

# **Inteligencia Artificial en la administración pública: ambigüedad y elección de sistemas de IA y desafíos de gobernanza digital**

**Fernando Filgueiras**

## **Fernando Filgueiras**

Profesor asociado de la Escuela de Ciencias Sociales de la Universidad Federal de Goiás (UFG), Brasil. Profesor afiliado en el Ostrom Workshop on Political Theory and Policy Analysis, Indiana University, Estados Unidos. Fue Director de Investigación y Estudios de Posgrado en la Escuela Nacional de Administración Pública (ENAP), Brasil. También fue Director de la Facultad de Filosofía y Ciencias Humanas de la Universidad Federal de Minas Gerais. Es doctor en Ciencias Políticas por el Instituto Universitario de Investigaciones de Río de Janeiro (IUPERJ). Entre sus publicaciones se destaca *Governance for the Digital World* (Palgrave, 2020).

Las comunicaciones con el autor pueden dirigirse a:  
E-mail: fernandofilgueiras@hotmail.com

## **Inteligencia Artificial en la administración pública: ambigüedad y elección de sistemas de IA y desafíos de gobernanza digital**

*A medida que las administraciones públicas adoptan la inteligencia artificial (IA) se observa que esta transición tiene el potencial de transformar al servicio público y a las políticas públicas, al ofrecer un cambio rápido en la toma de decisiones y la prestación de servicios. Sin embargo, una reciente serie de críticas ha señalado aspectos problemáticos de la integración de los sistemas de IA en la administración pública, señalando resultados problemáticos en términos de justicia y valores. El argumento que se proporciona aquí es que cualquier administración pública que adopte sistemas de IA debe considerar y abordar las ambigüedades e incertidumbres que rodean a dos dimensiones clave: los resultados de los algoritmos y cómo los administradores públicos toman decisiones para y sobre el diseño de los sistemas de IA. Este artículo señala la necesidad de diseñar instituciones que se sumerjan en la comprensión de los matices, los detalles y los posibles resultados de la gobernanza de la IA para la administración pública. Tales instituciones reconciliarían la lógica consecuencialista con una lógica de adecuación para ayudar a navegar y mediar las ambigüedades e incertidumbres.*

**Palabras clave:** Teoría de la Administración Pública; Gestión Pública; Gobernanza; Inteligencia Artificial; Toma de Decisiones; Perspectivas

### **Artificial Intelligence in Public Administration: Ambiguity and Choice of IA Systems and Challenges of Digital Governance**

*As public administrations embrace artificial intelligence (AI) it is noted that this transition has the potential to transform public service and public policies by offering rapid changes in decision making and service delivery. However, a recent series of critiques have pointed out the problematic aspects of integrating AI systems into public administration, noting some problematic outcomes in terms of justice and values. The argument provided here is that any public administration adopting AI systems must consider and address the ambiguities and uncertainties surrounding two key dimensions: the algorithms' outcomes and how public managers make decisions for and about the design of AI systems. This article points out the need to design institutions that immerse themselves in understanding the nuances, details, and potential outcomes of AI governance for public administration. Such institutions would reconcile consequentialist logic with a logic of appropriateness to help navigate and mediate ambiguities and uncertainties.*

Recibido: 24-11-2020 y 25-02-2021 (segunda versión). Aceptado: 26-02-2021.

**Key words:** Public Administration Theory; Public Management; Governance; Artificial Intelligence; Decisions Making; Perspectives

### **Introducción**

**E**l Oráculo de Delfos fue una de las instituciones más importantes del mundo griego antiguo. Su rol en la sociedad era producir predicciones sobre el futuro, resolver las preguntas sobre problemas personales o sobre si ir o no a la guerra. Emperadores y reyes como Alejandro o Licurgo asistieron al Oráculo de Delfos en busca de consejo sobre las guerras. El Oráculo estaba dedicado a Apolo, quien derrotó a Python y tomó su posesión. Pero dentro del Oráculo, los consultantes preguntaron a la Pitia (la suma sacerdotisa) sobre sus problemas. La Pitia inhaló los gases del templo y encarnó a Apolo, y luego Apolo dio la predicción.

La importancia del Oráculo en Delfos en el mundo griego antiguo era reducir las incertidumbres en el proceso de toma de decisiones, incluida la toma de decisiones políticas. Sin embargo, como ocurre con cualquier comunicación en un escenario de alto riesgo, requería una inspección minuciosa de la redacción y el significado o la interpretación. Las predicciones de la Pitia se tomaron como verdades absolutas y, por lo general, requerían la interpretación de los sacerdotes o los consejeros del rey para garantizar una comprensión completa.

Heródoto relata un episodio en el que Temístocles estaba seguro de ir a la batalla de Salamina, pero consultando a la Pitia recibió una respuesta negativa. Debido a que quería entender con más claridad e interpretar el significado de la predicción con cuidado, hubo una segunda consulta. La segunda respuesta fue más ambigua y desconcertante. La Pitia señaló una batalla divina en Salamina, lo que provocó que Temístocles apelara a los intérpretes del Oráculo, quienes se dedicaron a concretar lo que aconsejaba Apolo. Después de la interpretación de Timón, Temístocles fue a la guerra y los griegos ganaron la batalla de Salamina.

Este episodio narrado por Heródoto ilustra uno de los mayores desafíos que rodean el proceso de toma de decisiones. Los tomadores de decisiones siempre se enfrentan al problema de las incertidumbres y a cómo estas afectan directamente al proceso de las políticas. Hoy en día la tecnología parece ser la nueva Pitia, ya que la humanidad busca tecnología que reduzca las incertidumbres en los asuntos públicos, cubriendo temas que van desde la guerra hasta una simple asignación de beneficios sociales. Garantizar la información

**Los procesos de transformación digital en los gobiernos están reconstruyendo la idea de la tecnología predictiva que toma decisiones del sector público de manera autónoma.**

correcta que ingresa en ese proceso e interpretarla correctamente para obtener mejores resultados sigue siendo uno de los mayores desafíos para la humanidad.

Los procesos de transformación digital en los gobiernos están reconstruyendo la idea de la tecnología predictiva que toma decisiones del sector público de manera autónoma. Los gobiernos invierten recursos sustanciales en el desarrollo de tecnologías de inteligencia artificial (IA) para automatizar y personalizar los servicios públicos. Presionado por cuestiones fiscales, la perspectiva del gobierno es reducir los costos de los servicios públicos y proporcionar medios para los aspectos gubernamentales de la toma de decisiones. El rediseño de los servicios públicos mediante la aplicación de tecnologías de IA permite la transición a la gobernanza de la era digital, cambiando la relación entre los gobiernos y la sociedad a través de las tecnologías digitales (Dunleavy ...[et al], 2006). La razón de ser de la aplicación de la IA es generar una ventana de oportunidad para que las decisiones sobre los servicios públicos y las políticas puedan ser más imparciales, basadas en principios éticos que buscan las decisiones correctas mediadas por máquinas.

Sin embargo, el estado actual de la técnica de aplicación de las tecnologías de IA en la esfera pública plantea varios desafíos. Es importante comprender que la tecnología de IA no cambia un conjunto de problemas. Las decisiones en los servicios públicos y las políticas siguen siendo un trabajo ritualista con mensajes ambiguos que generan interpretaciones conflictivas (Zahariadis, 2016). Lo que se ha hecho evidente es que, si bien los algoritmos automatizan el proceso de decisión, también crean un nuevo orden de problemas. Los algoritmos producen y contribuyen a varias formas de injusticia, incluidos los prejuicios raciales o de género (Benjamin, 2019; Noble, 2018), formas de colonialismo de datos (Couldry y Mejias, 2019) y la producción de nuevas desigualdades (Eubanks, 2018; O'Neil, 2016). Por ejemplo, la aplicación de la IA a los grados "A-level" en Inglaterra y Escocia generó resultados controvertidos al examinar a los estudiantes para la admisión a la universidad. Los resultados dieron lugar a interpretaciones contradictorias que resultaron con el público exigiendo disculpas al gobierno y un enfoque alternativo del Departamento de Educación para que las universidades flexibilizaran el examen (Pagnamenta, 2020).

Como tecnología ritualista que produce predicciones para reducir las incertidumbres, no es exagerado ver por qué la IA ha sido exaltada

**Sin embargo, la experiencia muestra que las decisiones siguen siendo ambiguas y requieren interpretaciones conflictivas e incompletas.**

como la nueva Pitia para la administración pública. Sin embargo, la experiencia muestra que las decisiones siguen siendo ambiguas y requieren interpretaciones conflictivas e incompletas. Esto deja a muchos con preguntas fundamentales: ¿Por qué, incluso con la adopción de algoritmos de IA, las decisiones de política y servicio público siguen siendo ambiguas? ¿Cuáles son los desafíos que se materializan para la gobernanza de la IA en la administración pública? Llegar a una respuesta bien considerada a tales preguntas requiere una comprensión profunda de la transformación digital y el papel de las incertidumbres en la toma de decisiones.

### **1. IA en la administración pública: transformación digital y toma de decisiones autónoma**

La IA no tiene una definición precisa en la literatura (Wang, 2019). No significa automatizar actividades repetitivas a través de códigos. Tampoco significa una reproducción de los procedimientos “si... entonces” comunes en la programación. La IA implica un proceso de emulación de la inteligencia humana para establecer una decisión y hacer que la máquina desarrolle una habilidad para resolver problemas de manera más eficiente. La definición de IA se inserta como una abstracción de la mente humana para diseñar soluciones que imiten la racionalidad (Simon, 1957; Russell, 1997), funciones cognitivas (Russell y Norvig, 2010), capacidad (Minsky, 1985), comportamiento (Flach, 2012) y la estructura del pensamiento (Markram, 2006). Aplicada al sector público, el objetivo de la IA es producir predicciones que superen el rendimiento humano para decidir (Russell y Norvig, 2010). Entendida de esta manera, la IA representa sistemas que toman decisiones basadas en el aprendizaje, la creatividad y un mayor rendimiento. Aplicada a los gobiernos, la IA representa una herramienta importante para aumentar el desempeño (Wirtz, Weyerer y Geyer, 2019). Sin embargo, encarna una serie de nuevos riesgos para la gestión pública que requiere de mecanismos de gobernanza orientados a implementar tecnologías en los gobiernos (Filgueiras y Almeida, 2020).

La aplicación de la IA en la administración pública implica dos aspectos. Primero, aumentar la eficiencia de la máquina administrativa. Al expandir las capacidades de una administración pública para resolver diferentes problemas, la aplicación de IA aumenta la eficiencia del sector público. Puede ocurrir con la constitución de agentes virtuales en diferentes dominios de políticas, como salud e

**Más que un conjunto complejo de sistemas y códigos, las tecnologías digitales en la administración pública pueden entenderse como herramientas que median en la autoridad estatal para implementar políticas y servicios públicos.**

inmigración (Zheng ...[et al], 2018; Mehr, 2017), control de seguridad y monitoreo y reconocimiento facial para identificar criminales (Power, 2016), vehículos autónomos para el transporte público (Jefferies, 2016), *chatbots* para el servicio público (Mehr, 2017) y diagnóstico de imágenes para acelerar la atención médica (Collier ...[et al], 2017). En todas estas situaciones la IA cambia la relación costo-beneficio en el sector público, aumentando la eficiencia de la administración pública en diferentes sectores sociales.

El segundo aspecto involucra el proceso de elección (o decisión) de los administradores públicos. Desde el principio es importante pensar en la elección de tecnologías como una construcción social que involucra comportamientos, actitudes y cogniciones en el proceso de transformación digital (Mergel, Edelman y Haug, 2019). En muchas situaciones las decisiones de política sobre tecnología están impulsadas por una lógica de costo y beneficio, en la que la IA puede modificar los patrones organizacionales de la administración pública con el potencial de fomentar la coproducción de servicios públicos (Linders, 2012; Meijer, 2012), fomentar la colaboración entre organizaciones (Chen, 2017) y promover gobiernos abiertos (Janssen, Charalabidis y Zuiderwijk, 2012). Más que un conjunto complejo de sistemas y códigos, las tecnologías digitales en la administración pública pueden entenderse como herramientas que median en la autoridad estatal para implementar políticas y servicios públicos (Hood y Margetts, 2007).

Las herramientas digitales están sujetas a opciones técnicas y políticas que comprenden una combinación compleja que explica el grado de efectividad de la implementación de una política (Filgueiras, Cireno y Palotti, 2019). Las herramientas digitales especifican soluciones tecnológicas a aplicar para resolver problemas públicos. La elección de técnicas suele depender de factores que optimizan los resultados para la administración pública. En este caso, la IA debería tener como objetivo promover:

- **Entrega de servicios e implementación de políticas más rápidas:** las computadoras pueden analizar millones de situaciones individuales rápidamente.
- **Costos más bajos:** los sistemas pueden realizar tareas básicas y repetitivas con pocos recursos empleados.
- **Precisión:** las computadoras pueden manejar situaciones complejas y proporcionar resultados confiables.
- **Nuevas capacidades:** algunas aplicaciones son imposibles de alcanzar sin introducir computadoras.

Estas herramientas digitales se pueden aplicar para promover la gobernanza y ampliar la prestación de servicios públicos y las capacidades de políticas públicas (Veale y Brass, 2019). Permiten nuevas capacidades y amplían las posibilidades de acción de las organizaciones estatales, transformando el comportamiento de los individuos en la sociedad. La transformación digital implica la capacidad de automatizar y aumentar la administración pública. Las herramientas digitales transforman tanto las políticas como los aspectos organizativos y políticos de la administración pública. La IA transforma la cultura administrativa y modifica las ideologías presentes en la gestión pública, afirmando a veces valores importantes como el mérito o la neutralidad política, otras veces nuevos valores como la eficiencia y el control (Ramió, 2018). La transformación digital en las organizaciones públicas cambia varios aspectos institucionales de la administración pública, representando más que simplemente una reducción de los costos financieros de los servicios públicos. El fundamento de la transformación digital en la administración pública es provocar cambios en el funcionamiento institucional de la administración pública. Estos cambios están catalizados por el tema del control, los costos, la conveniencia y la conexión (Jeffares, 2021).

Este proceso organizacional disruptivo es el que la gobernanza digital debe abordar de manera tal que se aproveche el potencial de las tecnologías digitales en el proceso de gobernanza. El potencial disruptivo que las herramientas digitales brindan a las organizaciones es enorme. La transformación digital es un fenómeno importante para la investigación sobre la administración pública (Luna-Reyes y Gil-García, 2014), así como para los profesionales (Westerman, 2016). La transformación digital considera la tecnología como parte de una innovación compleja que involucra a organizaciones, instituciones, ciudadanos y empresas para cambiar la cadena de valor de la administración pública.

El aspecto central que se debe considerar es la adopción de algoritmos que produzcan la creciente automatización y aprendizaje dentro de la administración pública. Los algoritmos son secuencias de instrucciones para resolver problemas o realizar tareas y se han aplicado a los procesos de la administración pública durante décadas para ayudar a la toma de decisiones. La gran disrupción hoy en día es el análisis de *big data*, porque ahora los sistemas de inteligencia artificial pueden aprender de datos masivos y tomar decisiones de gobernanza de forma autónoma (Williamson, 2014). Se han



desarrollado algoritmos de aprendizaje automático para abordar problemas complejos mediante la aplicación de grandes volúmenes de datos (Domingos, 2015). También pueden constituir un aprendizaje supervisado o no supervisado para construir predicciones sobre los problemas planteados.

### **1.1 La reencarnación de la Pitia: calificar nuevas predicciones y sopesar los resultados**

En el proceso de transformación digital, la adopción de sistemas de inteligencia artificial significa que la administración pública ahora aprende de las máquinas. Los sistemas de IA juegan un papel fundamental como las nuevas Pitias en la administración pública: predicen alternativas de políticas y resuelven problemas. Por ejemplo, la red interconectada de dispositivos de recolección de datos proporciona una herramienta fundamental para mejorar las políticas y los servicios públicos (Greengard, 2015; Danaher ...[et al], 2017). Las administraciones públicas tienen como objetivo automatizar la prestación de servicios y hacer uso de algoritmos y datos distribuidos masivamente para anticipar demandas futuras (Williamson, 2014; Chen y Hsieh, 2014). De esta manera, los algoritmos pueden predecir comportamientos individuales y utilizar la capacidad predictiva para cambiar las expectativas, introducir incentivos e influenciar a los usuarios del servicio (Dunleavy, 2016; Kosters y Van der Heijden, 2015). El uso de sistemas de inteligencia artificial en la administración pública tiene el potencial de reducir la burocracia, a la vez que es más responsable y eficaz con el potencial de autoservicio por parte de los usuarios.

Por su capacidad predictiva, los sistemas de IA cambian varios aspectos de la administración pública y el orden social en áreas como educación, salud, justicia penal o políticas sociales (Gillingham, 2015; Danaher, 2016). También pueden afectar la cultura, modificar el conocimiento y construir realidades (Just y Latzer, 2017; Napoli, 2014). Pueden reemplazar a los burócratas de la calle y participar en actividades con discreción para implementar políticas y servicios (Jeffares, 2021; Alkhatib y Bernstein, 2019). De manera única, pueden representar y realizar soluciones de diseño sin interacción humana. Sin embargo, los sistemas de IA no deben verse solo dentro del proceso de gobernanza, sino que además ejercen una autoridad tecnocrática que tiene consecuencias políticas. En la toma de decisiones, los algoritmos primero deciden sobre un sesgo epistémico.

**Si hasta hace poco tiempo la adopción de sistemas de IA en la administración pública representaba una esperanza, actualmente se está acumulando una ola de críticas y problemas en torno a los resultados de su aplicación.**

Como tecnología de poder, los algoritmos de IA deciden técnicamente entre los gobiernos y los gobernados (Janssen y Kuk, 2016; Danaher, 2016). De esta forma, las decisiones importantes quedan a criterio de los algoritmos (Andrews, 2018).

Con eso en mente, ¿hasta qué punto se puede decir que estas decisiones tomadas por los sistemas de IA reemplazan a la administración pública? Si bien existe una fuerte esperanza de que los sistemas de IA puedan cumplir varias tareas y cambiar la administración pública en la dirección de una mayor personalización de los servicios públicos para los ciudadanos, simultáneamente existe una cuadrícula de engaños que trae problemas relacionados con las soluciones tecnológicas, la confianza y la eficacia. Los beneficios de aplicar sistemas de IA son innegables. De hecho alteran la relación costo-beneficio, lo que implica una racionalidad consecuente para la adopción de sistemas de IA en la administración pública; y, por otro lado, representan un paso hacia el diseño de soluciones que buscan la decisión más racional y eficiente. Por lo tanto, los sistemas de IA coinciden con el índice de eficiencia en tecnología para la gobernanza y la gobernanza de la tecnología.

Sin embargo, si hasta hace poco tiempo la adopción de sistemas de IA en la administración pública representaba una esperanza, actualmente se está acumulando una ola de críticas y problemas en torno a los resultados de su aplicación. Estas dificultades en la gobernanza de la inteligencia artificial plantean una serie de desafíos relacionados con el funcionamiento de la tecnología del poder, especialmente respecto a cómo la tecnología puede ejercer decisiones discrecionales que afectan la vida de los ciudadanos. Los problemas raciales, por ejemplo, implican cuestiones de derechos y valores que no siempre se reproducen en la sociedad. Cuando los sistemas de inteligencia artificial aprenden emulando el pensamiento humano, esto puede llevar a una discriminación racial repetida a gran escala y, aunque esto puede no comprometer la eficiencia del algoritmo, sí compromete su efectividad en el cumplimiento de importantes funciones sociales (Benjamin, 2019). Los criterios de eficiencia pueden suplantar los criterios de validez, creando dificultades para aceptar la legitimidad de las decisiones de los sistemas de IA.

Los gobiernos tienden a adoptar sistemas de inteligencia artificial como un proceso para tratar a los ciudadanos regularmente como objetos programables, usando *nudging* y creando personalización. Por una lógica consecuencialista, los sistemas de IA crean dilemas

sociales que reducen la capacidad de cooperación y refuerzan problemas como la tragedia de los comunes (Frischmann y Selinger, 2018; Almeida, Filgueiras y Gaetani, 2020). Los sistemas de IA buscan decisiones más eficientes, pero estas no son necesariamente las decisiones más correctas. En situaciones de conflicto social pueden reforzar los dilemas sociales, creando situaciones en las que una menor cooperación y equilibrio conducen a políticas con ganadores y perdedores. Los dilemas sociales creados por los sistemas de IA surgen de una racionalidad que los organiza mirando al pasado para predecir el futuro.

Como las nuevas Pitias en la administración pública, los sistemas de IA predicen ambiguamente decisiones que buscan ser eficientes, pero no necesariamente correctas. Este es un problema recurrente en su diseño: que las soluciones indicadas no siempre serán suficientes para constituir certezas. Si se parte de la definición de IA como la emulación del pensamiento humano, se debe considerar el papel de la ambigüedad relacionada con la interpretación del pensamiento. Las decisiones tomadas por los sistemas de IA son naturalmente ambiguas y dependen de la interpretación de las alternativas predichas.

## **2. Ambigüedades e incertidumbres en los sistemas de inteligencia artificial**

El mayor desafío de implementar la IA en la administración pública no es el hecho de que los sistemas puedan automatizar acciones repetitivas, sino utilizar su capacidad predictiva para que las máquinas puedan decidir y tomar el mejor curso de acción para el gobierno. La creciente personalización de los servicios públicos, en previsión de las demandas de los ciudadanos, requiere que las máquinas sean mucho más sofisticadas y no simplemente realizar actividades mecánicas y repetitivas. La aplicación de la IA requiere que las máquinas intercedan en el pensamiento, que aprendan y puedan emular las características del pensamiento humano y desarrollen la intuición (Simon, 1983; Newell y Simon, 1961). Por lo tanto, la IA en la administración pública desempeña un papel discrecional en el que las máquinas pueden aprender y ser sensibles a las diferentes demandas de los ciudadanos y a los problemas organizativos. Idealmente, entonces, estas máquinas son capaces de tomar decisiones sobre la asignación de recursos, derechos, cumplimiento de deberes y situaciones discursivas dentro del ámbito de la política.

**Las decisiones de la administración pública no siempre siguen una lógica racional dirigida a las consecuencias. La mayoría de las veces buscan la decisión más adecuada, apuntando a valores que definirán el motivador de la acción institucional.**

## **2.1 Avance y capacidad predictiva**

El avance en la aplicación de los sistemas de IA es innegable, la revolución del *big data* ha llevado su uso en la administración pública a otro nivel. Sin embargo, la ilusión hacia el uso de sistemas de IA está aumentando (Smith, 2018). El avance de los sistemas actuales se centra en emular el pensamiento humano y actualizar los niveles de la mecánica cuántica para expandir la capacidad computacional (Simon, 1979). Sin embargo, los diseñadores también deben tener en cuenta que el pensamiento humano y la arquitectura de decisiones ocurren en contextos de incertidumbre y ambigüedad (March, 1991).

La capacidad predictiva de los sistemas de IA presupone un trabajo racional de toma de decisiones, utilizando la lógica consecuencialista. A menudo, los diseñadores basan la solución predictiva para aplicar sistemas de inteligencia artificial en la lógica de costo y beneficio. El supuesto es que la IA puede ayudar a los responsables políticos a tomar mejores decisiones (O'Malley, 2014) y mejorar la eficiencia y la eficacia del gobierno (Milakovich, 2012). Los sistemas de IA tienen el potencial de construir decisiones más racionales, incluso en contextos políticos complejos, porque optimizan la información, produciendo una cadena de decisiones entre el analista de datos y el tomador de decisiones (Van der Voot ...[et al], 2019).

Lo que sigue siendo un desafío, sin embargo, es comprender que múltiples actores participan en las decisiones de la administración pública y cómo (Scharpf, 1988). Una característica esencial es que las decisiones en las organizaciones son ambiguas y la arquitectura de decisión no es un proceso propiamente racional y consecuente (Cohen, March y Olsen, 1972; March y Olsen, 1976). Las ambigüedades e incertidumbres comprenden características importantes de la teoría institucional en las ciencias sociales. Para esta teoría, los individuos no son atomistas, sino que están integrados en una serie compleja de relaciones con otros individuos y con colectividades (Peters, 2019). La toma de decisiones no depende del recuento de preferencias de los individuos atomísticos, sino de la colección de normas, reglas, entendimientos y rutinas que definen las acciones apropiadas en términos de relación entre roles y situaciones. Las decisiones de la administración pública no siempre siguen una lógica racional dirigida a las consecuencias. La mayoría de las veces buscan la decisión más adecuada, apuntando a valores que definirán el motivador de la acción institucional (March y Olsen, 2009). Estos valores que definen una lógica de adecuación de las decisiones públicas se basan

en valores éticos, ideologías o valor público, que informan cuándo una decisión será legítima en términos públicos.

Las decisiones en la administración pública se toman en contextos de racionalidad limitada (Simon, 1957), siempre que los problemas sean perversos y las soluciones dependan de una comprensión limitada de la situación de acción. Esta es una premisa importante de la teoría institucional que debería participar en la comprensión de la adopción de sistemas de IA en la administración pública<sup>1</sup>. Las decisiones a menudo se toman mirando al pasado para acumular aprendizaje, comprender el presente y predecir el futuro. En este proceso, el aprendizaje comprende incertidumbres y las ambigüedades dominan el escenario de decisión (March y Olsen, 1975). Las tecnologías no son claras y los contextos de las decisiones dependen de un proceso fluido de participación de los tomadores de decisiones. El encuentro entre problemas y soluciones no es un proceso racional ni lineal (Cohen, March y Olsen, 1972).

Por el contrario, los sistemas de IA pueden representar soluciones en busca de problemas. En este caso la definición del problema se vuelve fundamental para que los sistemas de IA se puedan implementar en una administración pública. En este contexto, es necesario reflexionar sobre cómo las ambigüedades se expanden y desafían la aplicación de la IA en la administración pública. Esto es válido tanto en la lógica de formulación de problemas, que deben ser interpretados, como en la lógica de los resultados obtenidos con los algoritmos de decisión autónomos.

Es necesario entender que los sistemas de IA están ocupando paulatinamente un lugar institucional en la administración pública. En otras palabras, los sistemas de IA emulan la inteligencia humana para tomar decisiones y crear aplicación con normas, reglas, rutinas y valores apropiados que apoyen el aparato institucional del Estado. Por ejemplo, cuando un sistema de IA asigna recursos a políticas sociales, calculando volúmenes masivos de datos para especificar quién es elegible para recibir el beneficio, define quién, dónde, cuándo y cómo tiene acceso a un derecho. Otro ejemplo es cuando se aplica un sistema de IA a la selección de personas, este establecerá cuáles son las expectativas con relación al perfil que se debe seleccionar y calculará entre varios datos quien mejor se ajusta al perfil deseado. En teoría, esta puede ser una forma de decisión más racional, libre de pasiones. Sin embargo, siempre será una decisión ambigua, porque depende de valores que se pusieron previamente

**El diseño de las soluciones de los sistemas de IA podría no coincidir con el diseño de las soluciones de los problemas de la administración pública. Esto se debe a una comunicación asimétrica entre los analistas de datos y los tomadores de decisiones, que no tienen los mismos antecedentes y conocimientos.**

en el diseño del sistema. Por ejemplo, ¿el algoritmo consideró cuestiones raciales o de género? Al decir que las decisiones del sistema de IA reproducen ambigüedades, se considera que los efectos de la decisión que toma un sistema autónomo pueden tener varios flujos diferentes, dependiendo de cómo los actores -los afectados por la decisión- entienden y valoran el resultado de aplicar el sistema de IA. Este es un gran desafío para la implementación de soluciones de IA en la administración pública, porque a medida que se toman las decisiones, estas pueden presentar múltiples flujos y reproducir incertidumbres (Cohen, March y Olsen, 1972; Kingdon, 1995).

Ahora bien, si se considera que los problemas públicos suelen ser perversos, mal formulados, inestables, no definitivos y vinculados a otros problemas (Peters, 2018), estos problemas crearon una situación de “demasiados cocineros estropean el caldo”. Los problemas mal formulados impactan directamente en la formulación de soluciones de IA (Simon, 1973). Considerando una lógica consecuencialista, los problemas mal formulados conducirán a soluciones de IA mal insertadas en contextos de ambigüedades e incertidumbres. Estos problemas surgen de incertidumbres previas (March y Olsen, 1975) y en el futuro se acumulan como problemas de memoria y aprendizaje (Simon, 1973). Por lo tanto, el diseño de las soluciones de los sistemas de inteligencia artificial podría no coincidir con el diseño de las soluciones de los problemas de la administración pública. Esto se debe a una comunicación asimétrica entre los analistas de datos y los tomadores de decisiones, que no tienen los mismos antecedentes y conocimientos. La administración pública no tiene un canal establecido para lidiar con el diseño de sistemas de inteligencia artificial, lo que genera dificultades para comunicar las demandas entre los tomadores de decisiones y los diseñadores de sistemas (Margetts y Dorobantu, 2019).

Además, las decisiones en la administración pública no ocurren solo en una lógica consecuencialista. Las decisiones que implican el comportamiento de los actores ocurren de una manera basada en reglas, requiriendo interpretaciones de la acción en contexto. Muchas acciones de formulación de políticas están impulsadas por reglas de comportamiento apropiado, organizadas en instituciones. La lógica de la adecuación implica que se espera que los actores se comporten de acuerdo con las normas y reglas de la sociedad (March y Olsen, 2009) y, por lo tanto, que los sistemas de IA deben observar reglas y situaciones que se insertan en el comportamiento adecuado, como

sentimientos, valores éticos, obligaciones y virtudes. A diferencia de una lógica consecuencialista, en la que las soluciones de IA pueden predecir e influir en el comportamiento, la lógica de la adecuación exige términos opacos, en los que su poder prescriptivo depende del análisis de conjuntos de situaciones que varían en precisión y claridad (March y Olsen, 2009).

La adopción de sistemas de IA en la administración pública se da en dos dimensiones. En primer lugar, al formular problemas de políticas donde se produce en una lógica consecuencialista, con el objetivo de reformar la administración pública a través de tecnologías digitales. Las presiones fiscales de los gobiernos explican las decisiones de los administradores públicos para los procesos de transformación digital, porque mejoran la relación entre costos y beneficios (Dunleavy y Margetts, 2013). Sin embargo, la adopción de sistemas de IA en la administración pública no se limita únicamente a la reducción de costes, sino que expande las capacidades de la administración pública (Veale y Brass, 2019) y tiene el potencial de expandir también el control, la conveniencia y la conexión con los ciudadanos (Jeffares, 2021). En cualquiera de estas situaciones, la lógica de la transformación digital con la adopción de sistemas de IA se moviliza por una lógica consecuencialista, reflejando los beneficios potenciales de la solución cuando se inserta en el proceso de toma de decisiones en la administración pública. El problema que se señala en este trabajo, siguiendo la teoría institucional, es que el proceso de toma de decisiones en la administración pública no solo sigue una lógica consecuencialista, sino también una lógica de adecuación incrustada en la ética, las rutinas, los roles y los valores. La formulación de problemas no siempre ocurrirá debido a la necesidad real de servicio o política pública, sino a una lógica consecuencialista motivada por una relación entre costo y beneficio. En la segunda dimensión, hay una evaluación de los resultados de los sistemas de IA motivada por una lógica de adecuación. La adopción de sistemas de IA en la administración pública incurre en una serie de críticas motivadas por criterios de equidad, género, raza, efectividad y valores éticos (O'Neil, 2016; Noble, 2018; Benjamin, 2019; Wachter, Mittelstadt y Russell, 2020).

El resultado de este proceso es un aumento de las ambigüedades de la presencia de los sistemas de IA en la administración pública. Motivado por la lógica consecuencialista, la adopción de sistemas de inteligencia artificial en la administración pública se vuelve irreconciliable con una

lógica de adecuación. Si bien la transformación digital de la administración pública requiere cálculos de utilidad, constituye una tensión con hábitos, emociones y valores que impulsan la lógica de adecuación. Esto requiere estructuras de gobernanza que impliquen equilibrar esta tensión entre diferentes lógicas de acción, lo que implica ambigüedades e incertidumbres tanto en el proceso de adopción de sistemas de IA, como en los resultados y seguimiento de dichos sistemas.

## **2.2 Interpretando y modificando las nuevas Pitias: construyendo caminos de la ambigüedad a la reconciliación**

Como las nuevas Pitias en la administración pública, los sistemas de IA ofrecen mensajes ambiguos que requieren la interpretación de reglas y principios internalizados en un contexto institucional. Sin considerar una reconciliación entre una lógica consecuencialista y una lógica de adecuación, pueden convertirse en una panacea para los problemas de la administración pública. Los períodos de crisis y cambios radicales en la administración pública, que de alguna manera se anticipan como parte del proceso de adopción de los sistemas de inteligencia artificial, hacen que las reglas y los arreglos institucionales sean irrelevantes o inaceptables. **En este contexto de cambio, sin embargo, es innegable el llamado al diseño de instituciones que sean resilientes y robustas para reconciliar una lógica consecuencialista con una lógica de adecuación. La interacción de estas lógicas implica la necesidad de soluciones de IA más robustas que estén orientadas hacia los valores de la sociedad.**

Conciliar las lógicas de adecuación y consecuencialista de los sistemas de IA significa explorar diferentes lógicas de acción que están formal e informalmente prescritas, autorizadas y permitidas, y cómo definen, a su vez, lo que está prohibido e ilegítimo para diferentes actores en diferentes circunstancias (March y Olsen, 2009). Así como Timón fue un intérprete de la predicción hecha por la Pitia en el Oráculo de Delfos, los sistemas de IA dependen de una actividad de interpretación crítica que requiere, a su vez, una gramática institucional guiada por normas que dan sentido a la acción (Crawford y Ostrom, 1995). Es decir, el mayor desafío para los sistemas de IA en la administración pública es que su aplicación ocurre dentro de arreglos de gobernanza que son capaces de crear una interpretación correcta del proceso de decisión autónomo.



### 3. Gobernando los sistemas de IA

Conciliar la lógica de adecuación y la lógica consecuencialista significa diseñar instituciones que puedan establecer normas para guiar el comportamiento de los sistemas. En el caso de la aplicación de sistemas de IA, diseñar instituciones capaces de crear un cuerpo normativo que regule su acción es nuevo y desafiante. Esto implica construir un puente entre los diseñadores y los administradores públicos para que los sistemas de IA puedan impactar la administración pública de una manera correcta y válida, apoyados no solo por una lógica consecuencialista, sino también por valores y normas presupuestos.

Desde esta perspectiva, el diseño de instituciones requiere delinear los resultados específicos a alcanzar, determinar qué valores están incrustados en las prácticas a promover a través de ellas y discernir las reglas y normas necesarias para lograr el resultado deseado (Goodin, 1996). Las instituciones se diseñan en procesos de transformación para definir normas constituidas y reconstituidas por acciones frecuentes y repetitivas que se incorporan en diversas situaciones de la vida humana (Crawford y Ostrom, 1995). Al definir una gramática, las instituciones están creadas para promover los resultados de los esfuerzos implícitos o explícitos de un conjunto de individuos para lograr el orden y la previsibilidad en situaciones definidas mediante: 1) la creación de posiciones; 2) indicando cómo los participantes entran o salen de sus posiciones; 3) explicando qué acciones se requiere, se permite o se prohíbe a los participantes realizar; 4) especificando qué resultado se requiere, se permite o se prohíbe afectar a los participantes; y 5) estipulando las consecuencias de la violación de las reglas, que en la mayoría de los casos se espera que estén asociadas a una sanción específica (Crawford y Ostrom, 1995). En este sentido, el diseño de las instituciones debe reflejar las expectativas de la sociedad sobre los valores que deben alcanzarse y resolver los problemas de la acción colectiva (Ostrom, 2005).

Los sistemas de IA requieren el reconocimiento de los derechos humanos como una estructura de valores esencial para el mundo digital (Donahoe y Metzger, 2019). El desafío es converger y sincronizar entre los profesionales de IA las normas, estándares y estrategias que aseguren la mayor efectividad de los servicios y políticas públicas, el bienestar y el control sobre la posibilidad de mal

**Diseñar instituciones adecuadas para gobernar los sistemas de IA en la administración pública requiere reducir las asimetrías entre los diseñadores de sistemas y los gerentes públicos, así como involucrar a los usuarios en la toma de decisiones.**

uso de las tecnologías digitales (IEEE, 2019). Diseñar instituciones adecuadas para gobernar los sistemas de IA en la administración pública requiere reducir las asimetrías entre los diseñadores de sistemas y los gerentes públicos, así como involucrar a los usuarios en la toma de decisiones (Figueiras y Almeida, 2020). Así, el diseño de instituciones capaces de regular los sistemas de IA en la administración pública necesita de los elementos que se describen a continuación.

### **3.1 Toma de decisiones sobre la adopción de sistemas de IA**

Con estos valores en mente, es esencial crear estructuras de gobierno colaborativas, basadas en la participación de diferentes partes interesadas en el diseño de sistemas de IA. Una perspectiva de gobernanza colaborativa implica la construcción de un foro común en el que diferentes partes interesadas participan en la toma de decisiones orientada al consenso (Ansell y Gash, 2008). La colaboración apoya procesos de toma de decisiones más abiertos, basados en el consenso entre los actores. Asimismo, genera confianza entre los participantes y una comprensión compartida de los problemas, las reglas y los resultados previstos. El liderazgo político, en este sentido, juega un papel esencial en la creación de adhesión y participación de la sociedad.

Además de la gobernanza colaborativa, la coproducción de la adopción de sistemas de IA también puede afectar el diseño de los sistemas que se inserta en disputas comerciales, de justicia, éticas, multijurisdiccionales e internacionales. La adopción de la gobernanza colaborativa para el diseño de sistemas tiene el potencial de reducir disputas y conflictos sobre los procedimientos y técnicas que se adoptarán, con el fin de ampliar la rendición de cuentas y un mayor aprendizaje para enfrentar acciones preventivas contra sesgos, injusticias y vigilancia de sistemas (Amsler, Martínez y Smith, 2020).

### **3.2 Transparencia algorítmica y responsabilidad**

Los sistemas de IA utilizan diferentes tipos de algoritmos, impactando directamente en la vida cotidiana de la sociedad. En consecuencia, la transparencia algorítmica y la responsabilidad son requisitos esenciales y desafiantes para la gobernanza de la IA. Al simular el pensamiento humano, los sistemas de inteligencia artificial utilizan datos previos y el historial de decisiones para producir una nueva decisión. Como el pasado es ambiguo y las decisiones se toman en

el contexto de incertidumbres (March y Olsen, 1975), los algoritmos pueden reforzar los sesgos sociales constituyendo injusticias (Noble, 2018; Benjamin, 2019), distribuyendo los recursos de manera desigual (Eubanks, 2018) y reforzando el carácter tecnocrático de la administración pública (Janssen y Kuk, 2016). En el contexto de ambigüedades, donde los sistemas de IA ejercen funciones discrecionales en la administración pública, es necesario establecer mecanismos institucionales para promover la transparencia y la rendición de cuentas. Los algoritmos son “cajas negras” y pueden obstaculizar la confianza de la sociedad, generando la reproducción de problemas de gobernanza sin tecnologías digitales (Shull, 2019). La transparencia algorítmica es el elemento mediante el cual los ciudadanos pueden conocer cómo los sistemas de decisión autónomos toman decisiones que impactan en sus vidas (Diakopoulos, 2016; Hildebrandt, 2012).

La transparencia algorítmica es un desafío e implica diferentes condiciones cognitivas entre los diseñadores de sistemas y los usuarios que se ven afectados por estos sistemas (Fine Licht y Fine Licht, 2020). Tiene límites en las posibilidades del secreto, tanto en los gobiernos como en las empresas, y en las capacidades cognitivas de los ciudadanos para crear comprensión (Crain, 2018; Ananny y Crawford, 2018). Estos problemas cognitivos crean dificultades e implican desafíos para la transparencia algorítmica, requiriendo que no se “abra la caja negra” completa, sino que existan procesos de justificación de decisiones en la adopción de la IA, generando una percepción de legitimidad por parte de los usuarios y afectados por las acciones de los sistemas (Fine Licht y Fine Licht, 2020). Para ampliar el potencial de la transparencia algorítmica, los intermediarios del conocimiento pueden jugar un papel fundamental como intermediarios entre los diseñadores de sistemas y los gestores públicos y los ciudadanos (Xu, Ramanathan y Ramnath, 2014).

La transparencia es esencial para generar un sentido de responsabilidad. Además del derecho a saber, los sistemas de inteligencia artificial requieren estructuras de responsabilidad institucional. Debido a que afectan el proceso de toma de decisiones e influyen en la vida de los ciudadanos, estos deben asumir la responsabilidad y el juicio sobre las decisiones tomadas por los sistemas de IA que perjudican el interés público. Los funcionarios administrativos insertan sistemas de inteligencia artificial en cuestiones normativas y de cumplimiento con reglas y legitimidad de las decisiones, o incluso sobre las decisiones de los gerentes con respecto a la gobernanza

de datos (Binns, 2018). Debido a que estos sistemas ejercen una función discrecional o tecnocrática en la administración pública, necesitan brindar retroalimentación, aprendiendo de la experiencia de la administración y con el consentimiento informado de los ciudadanos (Olsen, 2017). La rendición de cuentas se inserta en un complejo proceso de justificación y juicio de los actores ante un foro (Bovens, 2007). En este contexto, los sistemas de IA representan actores con poder de decisión, pero no representan a una persona que pueda ser considerada responsable o justificada en el foro. La principal limitación de la responsabilidad algorítmica es el hecho de que es posible que el foro no siempre pueda juzgar las acciones del agente. Como ya se ha dicho para las preocupaciones relacionadas con la transparencia algorítmica, los algoritmos son dinámicos y están compuestos por varias capas de sistemas basados en un orden técnico. La responsabilidad algorítmica es un asunto de expertos y este trabajo implica reducir el alcance del foro en el proceso de promover la responsabilidad del algoritmo junto con sus diseñadores y desarrolladores. La responsabilidad aplicada a los algoritmos está sujeta a controversias y escrutinio sobre los valores y principios involucrados; requiere permanentemente la justificación de su uso para el bien público (Binns, 2018).

### **3.3 Gobernanza de Datos**

Los sistemas de IA dependen de grandes volúmenes de datos para poder ampliar su capacidad predictiva (Domingos, 2015). Esto implica que la adopción de sistemas de IA en una administración pública depende de la capacidad de un gobierno para recopilar, almacenar, procesar y compartir datos entre las diferentes organizaciones que lo componen (Salvador y Ramió, 2020). Esta capacidad requiere, a su vez, una autoridad relacionada con la gobernanza de datos con el fin de proteger la privacidad de los ciudadanos, garantizar el uso seguro y, simultáneamente, proporcionar recursos para que los sistemas de IA puedan funcionar (Plotkin, 2013; Benfeldt, Persson y Madsen, 2020).

La gobernanza de datos es un recurso esencial para mejorar la construcción de soluciones de IA confiables, reduciendo los sesgos sistémicos, decisiones ilegales, crisis políticas y exposiciones financieras de los ciudadanos (Janssen ...[et al], 2020). Requiere estructuras institucionales sólidas para establecer una forma segura y confiable de administrar datos, un aspecto fundamental de la gobernanza, y

La gobernanza de datos requiere también que los administradores públicos sean cada vez más cuidadosos con la gestión del conocimiento, posibilitando no solo la expansión del *big data*, sino también la calificación de soluciones de inteligencia artificial que informan un nuevo estándar de la administración pública.

que sean responsables de garantizar su calidad, validez y seguridad, administrar los riesgos y hacer que las organizaciones públicas manejen la información con cuidado e integridad (Dawes, 2010).

La gobernanza de datos requiere también que los administradores públicos sean cada vez más cuidadosos con la gestión del conocimiento (Aguilar, 2020), posibilitando no solo la expansión del *big data*, sino también la calificación de soluciones de inteligencia artificial que informan un nuevo estándar de la administración pública (Ramió, 2018 y 2019).

### 3.4 Validez y carácter explicable de los sistemas de IA

El principal desafío para la constitución de sistemas de IA es establecer su validez como instrumento de toma de decisiones y que sus procedimientos sean explicables. Se trata de cuestiones técnicas que convergen en el mismo problema de asimetría entre diseñadores de sistemas, formuladores de políticas y usuarios. Establecer la validez de los sistemas de IA significa pensar en múltiples cuestiones técnicas: 1) ¿se ha formulado correctamente el objetivo?; 2) ¿el sistema de IA está libre de errores de *software*?; 3) ¿el sistema de IA se basa en datos adecuadamente representativos?; 4) ¿puede el sistema de IA hacer frente a anomalías y fallas de datos inevitables?; y 5) ¿es el sistema de inteligencia artificial lo suficientemente preciso? (Hand y Khan, 2020). Las cuestiones técnicas son importantes cuando se trata de la validez de la herramienta digital, buscando la evaluación y análisis continuo de riesgos y cumplimiento (Freeman, 2020). Sin embargo, como los sistemas de IA en la administración pública representan tecnologías de poder, es fundamental cuestionar no solo su carácter técnico, sino también la legitimidad de la herramienta para resolver los problemas que afectan a la sociedad. Este proceso involucra entonces no solo las cuestiones técnicas, sino también políticas.

Cuestionar una solución de IA significa movilizar criterios hacia una lógica de adecuación, en términos democráticos. El desafío es encajar diferentes motivaciones y lógicas de acción en un proceso de coproducción de sistemas de IA para que estos encuentren reglas y normas, formales e informales, que permitan evaluar la experiencia y los resultados obtenidos. El proceso de coproducción debe observar la pluralidad de valores y resultados obtenidos en términos de las interpretaciones de las partes interesadas (Osborne, Strokosch y Radnor, 2018). La validez de los sistemas de IA depende de una

motivación compleja que permite realizar pruebas basadas en valores plurales insertados en el proceso de interpretación de los resultados para anticipar cualquier problema en la adopción de tecnologías.

La gobernanza de la IA depende de un diseño de instituciones que consideren establecer normas para conciliar la lógica consecuencialista con una lógica de adecuación y producir tecnologías inmersas en normas y reglas democráticas, basadas en valores de cooperación. No se limita al aspecto regulatorio. La regulación es esencial para abordar cuestiones relacionadas con la privacidad de los ciudadanos, el control de los mecanismos de vigilancia y los incentivos para el uso adecuado de los datos. Pero la gobernanza de la IA también forma parte de las relaciones de poder, ya que los sistemas de IA toman decisiones discrecionales que afectan la vida de los ciudadanos, por lo que requiere de un diseño institucional complejo donde conviven diferentes estilos de políticas, mecanismos de mando y control regulatorios convergentes con la participación de redes, mecanismos de autogobierno y formas interactivas con la sociedad (Filgueiras y Almeida, 2020).

### **3.5 Gobernanza algorítmica**

Los algoritmos son secuencias de instrucciones o pasos para resolver un problema o tarea en particular. Los seres humanos los han puesto a trabajar en todo tipo de procesos de toma de decisiones durante siglos. El uso de algoritmos en la gobernanza no es nuevo (por ejemplo, conjuntos predeterminados de reglas para la resolución de problemas y rutinas preestablecidas para la toma de decisiones del gobierno). Desde el advenimiento de la “análisis de *big data*”, se ha implementado un amplio espectro de algoritmos para guiar el análisis de datos y los resultados de gobernanza en la sociedad, que van desde rutinas que están completamente programadas por seres humanos hasta sistemas de inteligencia artificial que aprenden de *big data* y toman decisiones de gobernanza de forma autónoma.

En 1959, Arthur Samuel definió el aprendizaje automático como un “campo de estudio que brinda a las computadoras la capacidad de aprender sin estar programadas explícitamente”. Los algoritmos de aprendizaje automático son programas que “aprenden de la experiencia” y la exposición a los datos (Samuel, 1959), y que logran buenos resultados cuando son “supervisados” o “entrenados” en un conjunto de datos predefinido que incorpora ejemplos de entrenamiento. Además, los algoritmos de aprendizaje automático

“no supervisados” pueden encontrar patrones y relaciones en conjuntos de datos relacionados con un problema específico sin estar capacitados. Con conjuntos de datos “más grandes” e incluso más “generalizados”, el aprendizaje automático se vuelve menos dependiente de los algoritmos codificados por humanos que dirigen los resultados y avanza hacia el objetivo de la inteligencia artificial general. El término inteligencia artificial, según la definición de Herbert Simon, se refiere en particular a los fenómenos que aparecen cuando las computadoras realizan tareas que, si las realizaran personas, requieren un pensamiento inteligente (Simon, 1969). Los algoritmos de aprendizaje automático son un subcampo de la IA.

Por ejemplo, la aplicación del rastreo de teléfonos móviles durante la pandemia de SARS-COV2 en China buscó formas de monitorear y vigilar el desapego social para prevenir un contagio rápido y aplanar la curva de contaminación. Tales medidas pueden cometer injusticias y crear zonas de exclusión. Los resultados de acciones gubernamentales como esta dependen de toda la cadena de comandos, basada en los datos de entrada, para constituir la solución al problema propuesto (Servick, 2020). Una preocupación aquí es cuán imparcial es realmente esta lógica programada y si respeta los derechos humanos y la calidad de los datos de entrada para que pueda tener soluciones efectivas y reales. De repente, los algoritmos son fundamentales para el proceso de políticas. Asimismo, en toda transformación digital, los algoritmos son fundamentales para la gobernanza. Los algoritmos hacen que las soluciones sean reales, incluso si son parciales o con fallas de gobierno.

Los gobiernos pueden utilizar sistemas algorítmicos de decisión en las políticas públicas en diferentes áreas, como educación, salud, sistemas de justicia penal y políticas sociales (Gillingham, 2015). Las herramientas digitales pueden cambiar el orden social mediante el análisis de los incentivos otorgados a varios actores políticos, que pueden predecirse mediante algoritmos, cumpliendo importantes funciones políticas (Danaher, 2016). Los algoritmos pueden filtrar y dirigir información para cambiar las percepciones de los ciudadanos, creando un contexto de mayor o menor confianza en el funcionamiento de las instituciones. Los algoritmos afectan la cultura, modifican el conocimiento y construyen realidades (Just y Latzer, 2017; Napoli, 2014).

La tendencia hacia la gobernanza algorítmica plantea muchos desafíos con respecto al sector privado y la toma de decisiones pública.

La gobernanza del algoritmo es un componente del proceso digital, puede variar desde los puntos de vista estrictamente legales y regulatorios hasta un punto de vista puramente técnico. La opacidad inherente y la falta de transparencia en la gobernanza algorítmica pueden tener consecuencias significativas para la agencia humana y la responsabilidad. Lo esencial es que la gobernanza algorítmica pueda ser capaz de inculcar en el diseño de sistemas de inteligencia artificial, especialmente para la administración pública, valores que conduzcan al desarrollo de sistemas autónomos de decisión basados en derechos humanos (Donahoe y Metzger, 2019), en parámetros éticos que orientan la práctica profesional de los analistas de datos (IEEE, 2019) y a los gestores públicos sobre los riesgos, problemas y desafíos que rodean la adopción de la inteligencia artificial en el sector público (Filgueiras y Almeida, 2020).

### **Conclusiones**

En el Templo de Apolo existían 147 inscripciones que buscaban orientar el trabajo de los consultantes cuando preguntaban a la Pitia. Algunas de estas inscripciones especifican valores diferentes como conocerse a sí mismo, obedecer la ley, escuchar y comprender, conocer su oportunidad, saber cuándo es un extraño, ayudar a sus amigos, perseguir el honor, elogiar la virtud, ser imparcial, ser justo en el habla y debatir, evitar hacer nada en exceso, mirar hacia el futuro, no hacer promesas arriesgadas, hablar desde el conocimiento. Las inscripciones del Templo de Apolo buscaban guiar al consultante con respecto a los valores básicos que deberían guiar las acciones, independientemente de las predicciones de la Pitia. Se esperaba que los consultantes interpretaran y absorbieran estos valores básicos antes de ingresar al templo para que la predicción pudiera ser válida y efectiva.

En el mundo actual los requisitos para el uso de tecnologías no son tan confiables ni tan claros en el proceso de toma de decisiones de una organización determinada. Los ejemplos de fallos en la adopción de sistemas de IA son variados y recurrentes en la administración pública. Sin embargo, una cosa es cierta: el engaño o la desconfianza con respecto a las predicciones realizadas por los sistemas de IA compromete la legitimidad de su aplicación en tareas y procesos de toma de decisiones. Teniendo esto en cuenta, existen tareas para mejorar el diseño de los sistemas de IA que incluyen la disponibilidad de más datos, la mejora de los algoritmos y la constitución de



**Los mayores desafíos para la gobernanza de la IA son establecer instituciones que puedan ensamblar y alinear los comportamientos correctos por parte de los diseñadores y los responsables políticos, de modo que la adopción de la IA en la administración pública se produzca en contextos más abiertos y responsables para los ciudadanos.**

parámetros éticos sobre el trabajo de los diseñadores. Sin embargo, también es necesario diseñar y crear instituciones que puedan establecer normas para guiar democráticamente a los actores en el proceso de diseño y adopción de la IA.

Al final, los mayores desafíos para la gobernanza de la IA son establecer instituciones que puedan ensamblar y alinear los comportamientos correctos por parte de los diseñadores y los responsables políticos, de modo que la adopción de la IA en la administración pública se produzca en contextos más abiertos y responsables para los ciudadanos. Las inscripciones en el Templo de Apolo sirven como ejemplo de cómo una lógica de adecuación puede ser central en el proceso de institucionalización, recordando los valores que los consultantes deben respetar para que las predicciones de la Pitia puedan volverse concretas (aunque ambiguas) y, por lo tanto, sujetas a diferentes interpretaciones. La IA puede ser un recurso valioso para fortalecer la administración pública, pero debe tener lugar dentro de un contexto institucional en el que los mecanismos de gobernanza pueden operar de manera transparente y responsable, y donde deben converger una lógica consecuencialista de herramientas digitales con un proceso de toma de decisiones más democrático y sostenible.

En este artículo se discute el papel que juegan los sistemas de IA en la transformación digital de la administración pública. Los sistemas de IA amplían las capacidades de la administración pública y tienen el potencial de cambiar el desempeño institucional de los servicios públicos y de las políticas públicas, promoviendo una mayor conexión entre gobiernos y ciudadanos, anticipando demandas, automatizando rutinas, desempeñando funciones discrecionales, entre otros. La adopción de sistemas de IA significa utilizar tecnologías que sean capaces de aprender a tomar decisiones y resolver problemas.

La adopción de sistemas de IA puede entenderse como un proceso técnico. Sin embargo, este artículo abordó la teoría institucional para demostrar que las decisiones en la administración pública no son solo procesos técnicos. La toma de decisiones en la administración pública no constituye apenas una lógica consecuencialista, esperando cambios en la relación entre costo y beneficio, sino que es un problema político basado en la racionalidad limitada de diferentes participantes en un contexto institucional, quienes pueden participar de manera fluida y establecer relaciones para movilizar colectivamente valores, normas, inquietudes éticas, entendimientos y roles diferentes.

La consecuencia de este proceso de toma de decisiones para el estudio de la IA en la administración pública es que el diseño de sistemas debe considerar las distintas lógicas del funcionamiento institucional, que comprenden un conjunto complejo de estándares de gobernanza para alinear una perspectiva de resultados con una perspectiva de valores. La gobernanza de la IA es un proceso desafiante que requiere la creación de instituciones para hacer frente a los diferentes problemas relacionados con el diseño de sistemas, que comprende un conjunto de normas y reglas que deben guiar el comportamiento de los diseñadores de sistemas de IA. Partiendo de la premisa de que el diseño de sistemas de IA es un problema político y debe considerar contextos institucionales, las políticas que orientan y regulan el diseño deben conciliar una lógica consecuencialista con una lógica de adecuación. Se han realizado esfuerzos en este sentido desde organizaciones profesionales como el IEEE (2019) o requeridos por organismos internacionales como la OECD (2019) y las Naciones Unidas (United Nations, 2019).

Una implicación importante es comprender que el diseño de los sistemas de IA y su adopción en la administración pública dependen del diseño de las instituciones de gobernanza de la IA. Estas deben ser capaces de dirigir políticas que movilicen instrumentos apropiados que alineen la lógica consecuencialista con la lógica de adecuación, permitiendo la conexión entre el propósito de la transformación digital y los valores políticos más amplios de democracia y derechos humanos.

## Notas

(1) Es interesante ver cómo el premio Nobel Herbert Simon sentó las bases de la teoría institucional que cubre temas de ciencia política y administración pública y, al mismo tiempo, es uno de los padres fundadores de la tecnología de inteligencia artificial, habiendo participado en el Seminario de Dartmouth College en 1956.

## Bibliografía

- Aguilar, L. F. (2020), *Gestión del conocimiento y gobierno contemporáneo*, México, Amazon Books.
- Alkhatib, A. y Bernstein, M. (2019), "Street-Level Algorithms: a Theory at the Gaps between Policy and Decisions", en *Proceedings of the 2019 CHI Conference on Human Factors in Computing Society Systems Proceedings*, New York, Association for Computing Machinery. Paper presented at the 2019 CHI Conference, Glasgow, Scotland, United Kingdom, May 4-9 (Paper N° 530), <https://doi.org/10.1145/3290605.3300760>.

- Almeida, V.; Filgueiras, F.; y Gaetani, F. (2020), "Digital Governance and Tragedy of Commons", en *IEEE Internet Computing*, Vol. 24 N° 4, pp. 41-46, <https://doi.org/10.1109/MIC.2020.2979639>.
- Amsler, L. B.; Martinez, J. K.; y Smith, S. E. (2020), *Dispute System Design: Preventing, Managing, and Resolving Conflict*, Stanford, Stanford University Press.
- Ananny, M. y Crawford, K. (2018), "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability", en *New Media and Society*, Vol. 20 N° 3, pp. 973-989, <https://doi.org/10.1177/1461444816676645>.
- Andrews, L. (2018), "Public Administration, Public Leadership and the Construction of Public Value in the Age of the Algorithm and 'Big Data'", en *Public Administration*, Vol. 97 N° 2, pp. 296-310, <https://doi.org/10.1111/padm.12534>.
- Ansell, C. y Gash, A. (2008), "Collaborative Governance in Theory and Practice", en *Journal of Public Administration, Research and Theory*, Vol. 18 N° 4, pp. 543-571, <https://doi.org/10.1093/jopart/mum032>.
- Benfeldt, O.; Persson, J. S.; y Madsen, S. (2020), "Data Governance as a Collective Action Problem", en *Information Systems Frontiers*, Vol. 22 N° 2, pp. 299-313, <https://doi.org/10.1007/s10796-019-09923-z>.
- Benjamin, R. (2019), *Race after Technology. Abolitionist Tools for the New Jim Code*, Cambridge, Polity Books.
- Binns, R. (2018), "Algorithmic Accountability and Public Reason", en *Philosophy and Technology*, Vol. 31 N° 4, pp. 543-556, <https://doi.org/10.1007/s13347-017-0263-5>.
- Bovens, M. (2007), "Analysing and Assessing Accountability: a Conceptual Framework", en *European Law Journal*, Vol. 13 N° 4, pp. 447-468, <https://doi.org/10.1111/j.1468-0386.2007.00378.x>.
- Chen, Y. C. (2017), *Managing Digital Governance: Issues, Challenges and Solutions*, New York, Routledge.
- Chen, Y. F. y Hsieh, T. C. (2014), "Big Data for Digital Government: Opportunities, Challenges, and Strategies", en *International Journal of Public Administration in the Digital Age*, Vol. 1 N° 1, pp. 1-14, <https://doi.org/10.4018/ijpada.2014010101>.
- Cohen, M. D.; March, J. G.; y Olsen, J. P. (1972), "A Garbage Can Model of Organizational Choice", en *Administrative Science Quarterly*, Vol. 17 N° 1, pp. 1-25, <https://doi.org/10.2307/2392088>.

- Collier, Matt; Fu, Richard; y Yin, Lucy (2017), "Artificial Intelligence: Healthcare's New Nervous System", Dublin, Accenture, [https://www.accenture.com/t20170418T023052Z\\_w\\_/au-en/\\_acnmedia/PDF-49/Accenture-Health-Artificial-Intelligence.pdf](https://www.accenture.com/t20170418T023052Z_w_/au-en/_acnmedia/PDF-49/Accenture-Health-Artificial-Intelligence.pdf).
- Couldry, N. y Mejias, U. (2019), "Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject", en *Television and New Media*, Vol. 20 N° 4, pp. 336-349, <https://doi.org/10.1177/1527476418796632>.
- Crain, M. (2018), "The Limits of Transparency: Data Brokers and Commodification", en *New Media and Society*, Vol. 20 N° 1, pp. 88-104, <https://doi.org/10.1177/1461444816657096>.
- Crawford, S. y Ostrom, E. (1995), "A Grammar of Institutions", en *American Political Science Review*, Vol. 89 N° 3, pp. 582-600, <https://doi.org/10.2307/2082975>.
- Danaher, J. (2016), "The Threat of Algocracy: Reality, Resistance and Accommodation", en *Philosophy and Technology*, Vol. 29 N° 3, pp. 245-268, <https://doi.org/10.1007/s13347-015-0211-1>.
- Danaher, J.; Hogan, M. J.; Noone, C.; Kennedy, R.; Behan, A.; De Paor, A.; Felzman, H.; Haklay, M.; Khoo, S. M.; Morison, J.; Murphy, M. H.; O'Brolchain, N.; Schafer, B.; y Shankar, K. (2017), "Algorithmic Governance: Developing a Research Agenda through the Power of Collective Intelligence", en *Big Data and Society*, Vol. 4 N° 2, pp. 1-21, <https://doi.org/10.1177/2053951717726554>.
- Dawes, S. S. (2010), "Stewardship and Usefulness: Policy Principles for Information-Based Transparency", en *Government Information Quarterly*, Vol. 27 N° 4, pp. 377-383, <https://doi.org/10.1016/j.giq.2010.07.001>.
- Diakopoulos, N. (2016), "Accountability in Algorithmic Decision Making", en *Communications of the ACM*, Vol. 59 N° 2, pp. 56-62, <https://doi.org/10.1145/2844110>.
- Domingos, Pedro (2015), *The Master Algorithm. How the Quest for Ultimate Machine Learning Will Remake Our World*, New York, Basic Books.
- Donahoe, E. y Metzger, M. M. (2019), "Artificial Intelligence and Human Rights", en *Journal of Democracy*, Vol. 30 N° 2, pp. 115-126, <https://doi.org/10.1353/jod.2019.0029>.
- Dunleavy, P. (2016), "Big Data and Policy Learning", en *Evidence-Based Policy Making in the Social Sciences. Methods that Matter*, G. Stoker y M. Evans (eds.), Chicago, Policy Press.
- Dunleavy, P. y Margetts, H. (2013), "The Second Wave of Digital-Era Governance: a Quasi-Paradigm for Government on the Web", en *Philosophical Transactions of the Real Society*, Vol. 371 Issue 1987, pp. 1-17, <https://doi.org/10.1098/rsta.2012.0382>.

- Dunleavy, P.; Margetts, H.; Bastow, S.; y Tinkler, J. (2006), "New Public Management Is Dead-Long Live Digital-Era Governance", en *Journal of Public Administration Research and Theory*, Vol. 16 N° 3, pp. 467-494, <https://doi.org/10.1093/jopart/mui057>.
- Eubanks, V. (2018), *Automating Inequality. How High-Tech Tools Profile, Police, and Punish the Poor*, New York, St. Martin's Press.
- Filgueiras, F. y Almeida, V. (2020), *Governance for the Digital World. Neither More State nor More Market*, London, Palgrave.
- Filgueiras, F.; Cireno, F.; y Palotti, P. (2019), "Digital Transformation and Public Service Delivery in Brazil", en *Latin American Policy*, Vol. 10 N° 2, pp. 195-219, <https://doi.org/10.1111/lamp.12169>.
- Fine Licht, K. y Fine Licht, J. (2020), "Artificial Intelligence, Transparency, and Public Decision-Making", en *AI and Society*, Vol. 35 N° 4, pp. 917-926, <https://doi.org/10.1007/s00146-020-00960-w>.
- Flach, P. (2012), *Machine Learning: the Art and Science of Algorithms that Make Sense of Data*, New York, Cambridge University Press.
- Freeman, L. (2020), "Testing and Evaluating Artificial Intelligence", en *Insights*, Vol. 23 N° 1, pp. 27-30, <https://doi.org/10.1002/inst.12281>.
- Frischmann, B. y Selinger, E. (2018), *Re-Engineering Humanity*, Cambridge, Cambridge University Press.
- Gillingham, P. (2015), "Predictive Risk Modelling to Prevent Child Maltreatment and other Adverse Outcomes for Service Users: Inside the 'Black Box' of Machine Learning", en *The British Journal of Social Work*, Vol. 46 N° 4, pp. 1044-1058, <http://10.1093/bjsw/bcv031>.
- Goodin, R. E. (1996), "Institutions and Their Design", en *The Theory of Institutional Design*, R. E. Goodin (ed.), Cambridge, Cambridge University Press.
- Greengard, S. (2015), *The Internet of Things*, Cambridge, MIT Press.
- Hand, D. J. y Khan, S. (2020), "Validating and Verifying AI Systems", en *Patterns*, Vol. 1 N° 3, pp. 1-3, <https://doi.org/10.1016/j.patter.2020.100037>.
- Hildebrandt, M. (2012), "The Dawn of a Critical Transparency Right for the Profiling Era", en *Digital Enlightenment Yearbook 2012*, Jacques Bus, Malcolm Crompton, Mireille Hildebrandt y George Metakides (eds.), Clifton, IOS Press, pp. 41-56, <https://doi.org/10.3233/978-1-61499-057-4-41>.
- Hood, C. y Margetts, H. (2007), *The Tools of Government in the Digital Age*, London, Palgrave Macmillan.

- IEEE (2019), *Ethically Aligned Design*, Piscataway, Institute of Electrical and Electronics Engineers. The Global Initiative on Ethics of Autonomous and Intelligent Systems, <https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/autonomous-systems.html>.
- Janssen, M.; Brous, P.; Estevez, E.; Barbosa, L. S.; y Janowski, T. (2020), "Data Governance: Organizing Data for Trustworthy Artificial Intelligence", en *Government Information Quarterly*, Vol. 37 N° 3, pp. 1-8, <https://doi.org/10.1016/j.giq.2020.101493>.
- Janssen, M.; Charalabidis, Y.; y Zuiderwijk, A. (2012), "Benefits, Adoption Barriers and Myths of Open Data and Open Government", en *Information Systems Management*, Vol. 29 N° 4, pp. 258-268, <https://doi.org/10.1080/10580530.2012.716740>.
- Janssen, M. y Kuk, G. (2016), "The Challenges and Limits of Big Data Algorithms in Technocratic Governance", en *Government Information Quarterly*, Vol. 33 N° 3, pp. 371-377, <https://doi.org/10.1016/j.giq.2016.08.011>.
- Jefferies, S. (2021), *The Virtual Public Servant. Artificial Intelligence and Frontline Work*, London, Palgrave.
- Jefferies, Duncan (2016), "The Automated City: Do We Still Need Humans to Run Public Services?", en *The Guardian*, London, September 20, <https://www.theguardian.com/cities/2016/sep/20/automated-city-robots-run-public-services-councils>, 02-07-2018.
- Just, N. y Latzer, M. (2017), "Governance by Algorithms: Reality Construction by Algorithmic Selection on the Internet", en *Media Culture and Society*, Vol. 39 N° 2, pp. 238-258, <https://doi.org/10.1177/0163443716643157>.
- Kingdon, J. (1995), *Agendas, Alternatives and Public Policies*, New York, Harper Collins.
- Kosters, M. y Van der Heijden, J. (2015), "From Mechanism to Virtue: Evaluating Nudge Theory", en *Evaluation*, Vol. 21 N° 3, pp. 276-291, <https://doi.org/10.1177/1356389015590218>.
- Linders, D. (2012), "From E-Government to We-Government: Defining a Typology for Citizen Coproduction in the Age of Social Media", en *Government Information Quarterly*, Vol. 29 N° 4, pp. 446-454, <https://doi.org/10.1016/j.giq.2012.06.003>.
- Luna-Reyes, L. F. y Gil-García, J. R. (2014), "Digital Government Transformation and Internet Portals: the Co-Evolution of Technology, Organizations, and Institutions", en *Government Information Quarterly*, Vol. 31 N° 4, pp. 545-555, <https://doi.org/10.1016/j.giq.2014.08.001>.

- March, J. G. (1991), "How Decisions Happen in Organizations", en *Human-Computer Interaction*, Vol. 6 N° 2, pp. 95-117, [https://doi.org/10.1207/s15327051hci0602\\_1](https://doi.org/10.1207/s15327051hci0602_1).
- March, J. G. y Olsen, J. P. (1975), "The Uncertainty of Past: Organizational Learning under Ambiguities", en *European Journal of Political Research*, Vol. 3 N° 2, pp. 147-171, <https://doi.org/10.1111/j.1475-6765.1975.tb00521.x>.
- \_\_\_\_\_ (1976), *Ambiguity and Choice in Organizations*, Bergen, Universitetsforlaget.
- \_\_\_\_\_ (2009), "The Logic of Appropriateness", en *The Oxford Handbook of Political Science*, R. E. Goodin (ed.), Oxford, Oxford University Press.
- Margetts, H. y Dorobantu, C. (2019), "Rethinking Government with AI", en *Nature*, N° 568, pp. 163-165, <https://doi.org/10.1038/d41586-019-01099-5>.
- Markram, H. (2006), "The Blue Brain Project", en *Nature Reviews Neuroscience*, Vol. 7 N° 2, pp. 153-160.
- Mehr, H. (2017), *Artificial Intelligence for Citizen Services and Government*, Cambridge, Harvard Kennedy School. Ash Center for Democratic Governance and Innovation, [https://ash.harvard.edu/files/ash/files/artificial\\_intelligence\\_for\\_citizen\\_services.pdf](https://ash.harvard.edu/files/ash/files/artificial_intelligence_for_citizen_services.pdf).
- Meijer, A. J. (2012), "Co-Production in an Information Age: Individual and Community Engagement Supported by New Media", en *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, Vol. 23 N° 4, pp. 1156-1172, <https://doi.org/10.1007/s11266-012-9311-z>.
- Mergel, I.; Edelman, N.; y Haug, N. (2019), "Defining Digital Transformation: Results from Experts Interviews", en *Government Information Quarterly*, Vol. 36 N° 4, <https://doi.org/10.1016/j.giq.2019.06.002>.
- Milakovich, M. E. (2012), "Anticipatory Government: Integrating Big Data for Smaller Government", Oxford, University of Oxford. Oxford Internet Institute. The Internet, Policy and Politics Conferences, <http://blogs.oii.ox.ac.uk/ipp-conference/sites/ipp/files/documents/Anticipatory%20Government%20Oxford.pdf>.
- Minsky, M. (1985), *The Society of Mind*, New York, Simon and Schuster.
- Napoli, P. M. (2014), "Automated Media: an Institutional Theory Perspective on Algorithmic Media Production and Consumption", en *Communication Theory*, Vol. 24 N° 3, pp. 340-360, <https://doi.org/10.1111/comt.12039>.

- Newell, A. y Simon, H. A. (1961), "Computer Simulation and Human Thinking", en *Science*, Vol. 134 N° 3495, pp. 2011-2017, <https://doi.org/10.1126/science.134.3495.2011>.
- Noble, S. U. (2018), *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York, New York University Press.
- OECD (2019), "Data in the Digital Age", Paris, OECD Publishing (OECD Going Digital Policy Note), <http://www.oecd.org/going-digital/data-in-the-digital-age.pdf>.
- Olsen, J. P. (2017), *Democratic Accountability, Political Order, and Change: Exploring Accountability Processes in an era of European Transformation*, Oxford, Oxford University Press.
- O'Malley, M. (2014), "Doing What Works: Governing in the Age of Big Data", en *Public Administration Review*, Vol. 74 N° 5, pp. 555-55, <https://doi.org/10.1111/puar.12260>.
- O'Neil, C. (2016), *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York, Crown.
- Osborne, S.; Strokosch, K.; y Radnor, Z. (2018), "Co-Production and Co-Creation of Value in Public Services. A Perspective from Service Management", en *Co-Production and Co-Creation, Engaging Citizens in Public Service*, Taco Brandsen, Trui Steen y Bram Verschuere (eds.), New York, Routledge, pp. 18-26.
- Ostrom, Elinor (2005), *Understanding Institutional Diversity*, Princeton, Princeton University Press.
- Pagnamenta, R. (2020), "The A-Levels Fiasco Has Revealed the Socially Toxic Consequences of Algorithmic Bias", en *The Telegraph*, London, August 17, <https://www.telegraph.co.uk/technology/2020/08/17/a-levels-fiasco-has-revealed-socially-toxic-consequences-algorithmic>.
- Peters, B. G. (2018), *Policy Problems and Policy Design*, Cheltenham, Edward-Elgar Publishing.
- \_\_\_\_\_ (2019), *Institutional Theory in Political Science: the New Institutionalism*, Cheltenham, Edward-Elgar Publishing.
- Plotkin, D. (2013), *Data Stewardship: an Actionable Guide to Effective Data Management and Data Governance*, Burlington, Morgan Kauffman.
- Power, D. J. (2016), "Big Brother' Can Watch Us", en *Journal of Decision Systems*, Vol. 25 Issue Sup1, pp. 578-588, <https://doi.org/10.1080/12460125.2016.1187420>.
- Ramió, C. (2018), "Inteligencia artificial, robótica y modelos de administración pública", en *Revista del CLAD Reforma y Democracia*, N° 72, pp. 5-42.



- \_\_\_\_\_ (2019), *Inteligencia artificial y administración pública: robots y humanos compartiendo el servicio público*, Madrid, Catarata.
- Russell, S. (1997), "Rationality and Intelligence", en *Artificial Intelligence*, Vol. 94 Nos. 1-2, pp. 57-77, [https://doi.org/10.1016/S0004-3702\(97\)00026-X](https://doi.org/10.1016/S0004-3702(97)00026-X).
- Russell, S. y Norvig, P. (2010), *Artificial Intelligence: a Modern Approach*, Englewood Cliffs, Prentice Hall.
- Salvador, M. y Ramió, C. (2020), "Capacidades analíticas y gobernanza de datos en la administración pública como paso previo a la introducción de la inteligencia artificial", en *Revista del CLAD Reforma y Democracia*, N° 77, pp. 5-36.
- Samuel, A. (1959), "Some Studies in Machine Learning Using the Game of Checkers", en *IBM Journal of Research and Development*, Vol. 44 Nos. 1-2, pp. 210-229, <https://doi.org/10.1147/rd.33.0210>.
- Scharpf, F. (1988), "The Joint Decision Trap: Lessons from German Federalism and European Integration", en *Public Administration Review*, Vol. 66 N° 3, pp. 239-278, <https://doi.org/10.1111/j.1467-9299.1988.tb00694.x>.
- Servick, K. (2020), "Cellphone Tracking Could Help Stem the Spread of Coronavirus. Is Privacy the Price?", en *Science*, Washington, March 22, <https://doi.org/10.1126/science.abb8296>.
- Shull, A. (2019), "Governing Cyberspace During a Crisis in Trust", en *Governing Cyberspace During a Crisis in Trust*, Waterloo, Centre for International Governance Innovation, pp. 4-8.
- Simon, H. A. (1957), *Models of Man: Social and Rational*, New York, John Wiley.
- \_\_\_\_\_ (1969), *The Science of Artificial*, Cambridge, MIT Press.
- \_\_\_\_\_ (1973), "The Structure of Ill Structured Problems", en *Artificial Intelligence*, Vol. 4 Nos. 3-4, pp. 181-201, [https://doi.org/10.1016/0004-3702\(73\)90011-8](https://doi.org/10.1016/0004-3702(73)90011-8).
- \_\_\_\_\_ (1979), *Models of Thought*, New Haven, Yale University Press, v. 1.
- \_\_\_\_\_ (1983), "Why Should Machines Learn?", en *Machine Learning*, R. S. Michalski; J. G. Carbonell y T. M. Mitchell (eds.), Berlin, Springer, [https://doi.org/10.1007/978-3-662-12405-5\\_2](https://doi.org/10.1007/978-3-662-12405-5_2).
- Smith, G. (2018), *The AI Delusion*, Oxford, Oxford University Press.
- United Nations (2019), *The Age of Digital Interdependency*, New York, United Nations. Report of the Secretary-General's High-Level Panel on Digital Cooperation, <https://www.un.org/en/pdfs/DigitalCooperation-report-for%20web.pdf>.

- Van der Voot, H. G.; Klievink, A. J.; Arnaboldi, M.; y Meijer, A. J. (2019), "Rationality and Politics of Algorithms. Will the Promise of Big Data Survive the Dynamics of Public Decision Making?", en *Government Information Quarterly*, Vol. 36 N° 1, pp. 27-38, <https://doi.org/10.1016/j.giq.2018.10.011>.
- Veale, M. y Brass, I. (2019), "Administration by Algorithm? Public Management Meets Public Sector Machine Learning", en *Algorithm Regulation*, K. Yeung y M. Lodge (eds.), Oxford, Oxford University Press.
- Wachter, S.; Mittelstadt, B.; y Russell, C. (2020), Why Fairness Cannot Be Automated: Bridging the Gap between EU Non-Discrimination Law and AI, <https://dx.doi.org/10.2139/ssrn.3547922>.
- Wang, P. (2019), "On Defining Artificial Intelligence", en *Journal of Artificial General Intelligence*, Vol. 10 N° 2, pp. 1-37, <https://doi.org/10.2478/jagi-2019-0002>.
- Westerman, G. (2016), "Why Digital Transformation Needs a Heart", en *MIT Sloan Management Review*, Vol. 58 N° 1, pp. 10-13.
- Williamson, B. (2014), "Knowing Public Services: Cross-Sector Intermediaries and Algorithmic Governance in Public Sector Reform", en *Public Policy and Administration*, Vol. 29 N° 4, pp. 292-312, <https://doi.org/10.1177/0952076714529139>.
- Wirtz, B. W.; Weyerer, J. C.; y Geyer, C. (2019), "Artificial Intelligence and Public Sector: Applications and Challenges", en *International Journal of Public Administration*, Vol. 42 N° 7, pp. 596-615, <https://doi.org/10.1080/01900692.2018.1498103>.
- Xu, Z.; Ramanathan, J.; y Ramnath, R. (2014), "Identifying Knowledge Brokers and Their Role in Enterprise Research through Social Media", en *Computer*, Vol. 3 N° 1, pp. 26-31, <https://doi.org/10.1109/MC.2014.61>.
- Zahariadis, N. (2016), "Delphic Oracles: Ambiguity, Institutions, and Multiple Streams", en *Policy Sciences*, Vol. 49 N° 1, pp. 3-12, <https://doi.org/10.1007/s11077-016-9243-3>.
- Zheng, Y.; Han, Y.; Cui, L.; Miao, C.; Leung, C.; y Yang, Q. (2018), "SmartHS: an AI Platform for Improving Government Service Provision", paper presented at the 30th AAAI Conference on Innovative Applications of Artificial Intelligence (IAAI18), Association for the Advancement of Artificial Intelligence, New Orleans, Louisiana, United States, February 2-7, pp. 7704-7711.