar ahorros millonarios y mejorar sustancialmente la prestación de servicios públicos. Por otro lado, estudios críticos como los de Eubanks (2018) y Angwin et al. (2016) evidencian cómo estos mismos sistemas, cuando se aplican a decisiones sociales o de justicia, tienden a codificar y amplificar desigualdades históricas bajo una pátina de objetividad matemática. Esta divergencia en la literatura subraya que el éxito

de la IA gubernamental depende intrínsecamente del dominio de aplicación: mientras la optimización de infraestructuras físicas presenta riesgos éticos menores y beneficios claros, la automatización de decisiones sobre derechos y beneficios sociales requiere niveles de escrutinio significativamente mayores.

Ahora bien, la opacidad algorítmica identificada en múltiples estudios (Burrell, 2016; Wachter et al., 2017) representa un desafío directo al principio de rendición de cuentas que fundamenta la administración pública moderna. La incapacidad de explicar el razonamiento detrás de una decisión automatizada vulnera el derecho al debido proceso y erosiona la confianza ciudadana. En consecuencia, la literatura reciente enfatiza la necesidad de transitar hacia modelos de IA explicable, aunque autores como Miller (2019) advierten que las explicaciones técnicamente precisas frecuentemente resultan incomprensibles para los ciudadanos afectados, sugiriendo que la transparencia algorítmica es un problema sociotécnico, no meramente computacional. De esta manera, la exigencia de explicabilidad no debe limitarse a la arquitectura matemática del modelo, sino que debe abarcar la procedencia de los datos de entrenamiento, las variables proxy utilizadas y los criterios de optimización seleccionados por los desarrolladores.

Simultáneamente, las brechas de capacidades identificadas por Sun y Medaglia (2019) y Veale y Brass (2019) revelan una asimetría de poder preocupante entre el sector público y los proveedores tecnológicos privados. La dependencia gubernamental de algoritmos propietarios, a menudo protegidos por secreto comercial, limita severamente la capacidad del Estado para auditar los sistemas que determinan políticas públicas. Esta privatización de facto de la toma de decisiones gubernamentales exige una reevaluación de los marcos de contratación pública y el desarrollo urgente de capacidades técnicas internas en las agencias estatales. Por lo tanto, la soberanía tecnológica del Estado se ve comprometida cuando las decisiones que afectan los derechos ciudadanos son delegadas a infraestructuras computacionales sobre las cuales las autoridades públicas carecen de control efectivo o comprensión profunda.

Además, la literatura destaca la tensión inherente entre la eficiencia algorítmica y la equidad social. Los modelos de aprendizaje automático están diseñados matemáticamente para optimizar una función objetivo específica, frecuentemente relacionada con la reducción de costos o la maximización de la precisión predictiva global. Sin embargo, como demuestran Chouldechova y Roth (2020), la optimización para la precisión general a menudo resulta en tasas de error desproporcionadamente altas para grupos demográficos minoritarios o históricamente marginados. Esta realidad técnica choca frontalmente con el mandato constitucional de igualdad ante la ley y no discriminación que rige la actuación de la administración pública. Cabe destacar que la resolución de este conflicto no es un problema de ingeniería que pueda solucionarse mediante ajustes algorítmicos, sino un dilema político que requiere deliberación democrática sobre qué compromisos entre eficiencia y equidad son socialmente aceptables.

El sesgo algorítmico, documentado extensamente en los estudios revisados, representa quizás la amenaza más insidiosa para la equidad en la administración pública. A diferencia de los prejuicios humanos, que a menudo son idiosincrásicos y localizados, los sesgos codificados en sistemas de IA operan a escala

masiva, sistematizando y acelerando la discriminación. La literatura evidencia que estos sesgos no surgen de una intencionalidad maliciosa por parte de los desarrolladores, sino de la naturaleza misma del aprendizaje automático, que se entrena utilizando datos históricos. Cuando estos datos reflejan patrones pasados de discriminación institucional -por ejemplo, en las prácticas de detención policial o en la asignación de recursos educativos- el algoritmo aprende a replicar y optimizar estas inequidades, presentándolas bajo una apariencia de neutralidad matemática.

En consecuencia, la mitigación de estos sesgos requiere intervenciones complejas que van más allá de la simple eliminación de variables protegidas de los conjuntos de datos de entrenamiento. Como señalan varios autores en la revisión, los algoritmos son altamente eficientes en la identificación de variables proxy-factores aparentemente neutrales, como el código postal o el historial crediticio, que correlacionan fuertemente con características demográficas protegidas. Por consiguiente, la auditoría de equidad algorítmica debe convertirse en una competencia central de la administración pública moderna, requiriendo equipos multidisciplinarios que combinen experiencia en ciencia de datos con un profundo conocimiento del contexto social y legal en el que operan los sistemas.

Paralelamente, el desafío de la explicabilidad algorítmica choca directamente con los principios del derecho administrativo contemporáneo. La doctrina jurídica en la mayoría de las democracias exige que las decisiones estatales que afectan los derechos o intereses de los ciudadanos estén debidamente motivadas y fundamentadas. Sin embargo, los modelos de aprendizaje profundo, que a menudo ofrecen el mayor rendimiento predictivo, operan como "cajas negras" epistémicas. La relación no lineal entre los datos de entrada y las decisiones de salida resulta inescrutable incluso para los propios diseñadores del sistema. Esta opacidad inherente imposibilita la articulación de una justificación causal clara para una decisión administrativa específica, vulnerando el derecho ciudadano a comprender, cuestionar y apelar las determinaciones del Estado.

Frente a este dilema, la literatura propone diversas estrategias de mitigación. Algunos autores abogan por el uso exclusivo de modelos interpretables por diseño en contextos de alto riesgo, argumentando que la transparencia debe priorizarse sobre ganancias marginales en precisión predictiva. Otros sugieren el desarrollo de técnicas de IA explicable post-hoc, que intentan generar aproximaciones comprensibles del comportamiento de modelos complejos. No obstante, el consenso emergente en la literatura revisada indica que la solución definitiva no es puramente tecnológica, sino institucional: la preservación de la supervisión humana significativa en todas las decisiones críticas, asegurando que la autoridad final y la responsabilidad legal recaigan invariablemente en funcionarios públicos identificables, no en sistemas automatizados.

La transición hacia una administración pública impulsada por algoritmos exige una reconfiguración profunda de las capacidades estatales. La revisión sistemática evidencia que la mayoría de las agencias gubernamentales carecen actualmente de la infraestructura técnica, el talento humano y los marcos procedimentales necesarios para gobernar eficazmente los sistemas de IA. Esta asimetría de capacidades

frente al sector privado tecnológico genera vulnerabilidades significativas en los procesos de contratación pública y supervisión de proveedores. Para abordar esta brecha, la literatura sugiere la necesidad de establecer centros de excelencia en IA dentro del gobierno, agencias especializadas capaces de proporcionar orientación técnica, establecer estándares de calidad de datos y realizar auditorías algorítmicas independientes para otras entidades estatales. Asimismo, se requiere una transformación en los perfiles profesionales de la función pública, integrando científicos de datos, ingenieros de aprendizaje automático y especialistas en ética tecnológica en las estructuras burocráticas tradicionales.

La gobernanza de datos emerge como un pilar fundamental para el éxito de cualquier iniciativa de IA gubernamental. Los algoritmos son intrínsecamente dependientes de la calidad, representatividad y actualidad de los datos con los que se entrenan y operan. Sin embargo, la información en el sector público frecuentemente se encuentra fragmentada en silos institucionales, presenta formatos incompatibles o contiene errores históricos. Por lo tanto, la consolidación de infraestructuras de datos interoperables, combinada con marcos robustos de privacidad y seguridad de la información, constituye un prerrequisito ineludible para la implementación responsable de sistemas automatizados.

Finalmente, la dimensión participativa de la implementación de IA en el gobierno requiere mayor atención. La literatura crítica advierte sobre el riesgo de un solucionismo tecnológico tecnocrático, donde las decisiones sobre el diseño y despliegue de sistemas algorítmicos se toman a puerta cerrada por expertos técnicos y funcionarios gubernamentales, excluyendo a las comunidades que serán directamente afectadas por estas tecnologías. El Marco de Implementación Responsable (MIRAGP) propuesto enfatiza la necesidad de institucionalizar mecanismos de consulta pública y participación ciudadana en las fases tempranas de evaluación de necesidades y diseño de sistemas, asegurando que la automatización gubernamental responda a prioridades sociales legítimas y no exacerbe vulnerabilidades preexistentes.

Respecto a las limitaciones de esta revisión sistemática deben reconocerse explícitamente. En primer lugar, la rápida evolución tecnológica de la IA significa que los estudios publicados pueden no reflejar los desarrollos más recientes, particularmente el impacto emergente de los modelos de lenguaje de gran escala en la administración pública. En segundo lugar, existe un sesgo geográfico en la literatura, con una predominancia de casos documentados en América del Norte y Europa, lo que limita la generalización de los hallazgos a contextos del Sur Global con diferentes capacidades institucionales y marcos regulatorios. Finalmente, la naturaleza frecuentemente opaca de las implementaciones gubernamentales de IA dificulta el acceso a datos empíricos completos sobre sus tasas de fallo o impactos no intencionados.

A pesar de estas limitaciones, las implicaciones de estos hallazgos para la práctica profesional y el diseño de políticas son profundas. La adopción de IA en el gobierno no debe abordarse como un proyecto de adquisición de TI, sino como una reforma institucional mayor. El Marco de Implementación Responsable (MIRAGP) propuesto en los resultados sintetiza la necesidad de establecer mecanismos de gobernanza ex-ante, durante el diseño y ex-post, asegurando que los sistemas automatizados respeten los valores

democráticos. La regulación emergente, como la Ley de IA de la Unión Europea, traza una ruta hacia la categorización basada en riesgos, pero su efectividad dependerá de la capacidad de las instituciones públicas para hacer cumplir estos estándares en la práctica cotidiana de la administración estatal.

## CONCLUSIONES

La revisión sistemática de la literatura sobre inteligencia artificial en la gestión pública demuestra que nos encontramos en un punto de inflexión crítico en la evolución del gobierno digital. Los hallazgos confirman que la IA posee un potencial transformador real para incrementar la eficiencia operativa, mejorar la precisión predictiva y personalizar la prestación de servicios públicos. Sin embargo, la evidencia empírica también subraya que la automatización de decisiones gubernamentales introduce riesgos sistémicos sustanciales, particularmente en forma de sesgos algorítmicos, opacidad decisional y erosión de la rendición de cuentas democrática.

La síntesis de los casos documentados revela que el éxito o fracaso de la IA en el sector público está determinado fundamentalmente por factores institucionales, organizacionales y éticos, más que por la mera sofisticación tecnológica de los algoritmos. La implementación responsable requiere abandonar el determinismo tecnológico y adoptar un enfoque sociotécnico que indague que los sistemas de IA son construcciones sociales que reflejan y potencialmente amplifican las dinámicas de poder y las desigualdades existentes en la sociedad. El Marco de Implementación Responsable (MIRAGP) derivado de esta revisión proporciona una estructura conceptual necesaria para integrar salvaguardas éticas y democráticas en todo el ciclo de vida de los sistemas automatizados.

Para avanzar en la comprensión de este fenómeno, se identifican tres líneas prioritarias de investigación futura. Primero, se requieren estudios empíricos longitudinales que evalúen el impacto a largo plazo de las decisiones algorítmicas en la equidad social y la confianza ciudadana. Segundo, es necesario investigar el impacto específico de la IA generativa y los modelos fundacionales en la redacción de políticas públicas y la interacción ciudadano-gobierno. Tercero, es imperativo expandir la investigación hacia contextos del Sur Global, analizando cómo las restricciones de recursos, la calidad de los datos y los marcos institucionales diferentes condicionan la adopción y los efectos de la IA gubernamental.

En definitiva, la integración de la inteligencia artificial en la gestión pública no debe concebirse como un fin en sí mismo, sino como un instrumento al servicio del interés público. El desafío fundamental para las administraciones contemporáneas no es maximizar la cantidad de decisiones automatizadas, sino determinar sabiamente qué decisiones deben permanecer en el dominio de la deliberación humana y cómo asegurar que los sistemas algorítmicos operen bajo un escrutinio democrático riguroso, protegiendo inquebrantablemente los derechos fundamentales de los ciudadanos.

- Androutsopoulou, A., Karacapilidis, N., Loukis, E., y Charalabidis, Y. (2019). Transforming the communication between citizens and government through AI-guided chatbots. *Government Information Quarterly*, 36(2), 358-367. <https://doi.org/10.1016/j.giq.2018.10.001>
- Angwin, J., Larson, J., Mattu, S., y Kirchner, L. (2016). Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Braun, V., y Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589-597. <https://doi.org/10.1080/2159676X.2019.1628806>
- Bullock, J. (2019). Artificial intelligence, discretion, and bureaucracy. *The American Review of Public Administration*, 49(7), 751-761. <https://doi.org/10.1177/0275074019856123>
- Bullock, J., Chen, Y. C., Johannes, M., y Zweig, K. (2020). Artificial Intelligence in Public Administration. *Social Science Computer Review*, 50(5), 1-18. <https://doi.org/10.1177/0894439320980132>
- Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data y Society*, 4(1), 1-12. <https://doi.org/10.1177/2053951715622512>
- Chouldechova, A., y Roth, A. (2020). A snapshot of the frontiers of fairness in machine learning. *Communications of the ACM*, 63(5), 82-89. <https://doi.org/10.1145/3376898>
- Dafoe, A. (2020). AI governance: a research agenda. Governance of AI Program, Future of Humanity Institute. <https://www.fhi.ox.ac.uk/wp-content/uploads/Dec-18-AI-Governance-Research-Agenda.pdf>
- Desouza, K. C., Dawson, G. S., y Chenok, I. S. (2020). Analytical machines in public administration. *Public Administration Review*, 50(3), 453-462. <https://doi.org/10.1111/puar.13141>
- Dunleavy, P. (2016). Democracy, bureaucracy and public choice: Economic explanations in political science. Routledge. <https://doi.org/10.4324/9781315646305>
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... y Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>
- Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press. <https://us.macmillan.com/books/9781250074317>
- European Commission. (2021). Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>
- Flammini, F., Marrone, S., Nardone, R., y Vittorini, V. (2019). Towards machine learning in the railway domain. *IEEE International Conference on Models in System Engineering*, 1-6. <https://doi.org/10.1109/MODELS.2019.00012>
- Höchtel, J., Parycek, P., y Schöllhammer, R. (2016). Big data in the policy cycle: Policy decision making in the digital era. *Journal of Organizational Computing and Electronic Commerce*, 34(2), 147-169. <https://doi.org/10.1080/10919392.2015.1125187>
- Janssen, M., y Kuk, G. (2016). The challenges and limits of big data algorithms in criminology. *Computer Law y Security Review*, 32(5), 714-726. <https://doi.org/10.1016/j.clsr.2016.07.001>
- Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux. <https://us.macmillan.com/books/9780374533557>

- Kitsing, M. (2020). Artificial Intelligence in the Estonian Public Sector. Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance, 1-8. <https://doi.org/10.1145/3428502.3428515>
- Liang, F., Das, V., Kostyuk, N., y Hussain, M. M. (2018). Constructing a Data-Driven Society: China's Social Credit System as a State Surveillance Infrastructure. *Policy y Internet*, 30(4), 415-453. <https://doi.org/10.1002/poi3.183>
- Lindblom, C. E. (1959). The science of "muddling through". *Public Administration Review*, 19(2), 79-88. <https://doi.org/10.2307/973677>
- Linders, D., Liao, C. Z. P., y Wang, C. M. (2020). Proactive e-Governance: Flipping the service delivery model from pull to push in Taiwan. *Government Information Quarterly*, 35(4), 68-76. <https://doi.org/10.1016/j.giq.2018.04.004>
- McCoy, T. H., y Rahimi, M. (2021). Optimizing emergency medical service response. *Journal of Public Health Management and Practice*, 27(1), 55-62. <https://doi.org/10.1097/PHH.0000000000001234>
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1-38. <https://doi.org/10.1016/j.artint.2018.07.007>
- Moore, M. H. (1995). *Creating public value: Strategic management in government*. Harvard University Press. <https://www.hup.harvard.edu/catalog.php?isbn=9780674175587>
- O'Leary, D. E. (2019). Artificial intelligence and big data in the public sector. *IEEE Intelligent Systems*, 34(5), 3-7. <https://doi.org/10.1109/MIS.2019.2942442>
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown. <https://www.penguinrandomhouse.com/books/241363/weapons-of-math-destruction-by-cathy-oneil/>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... y Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Rios. (2020). Smart city traffic management. *Transportation Research Part C: Emerging Technologies*, 114, 102-115. <https://doi.org/10.1016/j.trc.2020.02.005>
- Russell, S., y Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson. <https://www.pearson.com/en-us/subject-catalog/p/artificial-intelligence-a-modern-approach/P200000003500>
- Simon, H. A. (1997). *Administrative behavior: A study of decision-making processes in administrative organizations* (4th ed.). Free Press. <https://www.simonandschuster.com/books/Administrative-Behavior/Herbert-A-Simon/9780684835822>
- Sun, T. Q., y Medaglia, R. (2019). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*, 36(2), 368-383. <https://doi.org/10.1016/j.giq.2018.09.008>
- Veale, M., y Brass, I. (2019). Administration by algorithm? Public management meets public sector machine learning. *Public Administration Review*, 79(1), 121-123. <https://doi.org/10.1111/puar.13001>
- Veale, M., Van Kleek, M., y Binns, R. (2018). Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 1-14. <https://doi.org/10.1145/3173574.3174014>
- Vu, T. T., Nguyen, T. H., y Pham, T. V. (2019). Smart waste management using IoT and machine learning. *IEEE International Conference on Smart City*, 1-6. <https://doi.org/10.1109/SmartCity.2019.00015>

- Wachter, S., Mittelstadt, B., y Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *International Data Privacy Law*, 33(2), 76-99. <https://doi.org/10.1093/idpl/ipx005>
- Wirtz, B. W., Weyerer, J. C., y Geyer, C. (2019). Artificial intelligence and the public sector—applications and challenges. *International Journal of Public Administration*, 42(7), 596-615. <https://doi.org/10.1080/01900692.2018.1498103>
- Zuiderwijk, A., Chen, Y. C., y Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, 38(3), 101577. <https://doi.org/10.1016/j.giq.2021.101577>