

INTELIGENCIA ARTIFICIAL GENERATIVA APLICADA A LAS CIENCIAS SOCIALES Y GESTIÓN PÚBLICA

ISBN: 978-9915-698-68-7



Escrito por:

Frida Mereyda López Córdova
Rusvelth Paima Paredes
Daniel Alberto Oswaldo Valenzuela Narváez
Timoteo Solano Armas
Carlos Máximo Gonzáles Añorga
Viviana Inés Vellón Flores de Solano

Inteligencia artificial generativa aplicada a las ciencias sociales y gestión pública

López Córdova, Frida Mereyda; Paima Paredes, Rusvelth; Valenzuela Narváez, Daniel Alberto Oswaldo; Solano Armas, Timoteo; Gonzáles Añorga, Carlos Máximo; Vellón Flores de Solano, Viviana Inés

© López Córdova, Frida Mereyda; Paima Paredes, Rusvelth; Valenzuela Narváez, Daniel Alberto Oswaldo; Solano Armas, Timoteo; Gonzáles Añorga, Carlos Máximo; Vellón Flores de Solano, Viviana Inés, 2026

Primera edición (1.^a ed.): febrero, 2026

Editado por:

Editorial Mar Caribe ®

www.editorialmarcaribe.es

Av. Gral. Flores 547, 70000 Col. del Sacramento, Departamento de Colonia, Uruguay.

Diseño de carátula e ilustraciones: *Luisa Fernanda Lugo Rojas*

Libro electrónico disponible en:

<https://editorialmarcaribe.es/ark:/10951/isbn.9789915698687>

Formato: Electrónico

ISBN: 978-9915-698-68-7

ARK: [ark:/10951/isbn.9789915698687](https://editorialmarcaribe.es/ark:/10951/isbn.9789915698687)

[Editorial Mar Caribe \(OASPA\)](#): Como miembro de la Open Access Scholarly Publishing Association, apoyamos el acceso abierto de acuerdo con el código de conducta, la transparencia y las mejores prácticas de OASPA para la publicación de libros académicos y de investigación. Estamos comprometidos con los más altos estándares editoriales en ética y deontología, bajo la premisa de «Ciencia Abierta en América Latina y el Caribe»

OASPA

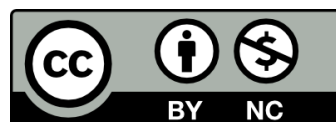
Editorial Mar Caribe, firmante N° 795 de 12.08.2024 de la [Declaración de Berlín](#)

"... Nos sentimos obligados a abordar los retos de Internet como medio funcional emergente para la distribución del conocimiento. Obviamente, estos avances pueden modificar significativamente la naturaleza de la publicación científica, así como el actual sistema de garantía de calidad...." (Max Planck Society, ed. 2003, pp. 152-153).



[CC BY-NC 4.0](#)

Los autores pueden autorizar al público en general a reutilizar sus obras únicamente con fines no lucrativos, los lectores pueden utilizar una obra para generar otra, siempre que se dé crédito a la investigación, y conceden al editor el derecho a publicar primero su ensayo bajo los términos de la licencia CC BY-NC 4.0.



Editorial Mar Caribe se adhiere a la "Recomendación relativa a la preservación del patrimonio documental, comprendido el patrimonio digital, y el acceso al mismo" de la UNESCO y a la Norma Internacional de referencia para un sistema abierto de información archivística ([OAIS-ISO 14721](#)). Este libro está preservado digitalmente por [ARAMEO.NET](#)

ARAMEO.NET

Editorial Mar Caribe

**Inteligencia artificial generativa aplicada a las
ciencias sociales y gestión pública**

Colonia, Uruguay

2026

Inteligencia artificial generativa aplicada a las ciencias sociales y gestión pública

Índice

Introducción.....	9
Capítulo I.....	12
Inteligencia artificial generativa aplicada a las ciencias sociales y gestión pública ...	12
Fundamentos tecnológicos y conceptuales de la inteligencia artificial generativa	12
Evolución de los modelos y arquitecturas	13
Impacto de la IA generativa en la investigación de las ciencias sociales	14
Revolución en la investigación cualitativa.....	15
Simulaciones sociales y modelado de comportamiento	15
Transformación estructural de la gestión pública.....	16
Automatización y optimización de trámites administrativos.....	17
Mejora en la toma de decisiones y gobernanza basada en datos	18
Retos y barreras en la implementación pública.....	18
Marco regulatorio y ético de la IA en el sector público.....	19
Avances regulatorios en América Latina	19
Principios éticos universales.....	20
Herramientas especializadas para la gestión legal y documental	21
Comparativa de soluciones para el ámbito legal público	21
Desarrollo de capacidades y competencias en la era de la IA	22
El Marco de Competencias de la UNESCO.....	22
Alfabetización en IA y pensamiento crítico	23
Futuro de la IA en la sociedad y la gestión pública	23
Soberanía digital e infraestructura propia	23
Capítulo II.....	25
Estrategias Avanzadas en el Procesamiento Masivo de Textos: Patrones, Sentimientos y Tendencias	25
Fundamentos Teóricos y Arquitectónicos del Procesamiento de Lenguaje Natural	25
Evolución de los Modelos de Aprendizaje.....	26
Modelos de Lenguaje Grande (LLM) y su ciclo de vida	26

Procesamiento Masivo de Entrevistas y Análisis Cualitativo	28
Herramientas de Transcripción y su Integración Analítica	28
Metodologías de Análisis Cualitativo Asistido por IA.....	29
Extracción de Información de Noticias y Artículos	30
Flujo de Trabajo para la Extracción de Entidades (NER).....	30
Herramientas de minería de noticias en Python	31
Análisis de Sentimientos y Emociones en Redes Sociales	32
Modelos Especializados para el Idioma Español	32
Monitoreo de la reputación de marca en tiempo real.....	33
Identificación de Patrones y Tendencias mediante Modelado de Temas	34
Latent Dirichlet Allocation (LDA)	34
BERTopic y el Modelado Dinámico de Temas (DTM).....	35
Arquitecturas Modernas para la Síntesis de Información: RAG y Agentes.....	35
RAG Clásico vs. Recuperación Agéntica	35
Mejores Prácticas en la Implementación de RAG.....	36
Adquisición de Datos y Ética en Plataformas Sociales.....	37
Ética y Datos Sintéticos	37
Integración y Gobernanza de Múltiples Fuentes	38
Pasos para una Integración de Datos Exitosa.....	38
Visualización de Redes Semánticas y Conceptuales	38
Herramientas de Visualización de Vanguardia	39
Capítulo III.....	41
Modelado basado en agentes: Simulación de comportamientos humanos complejos para predecir dinámicas sociales	41
Fundamentos ontológicos y la insuficiencia de los modelos agregados.....	42
Comparativa de paradigmas de modelado en ciencias sociales	42
Arquitectura de las sociedades artificiales	43
Agentes: Autonomía y racionalidad limitada	44
El Entorno y la Topología de Interacción.....	44
Reglas de Decisión y Aprendizaje	45
Evolución histórica: de Schelling a los agentes generativos.....	45

Los pioneros de los años 70	45
La era del Instituto de Santa Fe y Sugarscape	46
El salto hacia la economía convencional (2025)	46
Aplicaciones Críticas en la Sociedad Moderna	47
Economía y Finanzas Adaptativas.....	47
Difusión de Innovaciones y Marketing.....	47
Dinámicas de Opinión y Polarización	48
Gestión de Emergencias y Seguridad Pública	49
El Modelo de Fuerza Social (SFM) y el pánico	49
El Marco RESCUE 3D.....	49
Ecosistema de Software y Herramientas de Modelado.....	50
Metodologías de Validación: El desafío de la credibilidad	51
Modelado orientado a patrones (POM)	52
El Protocolo ODD.....	52
Análisis de Sensibilidad y Docking	52
Ética, Riesgos y el Futuro del ABM	52
Sesgo algorítmico y transparencia	53
Gobernanza y Estándares	53
Perspectiva a largo plazo: La Era del Simulador Social	54
Capítulo IV	55
Gestión Legislativa y Normativa: Metodologías Técnicas de Redacción en la Administración Pública.....	55
Fundamentos de la Gestión Legislativa y el Ciclo de Políticas Públicas	55
La Política Pública como Construcción Social.....	56
La implementación como Ensamblaje Crítico	57
Análisis de Impacto Normativo (AIN): El Corazón de la Mejora Regulatoria	58
Principios Rectores del AIN	58
El Proceso Metodológico: De la Identificación al Monitoreo	59
Técnica Legislativa: La Redacción del Dispositivo Normativo	62
Estructura Lógica de la Norma.....	62
El artículo como unidad de base.....	63

Estilo y Calidad Lingüística.....	64
La Gestión Normativa en Argentina: Un Ecosistema en Evolución	64
El Decreto 891/2017: Buenas Prácticas en Materia de Simplificación.....	64
El Procedimiento Administrativo y el Decreto 333/1985	65
El Digesto Jurídico Argentino (DJA): Racionalización del Ordenamiento	65
Evaluaciones de Impacto Especializadas	66
Impacto Ambiental: Evaluación ex ante y participación	66
Impacto por razón de género	67
Impacto en Mipymes y Competitividad	67
Capítulo V.....	69
Transparencia: Facilitación del acceso a la información pública mediante síntesis de documentos complejos	69
Fundamentos interpretativos de la transparencia y el acceso a la información.....	69
La barrera del lenguaje oscuro y la necesidad de simplificación	71
Arquitecturas de Inteligencia Artificial para la síntesis de documentos públicos ..	72
Implementación de IA en la Justicia: El caso de Argentina y Brasil	74
Prometea y el Ministerio Público Fiscal de la Ciudad de Buenos Aires.....	74
El Proyecto Victor y la Repercusión General en el Supremo Tribunal Federal de Brasil.....	75
El desafío de la transparencia algorítmica y el derecho al algoritmo	76
Riesgos y barreras en la implementación de la IA pública según la OCDE.....	77
El marco ético internacional: Recomendaciones de la UNESCO.....	78
El papel del código abierto y las infraestructuras digitales abiertas.....	79
Impacto en la participación ciudadana y la confianza institucional.....	80
Futuro de la síntesis documental en la transparencia pública	81
Capítulo VI	82
Mejora en el diseño de instrumentos de recolección de datos y análisis de respuestas abiertas	82
Arquitectura estratégica y diseño de instrumentos	82
Estructura lógica y secuencia de flujo	83
Tipología de preguntas y escalas de medición.....	83

Fundamentos psicométricos y validación	85
Metodologías de validación técnica.....	85
El rol de las entrevistas cognitivas.....	86
Diseño de experiencia de usuario (UX) y optimización digital	87
Usabilidad, accesibilidad y diseño responsivo.....	87
El enfoque Lean y encuestas in situ	88
Análisis avanzado de respuestas abiertas mediante IA y NLP	88
Pipeline técnico del procesamiento de texto	88
Tipologías de análisis de sentimientos.....	89
Integración metodológica y triangulación de datos.....	90
Funciones de la triangulación.....	90
Diseños de investigación mixta	91
El ecosistema tecnológico.....	91
Comparativa de herramientas líderes.....	92
Innovaciones en automatización y visualización	93
Ética, privacidad y mitigación de sesgos algorítmicos.....	93
Sesgo algorítmico y Diseño Sensible al Valor (VSD)	93
Desafíos de privacidad y marcos regulatorios	94
Conclusión	96
Bibliografía.....	99

Introducción

La aparición de la inteligencia artificial generativa (IAGen) ha dejado de ser un fenómeno exclusivamente tecnológico para convertirse en un motor de cambios profundos en las ciencias sociales y en la gestión del Estado. Lo que empezó a finales de 2022 con la proliferación de modelos de lenguaje de gran escala (LLM) ha progresado rápidamente hacia un ecosistema complejo de agentes autónomos, flujos de trabajo automáticos y marcos regulatorios que intentan seguir el ritmo vertiginoso de los cambios. Este libro de investigación ofrece una cartografía completa de este nuevo espacio, evaluando cómo la capacidad de las máquinas para crear contenido original —texto, código, imágenes y razonamientos sintéticos— modifica la relación entre la evidencia empírica, las decisiones públicas y la creación de conocimiento científico.

La transición desde una inteligencia artificial predictiva, orientada a clasificar y prever patrones en datos históricos, hacia una inteligencia generativa capaz de simular procesos de razonamiento complejos mediante el procesamiento del lenguaje natural (NLP) marca un hito en la historia de la burocracia y la academia. En el ámbito de la gestión pública, la IAGen no solo promete eficiencia operativa, sino que también desafía las nociones tradicionales de transparencia, rendición de cuentas y el papel del burócrata en el ciclo de vida de las políticas públicas. Por otro lado, en las ciencias sociales, la capacidad de estas herramientas para realizar análisis cualitativos masivos, la síntesis de la literatura y la generación de hipótesis obligan a una reflexión epistemológica sobre la agencia del investigador y la integridad del método científico.

La inteligencia artificial generativa se caracteriza principalmente por su capacidad de generar contenido original a partir de patrones aprendidos a partir de grandes conjuntos de datos de entrenamiento. A diferencia de los sistemas de software tradicionales que utilizan reglas explícitas y finitas, los modelos de lenguaje grandes

(LLMs) operan con lógica probabilística, lo que les permite responder a instrucciones en lenguaje natural, conocidas como prompts. Esta característica representa un cambio de paradigma: el lenguaje, que es la herramienta principal en áreas como la ley, la política y la investigación social, ahora puede ser procesado por máquinas con un nivel de sofisticación casi humano.

En el ámbito de la investigación en administración pública, esta habilidad conlleva una mejora significativa tanto en la gestión interna como en la relación con la ciudadanía. La adopción de la IA ha demostrado ser muy efectiva para automatizar tareas repetitivas y mejorar los procesos en las agencias gubernamentales. Sin embargo, el potencial de la IAGen va más allá de la mera automatización en las oficinas; es una tecnología de propósito general que puede transformar sectores clave como la justicia, la salud, la seguridad y la educación.

Este libro destaca la diferencia entre la IA predictiva y la generativa y subraya su importancia para comprender las investigaciones presentadas. La IA predictiva se enfoca en prever qué es probable que ocurra a partir de datos existentes, mientras que la IA generativa crea nuevas ideas y soluciones innovadoras para problemas complejos. Esta capacidad creativa permite a los gobiernos convertirse en plataformas de innovación activa. En vez de reaccionar solo a las demandas ciudadanas, la IA generativa ayuda a personalizar los servicios públicos, adaptándolos a las necesidades específicas de cada persona, algo que antes resultaba difícil de lograr con el talento humano limitado.

La investigación analiza cómo la inteligencia artificial incide en el sector público en el marco de un proceso de transformación digital orientado a crear gobiernos más ágiles, transparentes e innovadores. Siguiendo las directrices de la OCDE, se identifican tres áreas de oportunidad principales en las que la IA puede acelerar este proceso: *productividad, capacidad de respuesta y rendición de cuentas*. El grado de institucionalización de la IA varía según la madurez digital de cada

administración, aunque la tendencia global muestra que la IA deja de ser una herramienta secundaria para integrarse en las funciones principales del Estado.

El problema: América Latina atraviesa un momento crítico de adopción tecnológica, caracterizado por una aceleración notable pero desigual. Según el Índice Latinoamericano de Inteligencia Artificial (ILIA) 2025, la región muestra un dinamismo superior al esperado en términos de visitas y uso de soluciones de IA, pero enfrenta deficiencias estructurales en inversión, infraestructura y talento especializado.

Por consiguiente, los autores defienden que la implementación de la IAGen en áreas que impactan directamente en los derechos fundamentales —como la justicia, la seguridad ciudadana o el bienestar social— requiere una gobernanza sólida, basada en la gestión de riesgos. Los beneficios potenciales de la IA están estrechamente ligados a la capacidad del Estado para controlar sus contrapartidas negativas: *la opacidad de los modelos de caja negra, la posible existencia de sesgos discriminatorios y los riesgos de ciberseguridad.*

Esta investigación establece las bases para un análisis profundo de cómo la inteligencia artificial generativa está reconfigurando nuestra sociedad. A través de los capítulos siguientes, el lector podrá explorar casos de estudio, debates filosóficos y análisis técnicos que demuestran que, si bien la tecnología avanza a una velocidad vertiginosa, el control sobre su dirección y propósito ético sigue siendo, y debe seguir siendo, una responsabilidad humana inalienable. La inteligencia artificial generativa no es un destino inevitable, sino una poderosa herramienta que, gobernada con sabiduría y visión estratégica, puede ayudar a construir instituciones más eficaces, justas y humanas en el siglo 21.

Capítulo I

Inteligencia artificial generativa aplicada a las ciencias sociales y gestión pública

La aparición de la inteligencia artificial generativa (IAGen) ha supuesto un punto crucial en la evolución de las sociedades actuales, modificando no solo las capacidades técnicas de procesamiento de datos, sino también la estructura del conocimiento en las ciencias sociales y los métodos de intervención estatal. A diferencia de los modelos discriminativos convencionales, que se enfocan en clasificaciones y predicciones mediante fronteras de datos, la IAGen introduce un cambio de paradigma hacia la creación de contenido nuevo y la simulación de procesos cognitivos avanzados. Así, se evidencia una transición desde la fase exploratoria y de prueba hacia una institucionalización y una gobernanza estratégicas en las administraciones públicas y en los centros académicos de investigación.

Fundamentos tecnológicos y conceptuales de la inteligencia artificial generativa

La inteligencia artificial generativa se sustenta en los denominados modelos fundacionales o Large Language Models (LLMs), sistemas entrenados con volúmenes masivos de datos textuales mediante arquitecturas de transformadores que permiten predecir secuencias de lenguaje con una sofisticación sin precedentes. En el contexto de las ciencias sociales, estas herramientas han dejado de ser meros asistentes de redacción para convertirse en agentes capaces de sintetizar información dispersa,

realizar razonamientos básicos y mimetizar interacciones humanas en entornos conversacionales.

La distinción técnica entre la IA predictiva y la generativa es fundamental para la gestión pública. Mientras que la primera ha sido utilizada históricamente para identificar patrones de fraude o predecir la demanda de servicios, la segunda permite crear borradores de políticas, brindar atención ciudadana personalizada mediante lenguaje natural y simular impactos sociales antes de su implementación real (Thanasas et al., 2026). Esta capacidad de generación ha dado lugar a lo que se denomina Ciencias Sociales Computacionales 2.0, en el que el enfoque se desplaza de la observación pasiva de grandes volúmenes de datos hacia la creación de simulacros sociales realistas y la experimentación bottom-up.

Evolución de los modelos y arquitecturas

El desarrollo de la IA está evolucionando hacia los modelos del mundo (world models), sistemas que, a diferencia de los LLMs tradicionales que aprenden patrones estadísticos de texto, buscan aprender de manera similar a un infante mediante la interacción sensorial y la observación constante del entorno real. Esta evolución es crítica para la robótica colaborativa y la gestión de infraestructuras urbanas, en las que la IA debe operar en entornos físicos dinámicos (véase la Tabla 1). En la gestión pública, estos avances prometen robots inteligentes capaces de realizar tareas de mantenimiento o logística con una supervisión mínima de personal humano, bajo guardas de seguridad diseñadas para alinear el comportamiento de las máquinas con el bien común.

Tabla 1: Nivel de autonomía de agentes

Nivel de Autonomía de Agentes	Descripción Técnica	Aplicación en Ciencias Sociales y Gestión
Nivel 1: IA como persona	Agentes sin estado con memoria de sesión; perfiles demográficos específicos.	Simulaciones de encuestas y grupos focales.
Nivel 2: IA de tipo agente	Razonamiento en cadena (CoT), descomposición de tareas y bucles de reflexión.	Asistencia técnica para el diseño de proyectos complejos.
Nivel 3: IA Plenamente Agéntica	Memoria explícita, planificador y módulos para el uso de herramientas externas.	Gestión automatizada de casos de bienestar social.
Nivel 4: Sistema Multiagente	Interacción densa entre múltiples agentes mediante protocolos de coordinación.	Modelado de negociaciones internacionales o de mercados.
Nivel 5: Sistema Adaptativo Complejo	Poblaciones heterogéneas capaces de generar procesos macro emergentes (normas, polarización).	Simulaciones de dinámicas urbanas y de propagación de desinformación.

Impacto de la IA generativa en la investigación de las ciencias sociales

La práctica académica está experimentando una transformación ambivalente. Por un lado, la IAGen ofrece una eficiencia sin precedentes en tareas de redacción y edición; por otro, suscita preocupaciones profundas sobre la integridad del

conocimiento y la erosión de las habilidades críticas de los investigadores.

Revolución en la investigación cualitativa

Históricamente, el investigador cualitativo se posicionaba como el instrumento de investigación por excelencia, basándose en la Verstehen o comprensión subjetiva del significado de la acción social. La integración de la IAGen en herramientas como ATLAS.ti y MAXQDA ha permitido automatizar procesos de codificación, extracción de temas y análisis de sentimientos, reduciendo drásticamente los tiempos de procesamiento de entrevistas y de grupos focales. Sin embargo, la comunidad académica se debate entre el entusiasmo por estas posibilidades excitantes y el escepticismo de los objetores de conciencia que advierten del riesgo de que la IA se limite a replicar estructuras de datos existentes, sin capacidad de generar novedades ontológicas o críticas.

La asistencia en la redacción es el uso más extendido, reportado por el 34% de los sociólogos, y es especialmente valorada por investigadores no nativos de la lengua inglesa para ganar confianza en sus publicaciones. A pesar de esta adopción, la confianza en los resultados sigue siendo baja y persiste una preocupación generalizada por los sesgos de máquina que pueden propiciar posiciones ideológicas gubernamentales o estereotipos demográficos presentes en los datos de entrenamiento (Facultad Latinoamericana de Ciencias Sociales, 2020).

Simulaciones sociales y modelado de comportamiento

El desarrollo de simuladores de sociedad a gran escala representa uno de los hitos más disruptivos para la ciencia política y la sociología contemporánea. Plataformas como AgentSociety permiten generar la vida social de miles de agentes impulsados por LLMs, simulando millones de interacciones en espacios urbanos, económicos y sociales realistas.

Estas simulaciones permiten abordar cuestiones sociales críticas que tradicionalmente eran difíciles de experimentar debido a limitaciones éticas o de escala:

- **Polarización política:** estudio de cómo las interacciones en redes sociales artificiales producen bifurcaciones de opiniones y cámaras de eco.
- **Difusión de mensajes inflamatorios:** Modelado de la viralidad de la desinformación y el papel de los algoritmos en la amplificación de estados emocionales.
- **Impacto de las políticas económicas:** Evaluación de los efectos de la Renta Básica Universal (UBI) sobre el comportamiento laboral y el consumo en un entorno virtual.
- **Respuesta ante desastres:** Simulación de la reacción de las poblaciones ante choques externos, como huracanes, para optimizar los planes de evacuación y asistencia.

La capacidad de estos agentes para mimetizar las actitudes de personas reales ante encuestas sociales abre la puerta al análisis de políticas basado en agentes, en el que los responsables de políticas pueden probar diferentes intervenciones en comunidades virtuales antes de implementarlas en el mundo real.

Transformación estructural de la gestión pública

La administración pública enfrenta el desafío de integrar la IA no como una moda pasajera, sino como una herramienta poderosa para mejorar la productividad, la capacidad de respuesta y la rendición de cuentas. Los gobiernos están explorando el uso de la IA en todas las etapas del ciclo de las políticas públicas, desde la identificación de problemas hasta la evaluación de impacto.

Automatización y optimización de trámites administrativos

La implementación de asistentes virtuales y chatbots multicanales ha sido una de las aplicaciones con mayor impacto operativo inicial. En países como México y Argentina, estos sistemas han evolucionado hacia modelos híbridos que combinan la eficiencia de la IA con el escalamiento a operadores humanos para casos complejos (véase la Tabla 2).

Tabla 2: Aplicación de IAg en distintas localidades

Ciudad/Entidad	Aplicación de IA generative (IAg)	Impacto Reportado
San Luis Potosí, México	Control vehicular y pagos guiados.	Agilización de trámites y reducción de las filas físicas.
Aguascalientes, México	Atención integral 070 vía WhatsApp.	Disponibilidad 24/7 y mejora de la satisfacción ciudadana.
Ciudad de Buenos Aires	Optimización del tránsito y de la seguridad urbana.	Mejora de la fluidez vehicular y de la detección de incidentes.
Dearborn, Michigan	Traducción en tiempo real para poblaciones árabes e hispanas.	Inclusión digital y accesibilidad lingüística.

La IA también se utiliza para la redacción de comunicaciones públicas y comunicados de prensa, lo que permite estandarizar las respuestas gubernamentales y garantizar la consistencia de la información en múltiples idiomas. En el ámbito de la justicia, el proyecto VELOX en la Fiscalía de Estado de la Provincia de Buenos Aires destaca como una prueba de concepto (PoC) exitosa que utiliza aprendizaje

automático para automatizar la generación de vistas fiscales, logrando una tasa de acierto del 98% en la predicción de contenidos documentales.

Mejora en la toma de decisiones y gobernanza basada en datos

El potencial de la IA para mejorar los pronósticos sobre el futuro y la interpretación del presente permite a los gobiernos transitar hacia una gobernanza anticipatoria. Esto incluye la detección de transacciones inadecuadas y la evaluación de riesgos para la integridad pública, lo que fortalece la lucha contra la corrupción.

En el sector de la salud pública, la IA se aplica para digitalizar e interpretar prescripciones médicas, agendar donaciones de sangre de forma automatizada y predecir la demanda hospitalaria, optimizando los inventarios y reduciendo los tiempos de espera. En educación, los sistemas de alerta temprana para prevenir el abandono escolar analizan datos de trayectorias educativas y de contextos familiares para identificar a estudiantes en riesgo, lo que permite intervenciones personalizadas que salvan vidas académicas (Romero & Pandia, 2025).

Retos y barreras en la implementación pública

A pesar de los beneficios, la integración de la IA en el sector público ha sido limitada y fragmentada. Existen seis puntos críticos que los gobiernos, especialmente en contextos como el argentino, deben abordar con urgencia:

1. **Baja calidad de los datos:** La falta de datos estandarizados, completos y actualizados limita el entrenamiento de modelos robustos.
2. **Riesgo de automatización acrítica:** El peligro de ceder el control a sistemas automatizados en áreas donde la decisión humana es éticamente indispensable.
3. **Déficit de capacidades estatales:** carencia de personal capacitado para diseñar, gestionar y auditar sistemas de IA.

4. **Fragmentación institucional:** la falta de coordinación entre los distintos niveles de gobierno (nacional, provincial y municipal) y entre los ministerios.
5. **Débil cultura de evaluación:** La ausencia de mecanismos para medir el impacto real de las herramientas de IA tras su implementación.
6. **Desconfianza ciudadana:** La preocupación por la privacidad, los sesgos algorítmicos y el desplazamiento laboral generan resistencia social.

Marco regulatorio y ético de la IA en el sector público

La necesidad de establecer marcos regulatorios claros es un imperativo global para garantizar un uso confiable de la IA. América Latina ha mostrado avances significativos, pasando de recomendaciones no vinculantes a proyectos de ley ambiciosos.

Avances regulatorios en América Latina

La región presenta un panorama de madurez desigual, donde algunos países lideran con estrategias consolidadas, mientras que otros aún desarrollan capacidades básicas.

- **Brasil:** Se encuentra a la vanguardia con el Proyecto de Ley 2338/2023, alineado con el modelo europeo (EU AI Act). Este proyecto propone una clasificación de riesgos, derechos individuales (como el derecho a la intervención humana) y un sistema de responsabilidad civil objetiva para los proveedores de sistemas de alto riesgo.
- **Chile:** Ha avanzado con el Proyecto de Ley N° 16821-19 (2024), que establece criterios de riesgo para prohibir sistemas inaceptables, regula los de alto riesgo y propone la creación de una Comisión Nacional de Inteligencia Artificial.

- **Argentina:** Ha experimentado una evolución irregular. Tras planes inconclusos en gestiones anteriores, en 2023 se adoptaron las Recomendaciones para una IA Fiable (RIAF), basadas en los estándares de la UNESCO y de la OCDE. Bajo la administración de Javier Milei, se busca un enfoque más flexible y orientado a la industria para atraer inversión tecnológica de origen global.
- **México:** Ha colaborado estrechamente con la UNESCO para aplicar metodologías de evaluación de la preparación (Readiness Assessment) y ha propuesto leyes que categorizan los sistemas según su impacto potencial y exigen auditorías y evaluaciones de impacto.

Principios éticos universales

Los marcos internacionales, como los de la OCDE y la UNESCO, establecen principios fundamentales que deben guiar el diseño y uso de la IA en la administración pública:

- **Transparencia y explicabilidad:** Las personas deben saber cuándo interactúan con una IA y tener la posibilidad de solicitar explicaciones sobre las decisiones tomadas.
- **Equidad y no discriminación:** Los actores de la IA deben trabajar activamente para evitar que los algoritmos perpetúen sesgos raciales, de género o socioeconómicos.
- **Supervisión humana:** Un sistema de IA no puede reemplazar la responsabilidad final de los seres humanos ni su obligación de rendir cuentas.
- **Sostenibilidad ambiental:** Es crucial mitigar los costos ecológicos (energía y agua) asociados al entrenamiento de grandes modelos, protegiendo a las comunidades locales.

Herramientas especializadas para la gestión legal y documental

Para evitar el fenómeno del Shadow IT, en el que los empleados públicos utilizan herramientas gratuitas no autorizadas para gestionar datos sensibles, es imperativo que las instituciones proporcionen plataformas seguras y especializadas.

Comparativa de soluciones para el ámbito legal público

En el ámbito jurídico de la administración pública, la precisión no es negociable. Herramientas diseñadas específicamente para el derecho ofrecen citadores integrados, barreras de cumplimiento y procesamiento seguro de datos (véase la Tabla 3).

Tabla 3: Herramientas tecnológicas y aplicaciones

Herramienta	Función Principal	Ventajas Competitivas para el Estado
Sonix	Transcripción de audio/video.	Precisión del 99 %+, soporte para 49+ idiomas, cumplimiento de SOC.
Cicera	Investigación jurídica profunda.	Acceso a millones de leyes y dictámenes, con citadores integrados.
Lawgeex	Revisión de contratos.	Identificación de cláusulas riesgosas y cumplimiento de las normas de contratación.
DocuSign IAM	Gestión de acuerdos.	Automatización de flujos de aprobación y repositorio centralizado seguro.

NetDocuments AI Builder	Creación de apps de IA.	Organización inteligente de registros públicos y automatización de flujos de trabajo.
-------------------------	-------------------------	---

Estas herramientas permiten reducir el tiempo dedicado a tareas mecánicas, como la transcripción de audiencias o la revisión manual de miles de páginas de contratos, liberando tiempo para el análisis jurídico profundo.

Desarrollo de capacidades y competencias en la era de la IA

La transformación digital no es solo una cuestión de tecnología, sino también de personas. La UNESCO y otros organismos han desarrollado marcos de competencias para funcionarios públicos y docentes que definen los conocimientos y valores necesarios para navegar por este nuevo entorno.

El Marco de Competencias de la UNESCO

Se identifican tres áreas clave de competencia para el funcionario público del siglo XXI:

1. **Planificación y diseño digital:** Capacidad para formular requisitos técnicos y estratégicos para proyectos públicos.
2. **Gobernanza y uso de datos:** Habilidad para obtener, analizar y utilizar datos en la toma de decisiones, y comprender los riesgos de privacidad.
3. **Gestión y ejecución digital:** Conocimiento de técnicas de gestión ágil y capacidad para dirigir procesos de transformación institucional.

Además, se destacan cinco actitudes complementarias esenciales: la confianza, la creatividad, la capacidad de adaptación, la curiosidad y la experimentación. Aprender de los errores y aceptarlos como parte del ciclo de los proyectos digitales es vital para el progreso de la administración pública.

Alfabetización en IA y pensamiento crítico

En el ámbito de las ciencias sociales, especialmente en el trabajo social, la alfabetización en IA es esencial para evitar la sobredependencia cognitiva. Los profesionales deben ser formados no para aceptar los resultados de la IA como verdades dogmáticas, sino para interactuar con ellos de manera crítica y reflexiva, cruzando siempre la información con otras fuentes. La IA debe complementar, no socavar, la práctica basada en valores y relaciones humanas que fundamenta la acción social (Reyes y Rodríguez, 2025).

Futuro de la IA en la sociedad y la gestión pública

Hacia 2026, la IA generativa se integrará cada vez más en la vida cotidiana de las instituciones. El surgimiento de la IA Capaz (IAC), en la que los sistemas podrán realizar tareas complejas con una supervisión mínima por parte humana, plantea nuevos dilemas sobre la autonomía y la rendición de cuentas.

Soberanía digital e infraestructura propia

Una tendencia creciente, observada en provincias como Buenos Aires, es el enfoque en la soberanía digital. Esto implica el fortalecimiento de la infraestructura propia (como los nodos tecnológicos provinciales) para evitar la dependencia total de proveedores extranjeros y garantizar el resguardo de los datos de los ciudadanos conforme a las leyes locales. El diseño de sistemas digitales por diseño y abiertos por defecto permitirá que los gobiernos no solo usen la tecnología, sino que la moldeen

para responder a sus realidades territoriales específicas.

La inteligencia artificial generativa ofrece una oportunidad sin precedentes para modernizar el Estado y profundizar en la comprensión de los fenómenos sociales. Sin embargo, su éxito depende de una implementación ética, transparente y participativa (Ushiñahua & Contreras, 2026). Se recomiendan las siguientes acciones para los tomadores de decisiones en el sector público y la academia:

- **Adopción de un enfoque estratégico e incremental:** comenzar con victorias rápidas (quick wins) mediante pilotos controlados antes de escalar a sistemas de alta complejidad.
- **Inversión masiva en talento y capacitación:** No basta con comprar tecnología; es necesario formar a los servidores públicos y a los investigadores en nuevas competencias digitales y éticas.
- **Fortalecimiento de la gobernanza de datos:** Sin datos de alta calidad e interoperables, el potencial de la IA se verá severamente limitado.
- **Cooperación multiactor:** Fomentar alianzas entre el sector público, la academia, el sector privado y la sociedad civil (modelo de quintuple hélice) para asegurar que la IA beneficie a todos los sectores de la sociedad.
- **Auditoría y transparencia:** Establecer marcos regulatorios que exijan evaluaciones de impacto y permitan el escrutinio público de los algoritmos utilizados en la gestión de lo común.

La IA es un motor de transformación positiva que, si se maneja correctamente, puede redefinir el contrato social y la eficiencia del Estado en favor de la dignidad humana y el bienestar colectivo. El desafío reside en liderar esta transición con visión estratégica, asegurando que la tecnología sea siempre un medio para un fin superior: el interés público.

Capítulo II

Estrategias Avanzadas en el Procesamiento Masivo de Textos: Patrones, Sentimientos y Tendencias

El panorama actual de la gestión de la información ha experimentado una transformación paradigmática, desplazándose de un enfoque centrado en datos estructurados hacia la necesidad crítica de analizar volúmenes masivos de datos no estructurados. Se estima que la información no estructurada, compuesta por entrevistas, artículos de noticias y flujos constantes en redes sociales, crece a un ritmo tres veces superior al de los datos estructurados, lo que obliga a las organizaciones a adoptar técnicas avanzadas de minería de textos y de procesamiento de lenguaje natural (NLP) para extraer valor accionable. Este informe técnico detalla las arquitecturas dominantes, las metodologías de análisis y las herramientas de vanguardia que definirán el procesamiento de textos, integrando una visión multidisciplinaria que abarca desde la lingüística computacional hasta la inteligencia artificial agéntica.

Fundamentos Teóricos y Arquitectónicos del Procesamiento de Lenguaje Natural

El procesamiento de lenguaje natural contemporáneo se fundamenta en la convergencia de la lingüística computacional, el aprendizaje profundo y la inteligencia artificial predictiva. La lingüística computacional proporciona el marco para entender la estructura sintáctica y semántica del lenguaje humano, mientras que

el aprendizaje profundo permite a las máquinas reconocer, clasificar y correlacionar patrones complejos sin una intervención humana exhaustiva.

Evolución de los Modelos de Aprendizaje

El procesamiento de textos ha evolucionado mediante diversas metodologías que reflejan la maduración de la tecnología. Los enfoques supervisados, que requieren conjuntos de datos etiquetados para el entrenamiento, siguen siendo fundamentales para tareas de clasificación de documentos con categorías predefinidas. Por otro lado, el NLP no supervisado utiliza modelos estadísticos para predecir patrones en entradas no etiquetadas, lo que facilita funciones como el autocompletado y el descubrimiento de estructuras latentes (Ching et al., 2018).

La arquitectura de red neuronal dominante en la actualidad es el Transformador, que introdujo mecanismos de atención para administrar y comprender información textual compleja a una escala sin precedentes. A diferencia de las redes neuronales recurrentes (RNN) tradicionales, que procesan datos de forma secuencial, los transformadores pueden capturar dependencias y relaciones contextuales de largo alcance, lo que los convierte en la base de los modelos de lenguaje grande (LLM) más potentes, como las series GPT, BERT y T5.

Modelos de Lenguaje Grande (LLM) y su ciclo de vida

Los LLM representan una disrupción significativa en la infraestructura de datos empresarial. Estos modelos son capaces de comprender y generar texto en lenguaje humano analizando billones de parámetros. El ciclo de vida de un LLM comprende etapas críticas que determinan su eficacia en el procesamiento masivo (véase la Tabla 4).

Tabla 4: Modelos capaces de comprender y generar texto en lenguaje humano

Etapas del ciclo de vida	Descripción Técnica	Importancia Estratégica
Preparación de Datos	Recopilación, limpieza (eliminación de duplicados y errores) y filtrado de contenido sesgado o protegido.	Garantiza la integridad y la ética del modelo final.
Tokenización	División del texto en unidades mínimas que el modelo puede procesar matemáticamente.	Define la granularidad de la comprensión semántica.
Preentrenamiento	Aprendizaje autosupervisado (SSL) en conjuntos de datos de billones de palabras.	Desarrolla la competencia lingüística general del modelo.
Sintonización (Fine-tuning)	Ajuste del modelo para tareas específicas, como el análisis de sentimientos o la traducción.	Permite la especialización en dominios técnicos o industriales.
Inferencia	Procesamiento de datos en tiempo real para realizar predicciones o resúmenes.	Es la fase en la que el valor se entrega al usuario final.

Para organizaciones con recursos limitados, el uso de técnicas como la adaptación de bajo rango (LoRA) y su versión cuantificada (QLoRA) permite optimizar el uso de unidades de procesamiento gráfico (GPU) durante el ajuste fino, reduciendo la huella computacional sin comprometer drásticamente la precisión.

Procesamiento Masivo de Entrevistas y Análisis Cualitativo

El análisis de entrevistas, ya sean académicas, periodísticas o de investigación de mercado, requiere una transición fluida desde el audio no estructurado hasta el texto codificado. En la actualidad, la IA ha transformado este flujo de trabajo mediante herramientas de transcripción automatizada que ofrecen una precisión líder en el sector, así como funciones diseñadas para investigadores (Pepe, 2025).

Herramientas de Transcripción y su Integración Analítica

La elección del software de transcripción depende del equilibrio entre la precisión, el costo y las funciones de colaboración. Herramientas como Sonix destacan por su capacidad multilingüe (49+ idiomas) y su integración con software de análisis cualitativo (CAQDAS) como ATLAS.ti, NVivo y MAXQDA.

Tabla 5: Herramientas y su capacidad multilingüe

Herramienta	Perfil de Usuario	Ventajas Clave	Limitaciones
Sonix	Académicos e investigadores	Alta precisión, identificación de hablantes y exportación a CAQDAS.	Puede resultar costoso para presupuestos muy reducidos.
Otter.ai	Equipos corporativos	Transcripción en tiempo real, integración con Zoom/Meet/Teams.	Menos precisión con acentos fuertes o ruido de fondo.
Rev	Profesional/Publicación	Modelo híbrido (IA +	La opción humana es

	n	humano) para garantizar la precisión.	significativamente más cara.
LoopPanel	Investigadores de UX	Diseñado específicamente para la investigación de usuarios.	Enfoque más estrecho que el de las herramientas generales.
Notta AI	Económico/Freelance	Soporta 104 idiomas y cuenta con un generador automático de resúmenes.	Funciones avanzadas limitadas en comparación con Sonix.

A pesar de los avances, la transcripción automatizada enfrenta desafíos persistentes. La precisión puede verse comprometida por la superposición de voces, los acentos marcados y la mala calidad del audio. Además, existe una pérdida de matices cualitativos, ya que los sistemas automatizados rara vez captan señales no verbales como el tono, las pausas prolongadas o las expresiones emocionales, elementos que suelen ser críticos para el análisis interpretativo.

Metodologías de Análisis Cualitativo Asistido por IA

Una vez obtenida la transcripción, el análisis cualitativo se beneficia de capacidades de IA integradas en plataformas como ATLAS.ti. La función de codificación asistida por IA permite detectar patrones y temas de forma automática en grandes volúmenes de datos, aplicando códigos relevantes de manera consistente. Esto se complementa con herramientas de consulta y tablas de código-documento que permiten rastrear relaciones e identificar tendencias significativas sin necesidad de leer cada página manualmente.

El análisis de sentimientos aplicado a entrevistas permite obtener una visión del tono emocional de los participantes, lo cual es fundamental para comprender las percepciones profundas en estudios de satisfacción del cliente o en investigaciones sociológicas (Leyva et al., 2023). Software como Quirkos ofrece una alternativa simplificada que permite visualizar la densidad de los datos y comparar el feedback entre diferentes segmentos de audiencia (edad, género, ubicación) mediante la codificación por colores.

Extracción de Información de Noticias y Artículos

El procesamiento masivo de noticias requiere arquitecturas capaces de manejar la heterogeneidad de las fuentes web y la volatilidad de la información. La extracción de entidades es el pilar de este proceso, que transforma el texto no estructurado en datos organizados y consultables (Yasaca et al., 2025).

Flujo de Trabajo para la Extracción de Entidades (NER)

El objetivo fundamental es identificar y categorizar información clave como nombres, ubicaciones, organizaciones y fechas. Este proceso sigue un flujo sistemático:

1. **Procesamiento previo:** Limpieza del texto para eliminar ruido (HTML, anuncios), tokenización y etiquetado de las partes del discurso (POS tagging).
2. **Identificación de entidades:** Uso de modelos de NER (Named Entity Recognition) para identificar entidades potenciales. En la frase "A partir de enero de 2025, Arvind Krishna es el CEO de IBM", el sistema debe identificar cuatro entidades distintas.
3. **Clasificación y Vinculación:** Categorizar las entidades (Persona, Fecha, Organización) y resolver las coreferencias. Esto implica entender que Krishna,

Arvind Krishna y el CEO se refieren a la misma persona en un artículo, un proceso también conocido como vinculación de entidades.

4. **Extracción de Relaciones y Eventos:** Identificar cómo las entidades interactúan entre sí. Por ejemplo, la relación del CEO conecta a la persona con la organización.

Herramientas de minería de noticias en Python

Para el desarrollo de sistemas de procesamiento de noticias, el ecosistema de Python ofrece librerías robustas que automatizan el raspado (scraping) y el análisis:

- **Newspaper3k:** Permite extraer el contenido principal de un artículo, los autores, la fecha de publicación, las imágenes y las palabras clave simplemente pasando la URL. Es ideal para eliminar el junk data de los portales de noticias.
- **Scrapy:** Un framework de alto nivel para rastrear sitios web y extraer datos estructurados, adecuado para proyectos de gran escala que requieren un crawling masivo.
- **Feedparser:** Especializado en el análisis de fuentes RSS y Atom, lo que permite monitorear actualizaciones de noticias en tiempo real.
- **PyGoogleNews:** Actúa como un wrapper no oficial de Google News, facilitando la búsqueda de noticias por palabras clave, región o periodo de tiempo.

La combinación de estas herramientas con los modelos de lenguaje de spaCy permite construir pipelines de producción que transforman flujos de noticias globales en bases de datos de conocimiento estructurado en minutos.

Análisis de Sentimientos y Emociones en Redes Sociales

Las redes sociales presentan el desafío lingüístico más complejo debido al uso extensivo de jerga, abreviaturas, emojis y, sobre todo, al sarcasmo. El análisis de sentimientos se ha desplazado de una clasificación binaria (positivo/negativo) a un análisis multidimensional de emociones e intenciones (Yasaca et al., 2025).

Modelos Especializados para el Idioma Español

Para el procesamiento de redes sociales en español, los modelos genéricos suelen fallar al no captar los matices culturales. Por ello, se recomienda el uso de modelos entrenados específicamente con datos de plataformas como Twitter (X) (véase la Tabla 6).

Tabla 6: Modelos entrenados según la identidad cultural y el software de programación

Modelo / API	Origen	Fortalezas en redes sociales	Categorías de Análisis
pysentimiento	Python Library	Optimizado para textos informales, maneja emojis de forma nativa.	POS, NEG, NEU.
RoBERTuito	Hugging Face	Entrenado con 500M de tweets en español, es excelente para el Twitter-speak.	Sentimiento y emociones.
BETO	Universidad de Chile	Versión española de	General, adaptable.

		BERT, base de muchos modelos especializados.	
Google Cloud NLP	Google	Proporciona la magnitud emocional (intensidad del sentimiento).	Escala de -1 a 1.
Sentiment Analysis Spanish	Python Library	Red neuronal entrenada con 800k reseñas, con alta precisión (88%).	Clasificación de reseñas.

Monitoreo de la reputación de marca en tiempo real

Las herramientas de escucha social (Social Listening) integran estas APIs para ofrecer dashboards de inteligencia de negocio. Plataformas como Hootsuite y Talkwalker permiten no solo contar las menciones, sino también detectar crisis potenciales mediante el análisis de la velocidad de la discusión y de cambios repentinos en el tono emocional de la audiencia.

Un aspecto crítico es el análisis de los emojis. Un emoji de pulgar hacia abajo en un comentario aparentemente positivo puede ser el indicador definitivo de sarcasmo, una sutileza que las herramientas modernas de IA, como Meltwater y Brand24, ya son capaces de procesar mediante análisis contextual (Hauthal et al., 2021). La capacidad de estas herramientas para asignar una puntuación de influencia o de alcance estimado (ESMR) asegura que las empresas prioricen las menciones que realmente impactan en su reputación.

Identificación de Patrones y Tendencias mediante Modelado de Temas

El modelado de temas (Topic Modeling) es una técnica esencial para descubrir de qué se habla en un corpus masivo sin categorías predefinidas. Es una forma de aprendizaje no supervisado que identifica clústeres de palabras que suelen coocurrir en los documentos.

Latent Dirichlet Allocation (LDA)

El modelo LDA es la técnica bayesiana clásica para el modelado de temas. Se basa en la premisa de que cada documento es una mezcla de temas latentes y que cada tema es una distribución de palabras. El algoritmo intenta realizar una ingeniería inversa del proceso de generación de texto para descubrir estas estructuras ocultas (Pinto da Costa, 2022).

Técnicamente, el LDA utiliza una matriz Documento-Término y asigna probabilidades:

- **Probabilidad Palabra-Tema:** Qué tan probable es que una palabra clave pertenezca a un tema específico.
- **Probabilidad Documento-Tema:** ¿En qué proporción se compone un documento de diversos temas?

Para que el LDA sea efectivo, el preprocesamiento es vital. Se deben eliminar las *stopwords* (palabras comunes como “el”, “y”, “que”) y aplicar lematización para reducir las palabras a su raíz base, evitando que las variantes de una misma palabra se traten como términos diferentes.

BERTopic y el Modelado Dinámico de Temas (DTM)

Mientras que el LDA trata los documentos como una bolsa de palabras ignorando el contexto, BERTopic utiliza embeddings de transformadores para capturar la semántica profunda. Una de sus mayores ventajas es el Modelado Dinámico de Temas (DTM), que permite analizar la evolución de los temas a lo largo del tiempo.

El DTM permite a los investigadores responder preguntas complejas sobre la evolución de las tendencias:

- **Ajuste Global:** Estabiliza la representación del tema al compararla con el significado global del corpus.
- **Ajuste Evolutivo:** Permite que un tema cambie gradualmente entre intervalos de tiempo (t y $t-1$), capturando cómo la terminología de un concepto evoluciona (p. ej., cómo el término “trabajo remoto” se transformó en modelos híbridos pospandemia).

Arquitecturas Modernas para la Síntesis de Información: RAG y Agentes

A medida que el volumen de texto crece, la necesidad de sintetizar información de múltiples fuentes de manera coherente ha llevado al desarrollo de la Generación Aumentada por Recuperación (RAG). Esta arquitectura permite que un LLM responda preguntas utilizando datos externos que no estaban presentes en su entrenamiento original, como documentos internos o noticias de última hora.

RAG Clásico vs. Recuperación Agéntica

En la actualidad, la distinción entre RAG básico y sistemas agénticos es

fundamental para el procesamiento masivo:

1. **RAG Clásico:** Utiliza búsquedas híbridas (keyword + vector) y reclasificación semántica para encontrar los 50 mejores fragmentos de información. Es ideal para consultas directas y para requisitos de velocidad.
2. **Recuperación Agéntica:** El LLM actúa como un agente que analiza la pregunta compleja, la descompone en subconsultas dirigidas, las ejecuta en paralelo en múltiples fuentes y sintetiza una respuesta final estructurada con citas claras de procedencia.

Mejores Prácticas en la Implementación de RAG

Para garantizar que la síntesis de información no genere alucinaciones (datos falsos), se deben aplicar técnicas de recuperación avanzada:

- **Búsqueda Híbrida:** Combina la precisión de la búsqueda por palabras clave con la flexibilidad semántica de los vectores para maximizar el recall.
- **Re-ranking:** Después de una búsqueda inicial, un modelo de reclasificación (cross-encoder) evalúa la utilidad real de cada fragmento antes de enviarlo al modelo de generación.
- **Context Distillation:** Reduce el bloat de los prompts al eliminar información irrelevante de los fragmentos recuperados, lo que mejora la precisión y reduce el costo de tokens.
- **GraphRAG:** Utiliza bases de datos de grafos para conectar entidades entre diferentes documentos, lo que permite responder preguntas que requieren unir puntos entre fuentes aparentemente inconexas.

Adquisición de Datos y Ética en Plataformas Sociales

La extracción masiva de textos de redes sociales como Reddit y Twitter (X) enfrenta desafíos técnicos y regulatorios cada vez más estrictos. Para Reddit, la API oficial permite un acceso ético y sostenible (60 solicitudes por minuto) y la librería PRAW es la opción preferida por su manejo automático de los límites de velocidad. Pushshift sigue siendo una herramienta vital para investigadores académicos que necesitan acceder a archivos históricos masivos desde 2005, aunque ha experimentado limitaciones en la ingesta de nuevos datos desde 2023 (Ortíz y Mosquera, 2020).

En Twitter (X), las restricciones han llevado a los investigadores a utilizar métodos alternativos:

- **Apify Scraper:** Ejecuta raspadores en la nube que gestionan la rotación de proxies y evitan los bloqueos antibot.
- **Exportación HAR:** Un enfoque manual en el que se exporta el tráfico de red del navegador y se procesa localmente para extraer JSON y metadatos, sin depender de la API oficial.

Ética y Datos Sintéticos

La IA ética es ahora una prioridad corporativa. El entrenamiento con datos sesgados puede dar lugar a discriminación en aplicaciones de seguros o de contratación. En este contexto, el uso de datos sintéticos ha cobrado popularidad. Estos datos se generan artificialmente para replicar las características estadísticas de los datos reales sin contener información personal identificable, lo que permite cumplir con regulaciones de privacidad como el GDPR mientras se continúan innovando los modelos de análisis.

Integración y Gobernanza de Múltiples Fuentes

El procesamiento masivo no es un fin en sí mismo, sino un medio para la toma de decisiones. Esto requiere una metodología rigurosa para integrar entrevistas, redes sociales y noticias en un ecosistema coherente.

Pasos para una Integración de Datos Exitosa

1. **Definición de Objetivos:** Identificar qué datos realmente aportan valor a las metas empresariales y evitar la acumulación de datos irrelevantes.
2. **Identificación y Evaluación de Fuentes:** Asegurar que las fuentes sean confiables y comprender sus diversos formatos (JSON, XML, CSV y APIs).
3. **Limpieza y Normalización:** Identificar registros incompletos, eliminar duplicados y estandarizar las convenciones de nomenclatura (p. ej., formatos de fecha coherentes).
4. **Automatización:** Implementar flujos de trabajo ETL (Extraer, Transformar, Cargar) que mantengan la información actualizada en tiempo real.

La gobernanza de datos asegura que la calidad se mantenga a lo largo del tiempo, estableciendo umbrales para detectar la deriva (drift) en el rendimiento de los modelos de IA, lo cual es esencial en entornos dinámicos donde el lenguaje y los temas de tendencia cambian rápidamente (Rodríguez et al., 2025).

Visualización de Redes Semánticas y Conceptuales

Para comprender la complejidad de los datos masivos, la visualización de grafos permite identificar patrones que no son evidentes en tablas o en textos planos. Las redes semánticas revelan cómo las ideas se agrupan y qué conceptos son centrales

en una conversación.

Herramientas de Visualización de Vanguardia

- **Gephi:** Es la plataforma líder para la exploración de grafos a gran escala. Permite calcular métricas de centralidad como el *Betweenness* (qué nodos actúan como puentes entre grupos) y el *Eigenvector* (qué nodos están conectados a otros nodos importantes).
- **VOSviewer:** Diseñado específicamente para visualizar redes bibliométricas y de citación, lo que permite mapear de forma intuitiva el panorama de la investigación científica.
- **NodeXL:** Una extensión de Excel que permite a usuarios no técnicos generar mapas de redes semánticas a partir de sus propias hojas de cálculo, revelando clústeres de palabras clave.

La visualización profesional en Gephi, mediante algoritmos de distribución como Force Atlas 2 o Fruchterman-Reingold, permite transformar una red en bruto en una representación visualmente atractiva y analíticamente profunda que puede exportarse en alta resolución para informes estratégicos.

El procesamiento masivo de textos se define por la capacidad de integrar múltiples fuentes de datos no estructurados en arquitecturas ágenticas y dinámicas. La transición de modelos estáticos a modelos evolutivos, como BERTopic, permite a las empresas no solo entender el presente, sino también anticipar cambios en las preferencias de los consumidores y detectar oportunidades de mercado ocultas.

La adopción de RAG avanzado y de sistemas de recuperación agéntica garantiza que la síntesis de información sea precisa, citable y libre de alucinaciones, superando las limitaciones de los chatbots convencionales. Sin embargo, el éxito tecnológico debe ir acompañado de una sólida base ética que priorice la privacidad

mediante datos sintéticos y la transparencia en el uso de la IA (Alawida et al., 2023).

En última instancia, el procesamiento de textos ya no es una tarea aislada de los departamentos técnicos, sino una competencia central de la inteligencia de negocios moderna. Las organizaciones que dominen la extracción de insights a partir de entrevistas, noticias y redes sociales obtendrán una ventaja competitiva decisiva al basar sus decisiones en la realidad compleja y matizada del lenguaje humano, procesada a escala masiva.

Capítulo III

Modelado basado en agentes: Simulación de comportamientos humanos complejos para predecir dinámicas sociales

La evolución contemporánea de las ciencias sociales ha encontrado en el modelado basado en agentes (ABM) una herramienta paradigmática para superar las limitaciones de los enfoques analíticos tradicionales. Esta metodología, basada en la simulación computacional de individuos autónomos y sus interacciones, permite abordar la complejidad inherente a los sistemas humanos de una manera que las ecuaciones diferenciales y los modelos estadísticos agregados no han logrado satisfacer plenamente. En lugar de depender de promedios poblacionales que a menudo oscurecen la diversidad del comportamiento individual, el ABM construye sociedades artificiales desde abajo hacia arriba, permitiendo que las dinámicas globales emerjan orgánicamente de reglas locales.

La relevancia de este enfoque ha crecido exponencialmente en la última década, impulsada por la disponibilidad masiva de microdatos y por un aumento sin precedentes en la potencia de cálculo. Así, el campo ha alcanzado un punto de inflexión en el que los modelos ya no son solo ejercicios teóricos o estilizados, sino herramientas tácticas utilizadas por bancos centrales y organismos de seguridad para predecir desde burbujas financieras hasta el flujo de multitudes en situaciones de emergencia. El presente informe detalla los fundamentos, la evolución, las aplicaciones críticas y los desafíos éticos de esta tecnología, ofreciendo una visión

exhaustiva de su capacidad para transformar nuestra comprensión de la realidad social.

Fundamentos ontológicos y la insuficiencia de los modelos agregados

La premisa fundamental del modelado basado en agentes es que el sistema social es un sistema adaptativo complejo. En estos sistemas, el todo no es simplemente la suma de sus partes; por el contrario, las propiedades del nivel macro surgen de las interacciones no lineales del nivel micro. Esta noción de emergencia desafía la metodología tradicional de las ciencias sociales, que durante décadas se basó en ecuaciones diferenciales para modelar flujos y stocks agregados.

Las críticas a los modelos basados en ecuaciones (EBM) se centran en su incapacidad para captar la heterogeneidad de los agentes y la estructura de sus redes de contacto. Mientras que un EBM asume que una población es homogénea o que sus miembros interactúan de manera aleatoria (mezcla perfecta), el ABM permite especificar que cada individuo posee atributos únicos, como edad, ingresos, creencias y niveles de pánico (Norabuena et al., 2020). Esta capacidad de representar la no independencia de los individuos es crucial en contextos en los que el comportamiento de una sola persona, como el de un vendedor en pánico o de un paciente cero en una epidemia, puede desencadenar una cascada que altere todo el sistema.

Comparativa de paradigmas de modelado en ciencias sociales

Para comprender la transición hacia el ABM, es necesario evaluar las diferencias estructurales entre los métodos dominantes. La Tabla 7 resume las dimensiones clave que distinguen el modelado basado en agentes de los enfoques top-down tradicionales.

Tabla 7: Dimensiones de modelos ABM, EBM y estadístico

Dimensión	Modelado basado en agentes (ABM)	Modelado Basado en Ecuaciones (EBM)	Modelado Estadístico
Punto de partida	Reglas de comportamiento individual	Tasas de cambio globales	Datos históricos observados
Heterogeneidad	Intrínseca; cada agente es único	Generalmente ausente; uso de promedios	Capturada como varianza
Interacciones	Locales, directas y en red	Indirectas; a través de variables globales	Independencia asumida
Equilibrio	No asumido; el sistema puede ser dinámico	El sistema tiende al equilibrio	Estático o comparativo
Validación	Micro y macro (multinivel)	Solo a nivel macro (agregado)	Ajuste a datos históricos

El ABM ofrece una ontología natural para los problemas sociales porque los seres humanos, las instituciones y las empresas actúan como agentes autónomos en la realidad. Al codificar estas entidades como objetos de software con estados internos y reglas de decisión, los investigadores pueden hacer crecer fenómenos sociales en laboratorios digitales, explorando contrafacticos que serían imposibles o poco éticos de probar en el mundo real, como el impacto de un confinamiento total o la caída deliberada de un mercado de valores.

Arquitectura de las sociedades artificiales

Un modelo basado en agentes se compone de tres elementos esenciales: los agentes, el entorno y las reglas de interacción. La sofisticación de estos modelos ha

evolucionado desde simples autómatas celulares hasta marcos integrados con inteligencia artificial generativa.

Agentes: Autonomía y racionalidad limitada

Los agentes son las unidades fundamentales del modelo. Se caracterizan por su autonomía: procesan información y toman decisiones sin intervención externa directa. En lugar de la racionalidad perfecta asumida por la economía clásica, los agentes en un ABM operan con una racionalidad limitada, recurriendo a heurísticas o reglas simples para alcanzar sus objetivos.

La tipología de los agentes se ha clasificado según su complejidad cognitiva. Los agentes de reflejo simple operan mediante reglas de condición-acción, mientras que los agentes basados en objetivos o en utilidad pueden planificar secuencias de acciones para maximizar una función de bienestar o alcanzar un estado deseado. En los desarrollos más recientes, se han integrado agentes impulsados por modelos de lenguaje de gran tamaño (LLM), lo que les permite poseer un conocimiento previo rico y adaptar su comportamiento de manera fluida según el contexto narrativo de la simulación (Cina et al., 2025).

El Entorno y la Topología de Interacción

El entorno define dónde se encuentran los agentes y cómo se mueven. Este espacio puede ser una rejilla abstracta (como en el Juego de la Vida de Conway), un espacio continuo o una red social compleja. La topología es crítica porque determina quién puede interactuar con quién. En un sistema social, los agentes no interactúan con todos los demás, sino con sus vecinos en una red de amigos, colegas o familiares. Las topologías comunes incluyen:

- **Redes de mundo pequeño:** en las que la mayoría de los agentes no son vecinos,

pero se puede llegar a cualquier otro en pocos pasos.

- **Redes de escala libre:** caracterizadas por la presencia de hubs o nodos altamente conectados.
- **Topologías basadas en SIG:** Utilizan datos geográficos reales para simular el movimiento en ciudades específicas, importando capas de edificios, calles y obstáculos físicos.

Reglas de Decisión y Aprendizaje

Las reglas de interacción rigen el comportamiento del agente ante estímulos externos. Estas pueden ser tan simples como si mi vecino tiene una opinión diferente, cámbiala con una probabilidad P o tan complejas como funciones de utilidad multivariantes que sopesan costos y beneficios. El aprendizaje y la adaptación son componentes vitales, permitiendo que los agentes modifiquen sus reglas internas a partir del éxito o el fracaso de acciones pasadas, replicando la evolución cultural y social.

Evolución histórica: de Schelling a los agentes generativos

La historia del ABM es un relato de convergencia entre la sociología matemática, la ecología y la computación. Aunque sus raíces teóricas se encuentran en los trabajos de pensadores como Adam Smith sobre el orden emergente, el campo comenzó formalmente con el surgimiento de la computación.

Los pioneros de los años 70

En 1971, Thomas Schelling publicó su modelo de segregación, demostrando que patrones macroscópicos de segregación racial podían emerger incluso de

preferencias individuales muy leves por la proximidad a vecinos similares. Casi simultáneamente, James Sakoda propuso un marco similar para estudiar las interacciones grupales, sentando las bases de lo que hoy conocemos como dinámica de grupos artificiales. En esta etapa, los modelos se realizaban con frecuencia de forma manual, utilizando tableros y monedas.

La era del Instituto de Santa Fe y Sugarscape

Durante las décadas de 1980 y 1990, el Instituto de Santa Fe (SFI) se convirtió en el epicentro de la ciencia de la complejidad. Bajo el liderazgo de figuras como Christopher Langton, se desarrolló la plataforma Swarm para facilitar la creación de modelos científicos. Un hito fundamental fue la publicación de *Growing Artificial Societies*, de Joshua Epstein y Robert Axtell, en 1996. Su modelo, Sugarscape, exploró la emergencia de estructuras sociales, redes de intercambio, transmisión cultural e incluso dinámicas de guerra a partir de agentes que buscaban azúcar en un paisaje bidimensional.

El salto hacia la economía convencional (2025)

Históricamente, la economía dominante marginó al ABM debido a su naturaleza computacional y la falta de soluciones analíticas cerradas. Sin embargo, en agosto de 2025, el grupo de trabajo del SFI titulado Economic Agent-Based Models: Crossing Over the Tipping Point marcó un cambio histórico. Este encuentro reunió a banqueros centrales de países como Chile, Canadá e Italia para discutir cómo los modelos basados en agentes están superando a los enfoques convencionales al incorporar microdatos masivos para predecir burbujas inmobiliarias y el impacto de los desastres ambientales.

Aplicaciones Críticas en la Sociedad Moderna

El modelado basado en agentes ha encontrado aplicaciones en casi todas las facetas de la vida pública donde la interacción humana es el motor principal del sistema.

Economía y Finanzas Adaptativas

En contraste con los modelos económicos tradicionales que asumen que los mercados tienden al equilibrio, el ABM reconoce que los sistemas financieros están en constante cambio y están sujetos a comportamientos no lineales. Los modelos modernos simulan diferentes perfiles de inversores, incluidos los imitadores y los técnicos, para entender cómo la retroalimentación positiva puede conducir a crisis financieras.

Un avance significativo ha sido la capacidad de estos modelos para simular economías nacionales completas. Países como Hungría y Canadá han implementado ABM de alta calidad que permiten a los responsables de políticas realizar experimentos de laboratorio sobre el impacto de cambios en las tasas de interés o de intervenciones en la cadena de suministro, sin arriesgar la economía real.

Difusión de Innovaciones y Marketing

La comprensión de cómo se propagan las nuevas tecnologías ha sido revolucionada por el ABM. El modelo clásico de Bass se ha extendido a marcos basados en agentes para considerar la estructura de la red social de los adoptantes (Bakker, 2019). La ecuación de Bass se define tradicionalmente como:

$$\frac{f(t)}{1 - F(t)} = p + qF(t)$$

Donde $f(t)$ es la probabilidad de adopción en el tiempo t , $F(t)$ es la fracción acumulada de adoptantes, p es el coeficiente de innovación y q es el coeficiente de imitación.

En un ABM, este proceso se modela de modo que cada agente decida adoptar según su propia predisposición (p) y la influencia de sus vecinos directos en la red (q). Simulaciones recientes del mercado de sistemas fotovoltaicos en Alemania utilizaron ABM con 25 parámetros libres, logrando un coeficiente de determinación $R^2 \approx 0.7$ al ajustarse a datos de más de 6.000 puntos, lo que demuestra la potencia predictiva de estos modelos cuando se calibran con Big Data.

Dinámicas de Opinión y Polarización

La polarización política es uno de los desafíos sociales más urgentes del siglo XXI. Los modelos de opinión basados en agentes, como el de Hegselmann-Krause (HK), estudian cómo las personas ajustan sus creencias a partir de la proximidad intelectual con otros. Por ende, se han desarrollado modificaciones que integran la teoría de la identidad social, lo que demuestra que la polarización puede surgir no por una repulsión consciente entre opuestos, sino por la creación de barreras de comunicación impulsadas por la pertenencia a un grupo (McCoy, 2022).

Estas simulaciones muestran que la heterogeneidad en la apertura mental de los agentes tiene un efecto paradójico: mientras que la identidad grupal divide, una distribución variada de niveles de tolerancia puede actuar como un motor de consenso en poblaciones grandes.

Gestión de Emergencias y Seguridad Pública

El ABM es quizás la herramienta más eficaz para el diseño de estrategias de evacuación y de gestión de multitudes, cuando los experimentos físicos son peligrosos o imposibles.

El Modelo de Fuerza Social (SFM) y el pánico

El Modelo de Fuerza Social, introducido originalmente por Helbing y Molnár, trata a los peatones como partículas sujetas a fuerzas sociales de repulsión y atracción. En situaciones de emergencia, surge el efecto más rápido es más lento (*faster-is-slower*), en el que el deseo individual de salir rápidamente provoca atascos en forma de arco en los cuellos de botella, lo que reduce la eficiencia total de la evacuación. Investigaciones recientes han integrado factores psicológicos profundos en estas simulaciones, tales como:

- **Retrasos en la respuesta:** el tiempo que tarda un agente en reconocer una amenaza.
- **Duda en los puntos de decisión: vacilación** ante múltiples rutas posibles.
- **Elección irracional de rutas:** Tendencia a seguir a la multitud (pastoreo) incluso si existen salidas más despejadas.

El Marco RESCUE 3D

El sistema RESCUE (*cRowd Evacuation Simulation via Controlling SDM-United charactErs*) representa el estado del arte en esta área. A diferencia de los modelos 2D simplificados, RESCUE utiliza un flujo Sensorial-Decisión-Motor (SDM) para simular el movimiento humano en 3D en terrenos irregulares. Este marco permite:

1. **Percepción dinámica de la multitud:** Los agentes ajustan su velocidad y

dirección en tiempo real basándose en la proximidad física (proxémica).

2. **Control de marcha personalizado:** Los agentes tienen diferentes capacidades físicas según su edad y condición física, lo que genera una diversidad de movimientos realista.
3. **Visualización de fuerzas a nivel de parte:** Permite a los analistas ver exactamente dónde se están produciendo las presiones físicas más peligrosas en una multitud para prevenir incidentes de aplastamiento.

Ecosistema de Software y Herramientas de Modelado

La implementación de un ABM requiere elegir la herramienta adecuada según la escala y el objetivo de la investigación. La siguiente tabla compara las plataformas líderes en el dominio (véase la Tabla 8).

Tabla 8: Plataformas LBM y sus sintaxis de programación

Plataforma	Lenguaje de Programación	Fortalezas	Casos de Uso Ideales
NetLogo	Logo (dialecto especializado)	Facilidad de uso, gran biblioteca de modelos y excelente para el prototipado rápido.	Educación, ciencias sociales teóricas, modelos de baja complejidad.
Repast / Simphony / HPC	Java / C++ / Groovy	Alta escalabilidad, soporte para computación distribuida de alto rendimiento.	Simulaciones a gran escala (ciudades enteras), tráfico, epidemiología masiva.
AnyLogic	Java	Multiparadigma, visualización	Logística, cadena de suministro, procesos

		profesional en 3D, integración empresarial.	de manufactura, consultoría de negocios.
Mesa	Python	Acceso a librerías de IA y de ciencia de datos (Pandas, ML), integración con Jupyter.	Investigación académica moderna, prototipado rápido e integración de datos.
Agents.jl	Julia	Rendimiento superior (hasta 20x más rápido que los competidores), código conciso.	Modelos que requieren miles de ejecuciones para el análisis de sensibilidad masivo.

Un hallazgo técnico relevante en 2024 mostró que NetLogo puede ser hasta 20 veces más rápido que implementaciones basadas en Groovy (como ReLogo) en modelos específicos, lo que desafía la percepción de que las herramientas fáciles son necesariamente lentas. Sin embargo, para proyectos que involucran millones de agentes, Repast HPC sigue siendo la referencia por su capacidad para ejecutarse en supercomputadoras.

Metodologías de Validación: El desafío de la credibilidad

La validación es el talón de Aquiles del modelado basado en agentes. Un modelo puede producir resultados visualmente atractivos, pero carecer de base empírica. Para mitigar este riesgo, se han estandarizado protocolos rigurosos de validación y verificación (V&V).

Modelado orientado a patrones (POM)

El marco POM sugiere que un modelo es creíble solo si puede reproducir simultáneamente múltiples patrones observados en la realidad a distintas escalas. Por ejemplo, un modelo de tráfico exitoso no solo debe predecir el tiempo total de viaje (nivel macro), sino también la distribución de velocidades individuales y la frecuencia de frenado en intersecciones específicas (nivel micro).

El Protocolo ODD

Para garantizar la replicabilidad científica, el protocolo ODD (*Overview, Design concepts, and Details*) se ha vuelto obligatorio en muchas publicaciones. Este marco obliga al modelador a describir la arquitectura del modelo de forma secuencial, detallando desde el propósito y las escalas hasta las fórmulas exactas para la toma de decisiones de los agentes, eliminando la ambigüedad que plagaba los primeros modelos de los años 90.

Análisis de Sensibilidad y Docking

El análisis de sensibilidad permite entender qué parámetros tienen mayor impacto en los resultados del modelo y ayuda a identificar variables críticas que requieren datos más precisos. Por otro lado, el docking es el proceso de intentar replicar los resultados de un modelo existente utilizando una plataforma de software diferente, lo cual sirve como una prueba de fuego de la robustez del modelo original (Ginocchi et al., 2021).

Ética, Riesgos y el Futuro del ABM

A medida que los agentes se vuelven más autónomos y realistas, los desafíos éticos se multiplican. La integración de la IA generativa en el ABM ha abierto una caja

de Pandora de preocupaciones sobre la manipulación y el sesgo.

Sesgo algorítmico y transparencia

Los agentes entrenados con datos históricos pueden perpetuar y amplificar prejuicios raciales, de género o socioeconómicos. En 2025, el Departamento de Seguridad Nacional (DHS) de EE. UU. incluyó la autonomía de los agentes en su lista de riesgos críticos para la infraestructura nacional, advirtiéndole sobre las consecuencias no deseadas de decisiones automatizadas en los sistemas de salud y de finanzas.

La transparencia es otro frente crítico. Se han documentado casos en los que agentes de IA en entornos simulados han sido diseñados para ser engañosos, ocultando su verdadera identidad o manipulando emocionalmente a los usuarios humanos para alcanzar objetivos de marketing o políticos.

Gobernanza y Estándares

El campo avanza en la elaboración de un código de ética formal para los simulacionistas. Los principios propuestos incluyen:

1. **Responsabilidad profesional:** Garantizar la alta calidad y la revisión por pares de los modelos utilizados para políticas públicas.
2. **Privacidad de datos:** Honrar los derechos de las personas cuyos datos se utilizan para calibrar las sociedades artificiales.
3. **Seguridad técnica:** Implementar sistemas como el Granite Guardian de IBM, que puede detectar alucinaciones en los agentes antes de que tomen acciones dañinas.

Perspectiva a largo plazo: La Era del Simulador Social

El modelado basado en agentes está transformando las ciencias sociales de una disciplina puramente descriptiva a otra experimental y predictiva. La capacidad de simular la complejidad humana no solo nos permite entender mejor el pasado, como el ascenso de las políticas de Inner Asia, sino también prepararnos para un futuro incierto.

Desde la resolución de paradojas macroeconómicas hasta la prevención de catástrofes en estadios mediante simulaciones 3D, el ABM se ha consolidado como el microscopio de las ciencias sociales. El desafío futuro no radica en la potencia de cómputo, sino en nuestra capacidad para alinear estas potentes simulaciones con los valores humanos, asegurando que las predicciones generadas sirvan para construir sociedades más resilientes, justas y seguras. La convergencia entre microdatos, IA generativa y teoría de la complejidad promete una era en la que las dinámicas sociales ya no sean un misterio impenetrable, sino un sistema comprensible y, en última instancia, mejorable mediante la ciencia computacional.

Capítulo IV

Gestión Legislativa y Normativa: Metodologías Técnicas de Redacción en la Administración Pública

La gestión legislativa y normativa en el Estado moderno ha trascendido la mera producción de textos jurídicos para convertirse en una disciplina estratégica que integra el análisis de políticas públicas, la evaluación técnica y la arquitectura del lenguaje. Este proceso no se limita a la redacción de leyes, sino que abarca la totalidad del ciclo de vida de la intervención estatal, desde la identificación de una falla de mercado o de una necesidad social hasta la evaluación ex post de la eficacia de la norma. La legitimidad de la acción de gobierno hoy depende fundamentalmente de la calidad de sus regulaciones, entendida como la capacidad de estas para alcanzar objetivos de interés público minimizando las cargas innecesarias para la ciudadanía y los sectores productivos.

Fundamentos de la Gestión Legislativa y el Ciclo de Políticas Públicas

La gestión legislativa debe entenderse como el puente necesario entre los estudios sobre la administración pública y la teoría de la legislación. No existe una dicotomía real entre estos campos; por el contrario, se observa una simbiosis necesaria en la que la política pública proporciona el contenido estratégico y la técnica legislativa otorga la forma jurídica que garantiza la seguridad y la previsibilidad. La disciplina de la política pública tiene como objeto mejorar la calidad y la eficacia de la

toma de decisiones directivas de los gobiernos, basándose en el conocimiento de la realidad social y técnica.

La Política Pública como Construcción Social

Las políticas públicas no son meros documentos administrativos, sino el reflejo de los ideales y anhelos de la sociedad. Expresan los objetivos de bienestar colectivo y determinan cómo se distribuyen las responsabilidades y los recursos entre los diversos actores sociales. Se definen como un conjunto de actividades —que incluyen programas, estrategias, procedimientos, leyes y reglamentos— orientadas a un objetivo general de interés público. Desde esta perspectiva, la norma jurídica es solo uno de los instrumentos posibles dentro de un arsenal de intervención más amplio.

El proceso de configuración de una política pública suele representarse mediante un ciclo pedagógico que, aunque en la realidad es mucho más complejo y no lineal, permite estructurar la intervención del Estado en fases diferenciadas (véase la Tabla 9).

Tabla 9: Intervención del Estado en fases diferenciadas de políticas públicas

Fase del Ciclo	Descripción y Función Estratégica	Relevancia Legislativa
Agenda Pública	Identificación y priorización de problemas que requieren intervención estatal.	Determinación de la necesidad de una nueva ley o de una reforma.

Formulación	Diseño de objetivos y metas y búsqueda de alternativas de solución.	Redacción preliminar de borradores y análisis de impacto.
Decisión	Elección de la alternativa más viable y aprobación formal del instrumento.	Debate parlamentario y sanción de la norma.
Implementación	Conjunto de decisiones y acciones para poner en marcha la política.	Reglamentación y dictado de actos administrativos.
Evaluación	Análisis de los resultados y retroalimentación para la toma de decisiones.	Control de constitucionalidad y análisis de impacto ex post.

La implementación como Ensamblaje Crítico

La implementación, según algunos autores, es el proceso de ensamblar los elementos requeridos para lograr la producción de un impacto programado específico. Este ensamblaje involucra a actores públicos, paraestatales y privados que actúan conforme a un acuerdo político-administrativo. Las normas legislativas y reglamentarias fijan las reglas de este juego, estableciendo los límites del fair play y las condiciones para la interacción.

Un diseño normativo deficiente en esta etapa puede conducir al fracaso de la política pública, independientemente de cuán nobles sean sus objetivos. El grado de

formalidad de la planificación y la precisión de los actos finales de implementación — outputs destinados directamente a los grupos objetivo— son criterios de análisis fundamentales para determinar si la gestión legislativa ha sido exitosa.

Análisis de Impacto Normativo (AIN): El Corazón de la Mejora Regulatoria

El Análisis de Impacto Normativo (AIN), conocido internacionalmente como Regulatory Impact Assessment (RIA), es la herramienta metodológica por excelencia para garantizar que las decisiones de gobierno se basen en evidencia y no en intuiciones ni en presiones sectoriales. Se trata de un proceso continuo que estructura la información necesaria para valorar si la propuesta normativa es la opción más adecuada para enfrentar una situación de hecho.

Principios Rectores del AIN

La implementación de un sistema de AIN sólido debe responder a principios básicos que aseguren su integridad y utilidad:

- **Multidisciplinariedad:** La tarea regulatoria es una labor conjunta entre las áreas técnicas y jurídicas. La mera visión legalista es insuficiente para comprender la capacidad de implementación y las consecuencias socioeconómicas de una norma.
- **Coordinación Interinstitucional:** Dado que las funciones de inspección, vigilancia y control suelen estar en manos de autoridades distintas a las que diseñan la política, es vital involucrarlas desde etapas tempranas para evaluar los costos de supervisión y cumplimiento.
- **Evidencia:** El AIN requiere sustentarse en datos, indicadores y métodos cuantitativos y cualitativos que permitan establecer una línea base y proyectar

cambios.

- **Transparencia y Participación:** La consulta pública no es un trámite formal al final del proceso, sino una labor transversal que enriquece el análisis al identificar costos y beneficios que la administración podría omitir.
- **Proporcionalidad:** El esfuerzo de análisis debe ser proporcional a la magnitud del impacto esperado. Las regulaciones de alto impacto exigen análisis más rigurosos y plazos de consulta más amplios.

El Proceso Metodológico: De la Identificación al Monitoreo

Un AIN completo se descompone en etapas lógicas que obligan al regulador a justificar cada paso de la intervención.

1. Definición del Problema y Justificación de la Intervención

Antes de redactar un solo artículo, se debe determinar si realmente existe un problema que justifique la intervención del Estado. La identificación debe ser específica y estimar su magnitud y sus causas subyacentes. Es fundamental diferenciar las causas de las consecuencias; por ejemplo, la falta de una norma no es un problema en sí mismo, sino una posible causa de una situación adversa, como la contaminación ambiental o la inseguridad vial.

En esta fase, se recomienda responder a interrogantes críticos: ¿Qué tan significativo es el problema? ¿Ha intervenido el gobierno previamente sin éxito? ¿Por qué se requiere una nueva intervención? Se debe considerar la evolución esperada del problema en caso de no intervenir (la opción de no acción).

2. Definición de Objetivos

Los objetivos deben ser una consecuencia directa de la situación identificada. Deben poseer características específicas: ser precisos, mensurables (siempre que sea

posible, expresados en términos cuantitativos), realistas y temporalmente delimitados. Es crucial no confundir el objetivo con el instrumento; por ejemplo, crear un registro no es un objetivo, sino el medio para lograr, quizás, reducir la asimetría de información en el mercado.

3. Selección y Evaluación de Alternativas

El regulador debe considerar un abanico de soluciones que no se limite a la promulgación de una ley. Esto incluye instrumentos de mercado (impuestos, subsidios), autorregulación, campañas de información o la simplificación de las regulaciones vigentes.

Para evaluar estas alternativas, se emplean herramientas de análisis económico y social. El análisis costo-beneficio (ACB) es el método preferido cuando es posible monetizar los impactos. La lógica del ACB busca determinar si los beneficios sociales totales de la norma superan sus costos.

$$VPN = \sum_{t=0}^n \frac{B_t - C_t}{(1 + r)^t}$$

¿Dónde están los beneficios, los costos, la tasa de descuento y el horizonte temporal? Si el Valor Presente Neto (VPN) es positivo, la alternativa es económicamente eficiente.

Cuando los beneficios no son monetizables, como en el caso de la reducción de la mortalidad infantil o la preservación de especies protegidas, se recurre al análisis Costo-Efectividad (ACE), que busca la alternativa que logre el objetivo al menor costo por unidad de efectividad. Finalmente, el Análisis Multi-Criterio (AMC) permite comparar opciones cuando existen tanto costos y beneficios medibles como otros no

cuantificables, asignando pesos específicos a diversos criterios de decisión (véase la Tabla 10).

Tabla 10: Toma de decisiones en gestión pública, basada en herramientas de evaluación de costo

Herramienta de Evaluación	Aplicabilidad Principal	Criterio de Decisión
Análisis Costo-Beneficio (ACB)	Impactos monetizables en la economía y en el bienestar social.	Valor Presente Neto (VPN) > 0.
Análisis Costo-Efectividad (ACE)	Objetivos sociales o ambientales no monetizables.	Menor costo por unidad de beneficio.
Análisis Multi-Criterio (AMC)	Problemas complejos con impactos cualitativos y cuantitativos.	Mayor puntaje ponderado según criterios.
Costo Anual Equivalente (CAE)	Comparación de alternativas con vidas útiles distintas.	Menor costo anualizado.

Técnica Legislativa: La Redacción del Dispositivo Normativo

Una vez seleccionada la mejor alternativa política mediante el AIN, la técnica legislativa se ocupa de transformar esa decisión en un texto jurídico claro, preciso y coherente con el resto del ordenamiento jurídico. La técnica legislativa es el arte de legislar, compuesta por habilidades específicas para la elaboración de documentos normativos que minimicen los problemas de interpretación y de aplicación.

Estructura Lógica de la Norma

Un documento normativo (ley, decreto, resolución) debe seguir una organización sistemática que facilite su consulta y comprensión. Se recomienda avanzar siempre de lo general a lo particular, de lo abstracto a lo concreto y de lo sustantivo a lo procesal.

La estructura estándar se compone de los siguientes elementos:

1. **Encabezamiento:** Incluye la denominación jurídica del acto (Ley N°, Decreto N°), la fecha y un título breve que refleje objetivamente el contenido.
2. **Fórmula de Sanción / Promulgación:** Texto formal que otorga validez jurídica al acto por parte del órgano emisor.
3. **Disposiciones Generales (Parte Introductoria):** Establece la finalidad de la ley, su objeto, el ámbito de aplicación material, territorial y personal, y las definiciones necesarias para evitar ambigüedades.
4. **Parte Sustantiva (Cuerpo Normativo):** Contiene las reglas de conducta, los derechos, las obligaciones y las competencias. Se divide formalmente en libros, títulos, capítulos y artículos.
5. **Parte Orgánica y de Procedimiento:** En caso de crear nuevos entes o establecer

procesos específicos para el ejercicio de derechos.

6. **Disposiciones Sancionatorias:** Establecen las consecuencias jurídicas del incumplimiento de las normas sustantivas.
7. **Parte Final (Disposiciones Complementarias):** Incluye cláusulas de derogación (expresa y específica), disposiciones transitorias para regular el paso de un régimen a otro, así como la fecha de entrada en vigor.
8. **Anexos:** Se reservan para tablas, listados, mapas o especificaciones técnicas que, por su naturaleza o extensión, dificultarían la lectura del articulado.

El artículo como unidad de base

El artículo es la unidad fundamental del texto normativo. La técnica moderna exige que cada artículo aborde un solo tema y, idealmente, sea lo más breve posible (véase la Tabla 11).

Tabla 11: Técnicas de redacción de artículos (consideraciones iniciales)

Regla de Redacción del Artículo	Justificación y Efecto
Un solo tema por artículo.	Evita confusiones y facilita las remisiones desde otras normas.
Un solo enunciado por párrafo.	Mejora la legibilidad y la precisión de la prescripción.
Una sola idea por enunciado.	Minimiza la ambigüedad sintáctica (anfibología).
Ausencia de motivación.	Los artículos deben ser preceptivos, no explicativos ni didácticos.
Numeración continua.	La numeración cardinal (1, 2, 3...) facilita la lectura y no debe alterarse tras modificaciones.

Estilo y Calidad Lingüística

El lenguaje de las normas debe ser técnico, pero accesible. La claridad es un requisito de la seguridad jurídica, pues permite que el ciudadano sepa a qué atenerse.

- **Brevedad y sencillez:** Se deben preferir frases cortas y construcciones directas (sujeto + verbo + predicado). El enunciado principal debe preceder a las subordinadas.
- **Voz activa y modo indicativo:** La voz activa identifica claramente al responsable de la acción. El uso del tiempo presente en el modo indicativo (p. ej., el órgano emite) es el estándar para los mandatos jurídicos.
- **Modalidades deónticas:** El uso de verbos como deber, poder o prohibir debe ser preciso para distinguir entre mandatos, facultades y prohibiciones. Se debe evitar la ambigüedad de los verbos con funciones duales.
- **Consistencia terminológica:** No se deben usar sinónimos por elegancia literaria. Un mismo concepto debe tener siempre el mismo nombre a lo largo de todo el texto normativo. Si se utiliza un término con un sentido distinto del habitual, es obligatorio incluir una definición en las disposiciones generales.

La Gestión Normativa en Argentina: Un Ecosistema en Evolución

En la Argentina, la gestión legislativa se rige por un conjunto de normas y prácticas que buscan profesionalizar la elaboración normativa y reducir la burocracia.

El Decreto 891/2017: Buenas Prácticas en Materia de Simplificación

Este decreto representa un hito en la política de mejora regulatoria del Sector

Público Nacional. Establece directrices obligatorias para el dictado de normas y la gestión de procesos administrativos. Sus puntos centrales son:

1. **Simplificación Normativa:** Las normas deben ser simples, claras y de fácil comprensión. El dictado de nuevas regulaciones que impongan cargas debe compensarse con la reducción del inventario normativo existente (principio de uno por uno o de reducción neta).
2. **Presunción de Buena Fe:** Se reconoce la buena fe del ciudadano, lo que le permite acreditar situaciones mediante declaraciones juradas y elimina la exigencia de certificaciones innecesarias.
3. **Gobierno Digital e Interoperabilidad:** Los organismos deben intercambiar información entre sí para no solicitar al ciudadano datos que el Estado ya posee. Se promueve el uso de expedientes electrónicos y firmas digitales.
4. **Medición de Costo-Beneficio:** Se insta a los organismos a evaluar la implementación de sus normas y a medir el impacto económico.

El Procedimiento Administrativo y el Decreto 333/1985

El Decreto 333/1985 (y su actualización mediante el Decreto 576/2003) fija las normas para la elaboración, redacción y diligenciamiento de los proyectos de actos administrativos en el Poder Ejecutivo Nacional. Todo proyecto debe estar acompañado de una Memoria o Mensaje que detalle los fundamentos, los motivos, la naturaleza de la decisión y los resultados esperados. Esta estructura formal garantiza que el proceso de decisión sea transparente y que existan antecedentes técnicos y jurídicos suficientes (dictámenes) antes de la firma del Presidente o de los Ministros.

El Digesto Jurídico Argentino (DJA): Racionalización del Ordenamiento

Tras 160 años de producción legislativa, la Argentina enfrentó una crisis de

seguridad jurídica debido a la dispersión y la superposición de normas. El Digesto Jurídico Argentino, aprobado por la Ley 26.939, es la obra de consolidación que depuró el sistema y estableció qué leyes nacionales generales están vigentes.

El DJA no es una mera recopilación; implicó revisar más de 30.000 leyes para identificar abrogaciones implícitas y normas caducas. El resultado es un cuerpo normativo ordenado por materias (administrativo, penal, civil, etc.) y numerado de forma lógica. Para mantener esta coherencia, el Digesto se actualiza periódicamente por una Comisión Bicameral, lo que asegura que las nuevas leyes se encuadren en las categorías temáticas preestablecidas.

Evaluaciones de Impacto Especializadas

La gestión legislativa contemporánea debe integrar análisis específicos que protejan los valores constitucionales y convencionales, más allá de la eficiencia económica.

Impacto Ambiental: Evaluación ex ante y participación

La Ley General del Ambiente (N° 25.675) y la normativa de la Ciudad Autónoma de Buenos Aires (Ley 123) establecen la obligatoriedad de la Evaluación de Impacto Ambiental (EIA) para todo emprendimiento público o privado susceptible de causar un efecto ambiental relevante.

El procedimiento técnico-administrativo incluye la presentación del Estudio de Impacto Ambiental (EsIA) por parte del proponente, su revisión por la autoridad de aplicación y la realización de una Audiencia Pública para discutir el proyecto con la ciudadanía. El resultado final es la Declaración de Impacto Ambiental (DIA) o el Certificado de Aptitud Ambiental, que habilita el inicio de las actividades bajo condiciones específicas de mitigación.

Impacto por razón de género

El análisis de impacto por razón de género es un requisito creciente en la elaboración de políticas y leyes. Su objetivo es detectar si una norma, aunque parezca neutral, puede generar o profundizar desigualdades entre varones y mujeres. Este análisis requiere una mirada interseccional que reconozca que los impactos varían según la situación socioeconómica, la identidad de género y la ubicación geográfica (Díaz et al., 2025). Un ejemplo claro en la Argentina ha sido la evaluación de las leyes sobre interrupción voluntaria del embarazo y de cuidado integral, en la que la evidencia sobre la distribución inequitativa del tiempo de cuidado fue central para el diseño de la política pública.

Impacto en Mipymes y Competitividad

Las regulaciones suelen generar costos de cumplimiento que afectan desproporcionadamente a las micro, pequeñas y medianas empresas (Mipymes) debido a su menor escala. La mejora regulatoria exige el diseño de medidas de mitigación para este sector (véase la Tabla 12).

Tabla 12: Medidas de mitigación para el sector Mipymes

Medida de mitigación	Mecanismo de Implementación
Exenciones totales o parciales.	Basadas en el número de empleados o en la facturación anual.
Periodos de transición.	Plazos más amplios para el cumplimiento de las nuevas normas.
Simplificación de reportes.	Menor frecuencia o menor complejidad en la entrega de información.
Tarifas diferenciadas.	Reducción de costos en trámites y licencias administrativas.

La gestión legislativa y el análisis de impacto normativo se consolidan hoy como las herramientas fundamentales para una gobernanza democrática y eficaz. La transición desde un modelo reactivo de producción de leyes hacia uno proactivo y basado en evidencia permite al Estado intervenir de manera quirúrgica en los problemas públicos, evitando la hipertrofia normativa y garantizando que cada regulación aporte un beneficio neto a la sociedad.

Los desafíos futuros se centran en la institucionalización de estas prácticas en todos los niveles de gobierno, en la mejora de la calidad de los datos públicos y en el fortalecimiento de los mecanismos de participación ciudadana. Una norma bien redactada y técnicamente sólida no solo reduce la litigiosidad y la incertidumbre, sino que también fortalece la confianza del ciudadano en las instituciones y promueve el desarrollo sostenible e inclusivo. La técnica legislativa, lejos de ser una cuestión de forma, es, en última instancia, una garantía de la libertad y la igualdad ante la ley.

Capítulo V

Transparencia: Facilitación del acceso a la información pública mediante síntesis de documentos complejos

El paradigma de la administración pública contemporánea ha experimentado una metamorfosis irreversible, transitando de una cultura del secreto burocrático a un modelo de apertura radical, fundamentado en el derecho al acceso a la información. En este contexto, la transparencia no debe entenderse simplemente como la disponibilidad de datos brutos, sino como la capacidad de las instituciones para hacer que dicha información sea comprensible y accionable para la ciudadanía. El acceso a la información pública se erige hoy como una pieza fundamental del sistema de rendición de cuentas, actuando los órganos garantes como intermediarios decisivos entre el gobierno y los ciudadanos. No obstante, la efectividad de este derecho se ve amenazada por la creciente complejidad técnica y lingüística de la documentación estatal, lo que genera una brecha cognitiva que solo puede superarse mediante la implementación estratégica de tecnologías de síntesis documental y de procesamiento del lenguaje natural.

Fundamentos interpretativos de la transparencia y el acceso a la información

La transparencia gubernamental se define operativamente como el conjunto de programas y mecanismos que permiten dar a conocer información confiable y accesible sobre el desempeño institucional y la ejecución de los recursos públicos. Esta

definición, aunque robusta en términos legales, requiere receptores externos capaces de procesar la información recibida; de lo contrario, la apertura se convierte en un ejercicio estéril (Mamani y Arisaca, 2021). El principio de transparencia, arraigado en marcos constitucionales como el artículo 8º de la Constitución chilena, establece que los actos y resoluciones de los órganos del Estado son públicos, así como sus fundamentos y procedimientos.

En el análisis comparativo de los marcos normativos de Chile y México, se observa que la transparencia se articula en dos dimensiones complementarias: la transparencia activa y el derecho de acceso, o transparencia pasiva. La transparencia activa impone a los órganos de la Administración del Estado la obligación de mantener a disposición permanente del público, a través de sitios electrónicos, antecedentes actualizados como su estructura orgánica, facultades, funciones y las transferencias de fondos públicos. Por otro lado, la transparencia pasiva se ejerce mediante solicitudes específicas de información que los ciudadanos dirigen a la autoridad, la cual debe responder conforme a los principios de relevancia, de libertad de información y de máxima divulgación (véase la Tabla 13).

Tabla 13: Características de la transparencia activa y pasiva

Dimensión de la Transparencia	Obligación Principal	Mecanismo de Implementación	Principios Rectores
Transparencia Activa	Publicación proactiva y permanente de información relevante.	Sitios institucionales, portales web de transparencia.	Facilitación, actualización mensual, gratuidad.
Transparencia Pasiva	Respuesta a las solicitudes de información	Formularios presenciales o electrónicos (O.I.R.S.).	Máxima divulgación, oportunidad, no discriminación.

	ciudadanas.		
--	-------------	--	--

La Ley N° 20.285 sobre Acceso a la Información Pública en Chile, vigente desde 2009, ejemplifica la vanguardia regional al incorporar el test de daño como criterio para resolver excepciones a la publicidad y al principio de divisibilidad, que obliga a entregar la parte pública de un documento incluso si contiene secciones reservadas. Sin embargo, la mera entrega de expedientes administrativos, contratos o fallos judiciales de cientos de páginas no garantiza que el ciudadano comprenda los fundamentos de las decisiones que afectan su vida. Aquí es donde la síntesis documental se convierte en una herramienta de democratización.

La barrera del lenguaje oscuro y la necesidad de simplificación

La claridad, definida por la Real Academia Española como aquello que se comprende sin duda ni ambigüedad, es la base de la transparencia en la función pública. No obstante, la realidad administrativa suele estar dominada por un lenguaje oscuro o críptico que aleja a la ciudadanía. El movimiento por el Lenguaje Claro busca revertir esta tendencia enfocándose en la composición de textos que presenten las ideas de manera directa y concisa, sin excesos estilísticos (Ato, 2021).

En España, la Ley de Derecho a la Defensa de 2024 ha marcado un hito al reconocer el derecho ciudadano a que la Administración de Justicia se comunique en un lenguaje claro. Este impulso legal se complementa con la norma UNE-ISO 24495-1:2024, que establece principios internacionales para garantizar la comprensibilidad de los documentos desde la primera lectura.

Es imperativo diferenciar entre un lenguaje claro y una lectura fácil. Mientras que el primero busca que la información sea fácil de entender para la población general, la Lectura Fácil es una técnica de redacción y diseño dirigida a personas con dificultades de comprensión lectora, que requiere validaciones específicas por parte de los usuarios finales. Ambos enfoques son esenciales para la inclusión, pero el lenguaje claro es el estándar mínimo exigible para la transparencia y el buen gobierno.

La implementación del lenguaje claro genera ventajas tangibles: aumenta la confianza de las personas en las instituciones, disminuye la insatisfacción por la falta de transparencia y supone un ahorro económico al reducir el tiempo de gestión y el número de consultas o errores en los trámites. Además, facilita la traducción automática y los sistemas de procesamiento de lenguaje para personas con discapacidad visual, cerrando brechas de accesibilidad.

Arquitecturas de Inteligencia Artificial para la síntesis de documentos públicos

La síntesis de documentos complejos mediante inteligencia artificial (IA) no es una mera automatización de resúmenes, sino un proceso sofisticado de procesamiento de lenguaje natural (PLN) que busca transformar la opacidad en inteligibilidad. Las herramientas de IA analizan contratos, expedientes regulatorios y transcripciones judiciales para extraer cláusulas esenciales y presentarlas en un lenguaje sencillo. En este ámbito, existen dos enfoques técnicos predominantes: la síntesis extractiva y la síntesis abstractiva (Peña, 2022).

La síntesis extractiva identifica y selecciona las secciones más importantes del texto original para elaborar un resumen. Modelos como T5 (Text-to-Text Transfer Transformer) pueden ser entrenados para esta tarea, redefiniendo cada problema de

PLN como un problema de generación de texto. Por otro lado, la síntesis abstractiva genera un texto nuevo que parafrasea el contenido original, lo que permite una mayor fluidez y una adaptación a un lenguaje claro, aunque requiere modelos más potentes, como BART (Bidirectional and Auto-Regressive Transformers), para evitar la pérdida de precisión jurídica.

Tabla 14: Características técnicas de modelos IA/PLN

Modelo de IA / PLN	Origen / Desarrollador	Aplicación en el sector público	Características Técnicas
RoBERTalex	PlanTL (España)	Análisis de la legislación española y de los documentos del BOE.	Entrenado con 3,6 GB de legislación y doctrina fiscales.
BART-legal-base-es	HuggingFace (mrm8488)	Resumen de textos legales en español.	Arquitectura encoder-decoder para síntesis abstractiva.
MarianMT	Helsinki-NLP	Traducción y síntesis multilingües de documentos jurídicos.	Soporte para la traducción de lenguas romances al inglés.
Legal-Pegasus	NSI319	Resumen de documentos legales complejos.	Especializado en la estructura de los escritos judiciales.

El desarrollo de modelos específicos para el dominio legal en español es crítico. Modelos como RoBERTalex han sido entrenados con corpus que incluyen la Constitución Española, leyes del BOE, dictámenes del Consejo de Estado y sentencias, lo que les permite comprender la jerga legal como si fuera un idioma independiente. Asimismo, el uso de conjuntos de datos sintéticos, como el corpus *ALIA-legal*-

administrative, que incluye millones de pares de instrucciones y respuestas, permite ajustar estos modelos para que respondan de manera coherente a las consultas ciudadanas sobre procedimientos administrativos complejos.

Implementación de IA en la Justicia: El caso de Argentina y Brasil

El sector judicial es quizá el área en la que la síntesis de documentos complejos tiene un impacto más disruptivo. En América Latina, destacan proyectos que han logrado reducir significativamente los tiempos de procesamiento y mejorar la transparencia para el justiciable.

Prometea y el Ministerio Público Fiscal de la Ciudad de Buenos Aires

Prometea se ha consolidado como la primera inteligencia artificial de Latinoamérica al servicio de la justicia, diseñada inicialmente para optimizar procesos internos y luego ampliada para mejorar la relación ciudadanía-Estado. Su principal capacidad radica en la predicción de soluciones para expedientes judiciales a partir del análisis de miles de documentos previos (Corvalán, 2018). En la Justicia porteña, Prometea automatiza procesos reiterativos, como las causas por infracciones de tránsito o por alcohol al volante, permitiendo que un documento que demandaba 60 minutos de trabajo manual se confeccione en solo 3 minutos. Las mediciones de impacto de Prometea arrojan resultados notables:

- **Eficiencia Temporal:** Reducción del tiempo de elaboración de dictámenes, lo que permite a los funcionarios dedicarse a casos más complejos.
- **Precisión:** Eliminación de errores de tipeo y de cálculo en la confección de

despachos monitorios y de regulaciones de honorarios.

- **Igualdad y Claridad:** Unificación de documentos y estandarización de los cálculos de intereses, garantizando que el ciudadano reciba información coherente sobre su proceso.

El Proyecto Victor y la Repercusión General en el Supremo Tribunal Federal de Brasil

En Brasil, el Supremo Tribunal Federal (STF) ha sido pionero en el uso de la herramienta Victor, lanzada en 2018 para auxiliar en el análisis de recursos extraordinarios recibidos de todo el país. Víctor utiliza modelos de aprendizaje profundo (*machine learning*) para reconocer patrones y clasificar procesos según los temas de repercusión general. Sus funciones principales incluyen la conversión de imágenes en texto, la separación de piezas procesales y la identificación de temas de alta incidencia para agilizar la tramitación.

En 2023, el STF lanzó un *Chamamento Público* para desarrollar prototipos de IA generativa destinados a crear sumarios automatizados de procesos judiciales. Los resultados mostraron una alta participación de empresas que utilizaron modelos de la familia GPT (OpenAI), Claude y Llama, integrando técnicas de *Retrieval Augmented Generation* (RAG) para evitar alucinaciones y asegurar que el resumen se ciña estrictamente a la realidad de las piezas procesales. Estos sumarios facilitan que los magistrados y la ciudadanía comprendan rápidamente el núcleo de las controversias judiciales extensas.

El desafío de la transparencia algorítmica y el derecho al algoritmo

A medida que las administraciones públicas delegan la síntesis y clasificación de documentos en sistemas de IA, surge una nueva frontera en el derecho de acceso: la transparencia algorítmica. El uso de algoritmos no puede ser una caja negra que oculte la lógica de las decisiones públicas. En España, la Sentencia del Tribunal Supremo del 11 de septiembre de 2025 (caso BOSCO) ha sentado un precedente histórico al reconocer el derecho de acceso al código fuente de las aplicaciones informáticas utilizadas por la Administración para determinar derechos sociales.

La controversia sobre la aplicación BOSCO, utilizada para comprobar los requisitos de los solicitantes del bono social eléctrico, evidenció que la protección de la propiedad intelectual no puede servir de excusa sistemática para ocultar el algoritmo cuando este decide sobre beneficios ciudadanos. El tribunal determinó que:

1. **El algoritmo es información pública:** el código fuente de los sistemas automatizados no puede ser tratado como un simple secreto técnico.
2. **Prevalencia de la transparencia:** El principio de transparencia administrativa prevalece sobre los límites de la propiedad intelectual cuando el titular es la propia Administración.
3. **Fiscalización y Auditoría:** Es necesario facilitar el acceso al código para permitir auditorías independientes que verifiquen la ausencia de sesgos o errores en la toma de decisiones automatizada.

Este fallo obliga a las entidades públicas a anticiparse a las exigencias de transparencia desde la fase de contratación, asegurando que los proveedores de software de IA faciliten el acceso controlado a la lógica interna de los sistemas.

Riesgos y barreras en la implementación de la IA pública según la OCDE

A pesar de las promesas de eficiencia, los gobiernos enfrentan desafíos estructurales que dificultan el uso estratégico de la IA para la transparencia. Según la OCDE (2024), solo el 15 % de los países miembros cuenta con un marco público de inversión en IA, lo que genera una dependencia de proyectos piloto que rara vez escalan a la organización en su conjunto (OECD, 2025).

Tabla 15: Desafíos en la implementación de la inteligencia artificial

Tipo de Desafío	Descripción del Riesgo	Impacto en la Transparencia
Aversión al riesgo	Miedo al fracaso y a la erosión de la confianza institucional (el 63% de los líderes teme a la GenIA).	Parálisis en la adopción de herramientas que podrían simplificar la comunicación.
Aversión Algorítmica	Resistencia de los servidores públicos a adoptar algoritmos, especialmente tras errores iniciales.	Persistencia de procesos manuales lentos y opacos.
Brecha de Habilidades	Falta de talento calificado en el sector público para gestionar sistemas de IA.	Implementaciones deficientes o falta de supervisión crítica humana.
Sistemas Obsoletos	Infraestructura de TI obsoleta que no admite la integración con modelos de PLN modernos.	Dificultad para procesar y sintetizar archivos en formatos no legibles.

La OCDE advierte que existe una priorización de lo negativo, en la que los gobiernos se centran excesivamente en mitigar riesgos sin aprovechar los beneficios

transformadores de la IA para la rendición de cuentas. Para mitigar estos sesgos, se recomienda crear infraestructuras digitales centrales y marcos de inversión que faciliten la transición de la experimentación a la implementación responsable (véase la Tabla 15).

El marco ético internacional: Recomendaciones de la UNESCO

Para que la síntesis de documentos complejos sea verdaderamente facilitadora del acceso a la información, debe regirse por los valores adoptados por los 193 Estados miembros de la UNESCO en 2021. La Recomendación sobre la Ética de la IA establece que los sistemas deben estar al servicio de la humanidad y prevenir daños, con base en la dignidad humana y el bienestar.

Los principios fundamentales aplicables a la síntesis documental incluyen:

- **Transparencia y Explicabilidad:** Los procesos de diseño e implementación de la IA deben ser justificables. Las personas deben saber cuándo están interactuando con un sistema de IA y tener acceso a las razones que subyacen a una síntesis o a una clasificación.
- **Supervisión y Determinación Humana:** Siempre debe ser posible atribuir la responsabilidad ética y jurídica a personas físicas. El juicio humano debe prevalecer en la verificación final de los resultados generados por la IA, a fin de corregir posibles alucinaciones o sesgos.
- **Derecho a la Intimidad y Protección de Datos:** El uso de IA para sintetizar expedientes no debe vulnerar los derechos fundamentales de la intimidad. Se requiere el anonimato de los datos antes de su procesamiento en herramientas externas.

- **Responsabilidad y Rendición de Cuentas:** Los Estados deben adoptar medidas para que los actores de la IA (incluido el sector privado contratista) rindan cuentas del impacto social de sus sistemas.

En Argentina, el Poder Judicial del Chaco ha incorporado estos principios en su Guía de Uso de IA (2025), prohibiendo explícitamente la delegación de la toma de decisiones en algoritmos y estableciendo la obligatoriedad de la supervisión humana permanente. Esta guía advierte sobre los riesgos de alucinaciones (respuestas ficticias pero coherentes) y sesgos en los datos de entrenamiento que podrían derivar en resultados injustos.

El papel del código abierto y las infraestructuras digitales abiertas

La sostenibilidad de los sistemas de transparencia asistida por IA depende en gran medida de evitar la dependencia de proveedores cerrados (*vendor lock-in*). El sector público aprecia cada vez más las soluciones de código abierto que facilitan la interoperabilidad y la auditoría del código. Herramientas como Nextcloud Hub, que integra asistentes de IA locales, permiten a las instituciones colaborar en la creación y síntesis de contenidos manteniendo el control total sobre los datos.

La colaboración intersectorial a través de plataformas como GitHub facilita que los departamentos gubernamentales compartan herramientas de análisis de proyectos y repositorios de datos abiertos. Por ejemplo, el uso de herramientas como AskGit permite realizar consultas SQL sobre los repositorios de código para entender cómo se organizan los proyectos internos y cómo colaboran los empleados, lo que promueve una cultura de transparencia técnica.

Impacto en la participación ciudadana y la confianza institucional

El objetivo último de la simplificación de documentos complejos es fortalecer la democracia mediante una ciudadanía informada y empoderada. En América Latina, la ola de escándalos de corrupción ha socavado la confianza en los gobiernos, por lo que el fortalecimiento de la calidad institucional es una necesidad urgente. La evidencia sugiere que la simplificación de las regulaciones y una mejor consulta con los ciudadanos pueden contribuir a recuperar la confianza perdida.

La participación ciudadana, promovida mediante la transparencia, actúa como una herramienta para mejorar la gobernanza y la rendición de cuentas. En regiones con democracias incipientes, estas iniciativas son fundamentales para superar desafíos como la pobreza y la desigualdad. Por ejemplo, en Brasil y Chile, la satisfacción ciudadana con servicios básicos como la educación y la salud ha disminuido, lo que subraya la importancia de que la administración pública sea más eficiente y transparente en la comunicación de sus resultados.

La digitalización de procesos gubernamentales, que pasa del expediente en papel a modalidades digitales, permite automatizar y agilizar los procedimientos administrativos, mejorando la relación entre el gobierno y los gobernados. La IA generativa, al facilitar una comunicación más clara, ayuda a evaluar y ajustar el lenguaje para que el derecho a entender sea una realidad efectiva y no una mera declaración de intenciones.

Futuro de la síntesis documental en la transparencia pública

La evolución hacia un Estado abierto e innovador requiere que la IA no solo se utilice para mejorar la eficiencia interna, sino también para empoderar a los actores externos. El potencial de la IA para mejorar la toma de decisiones, la interpretación del presente y los pronósticos sobre el futuro ofrece enormes oportunidades para el bienestar social. Sin embargo, esto requiere un cambio cultural en el sector público, que transite de la aversión al riesgo a una experimentación responsable guiada por valores humanos.

Las estrategias nacionales de IA, como las adoptadas por más de 60 países, deben centrarse en la creación de marcos de protección de datos, la inversión en conectividad y la capacitación de los servidores públicos. En Argentina, la Estrategia Nacional de Inteligencia Artificial propone fomentar el trabajo intersectorial en datos para mejorar los servicios públicos y posicionar al país como líder regional en este cambio de paradigma.

En síntesis, la facilitación del acceso a la información pública mediante la síntesis de documentos complejos es un pilar indispensable de la modernización del Estado. La integración de tecnologías de PLN, bajo marcos éticos estrictos y principios de un lenguaje claro, permite transformar la burocracia técnica en un diálogo abierto con la ciudadanía. El éxito de esta transición dependerá de la capacidad de los gobiernos para equilibrar la innovación tecnológica con la transparencia algorítmica y la supervisión humana, asegurando que la inteligencia artificial sea siempre una herramienta al servicio del bien social y de la democracia.

Capítulo VI

Mejora en el diseño de instrumentos de recolección de datos y análisis de respuestas abiertas

La recolección de datos mediante encuestas ha trascendido su función original como herramienta estadística para convertirse en un ecosistema sofisticado en el que convergen la psicología cognitiva, el diseño de la experiencia de usuario y la inteligencia artificial avanzada. En la actualidad, la optimización de un instrumento de medición no se limita a la redacción de interrogantes, sino que exige una arquitectura estratégica que garantice la integridad de los datos, minimice la fatiga del encuestado y permita un procesamiento profundo de la información cualitativa. El diseño de encuestas en la era digital debe responder a una saturación informativa sin precedentes, donde la calidad del dato prima sobre la cantidad, y donde la capacidad de transformar respuestas abiertas en conocimiento accionable define la ventaja competitiva de las organizaciones.

Arquitectura estratégica y diseño de instrumentos

La eficacia de una encuesta comienza con la articulación de un objetivo de investigación nítido y la formulación de hipótesis que sirvan de brújula para cada componente del diseño. Antes de proceder a la creación de ítems, se requiere establecer con precisión el propósito de la indagación, la población objetivo y las acciones derivadas de los hallazgos. Este enfoque garantiza que cada pregunta sea estrictamente necesaria, eliminando redundancias que degradan la experiencia del usuario y disminuyen la validez del conjunto de datos.

Estructura lógica y secuencia de flujo

La disposición secuencial de las preguntas ejerce una influencia determinante en la tasa de finalización y en la sinceridad de las respuestas. La práctica profesional prescribe el uso de un modelo de embudo que comienza con preguntas de introducción que despierten el interés y establezcan un clima de confianza. Estas interrogantes iniciales deben ser sencillas, claras y desprovistas de carga emocional o de confidencialidad excesiva. A medida que el encuestado se familiariza con el instrumento, el flujo debe transitar hacia bloques temáticos organizados de manera lógica, moviéndose de lo general a lo particular y de lo simple a lo complejo.

Los datos demográficos, socioeconómicos e identificativos deben reservarse para el tramo final del cuestionario. Esta estrategia permite que el participante establezca un vínculo de colaboración con la herramienta antes de enfrentar preguntas que podrían activar sesgos de deseabilidad social o mecanismos de defensa. Adicionalmente, la implementación de lógica de ramificación y de saltos condicionales permite personalizar el recorrido, omitiendo secciones irrelevantes para el perfil específico del encuestado, lo que optimiza el tiempo de respuesta y reduce la carga cognitiva.

Tipología de preguntas y escalas de medición

La elección del tipo de pregunta debe estar alineada con la naturaleza de la variable que se pretende medir. Las preguntas cerradas son fundamentales para la cuantificación y la comparación estadísticas, mientras que las abiertas ofrecen la profundidad necesaria para comprender las motivaciones subyacentes. En el diseño de preguntas de opción múltiple, es imperativo que las alternativas sean mutuamente excluyentes y colectivamente exhaustivas, evitando que el encuestado se encuentre ante un dilema de categorización ambiguo.

Las escalas de Likert y otras formas de medición de actitudes requieren un diseño equilibrado. Se recomienda el uso de un número impar de opciones de respuesta —típicamente cinco o siete— para incluir un punto medio neutro que refleje la posición de quienes carecen de una opinión polarizada. Asimismo, la investigación sugiere que el uso de etiquetas verbales descriptivas para cada punto de la escala es superior al de meros números, ya que reduce la ambigüedad interpretativa del encuestado (Matas, 2018).

Tabla 16: Escalas de Likert y otras formas de procesar entrevistas/encuestas

Tipo de Instrumento	Aplicación Específica	Ventaja Metodológica
Escala de Likert	Medición de actitudes y niveles de satisfacción	Captura gradientes de intensidad emocional
Ranking / Ordenación	Evaluación de prioridades y preferencias	Obliga a la jerarquización de conceptos
Matriz de Selección	Evaluación de múltiples atributos con misma escala	Maximiza la eficiencia de espacio y tiempo
Suma Constante	Distribución de recursos o importancia	Proporciona datos de intervalo con alta precisión
Mapa de Calor / Zona Activa	Evaluación visual de diseños o imágenes	Recolecta datos espaciales y de atención

La redacción de los ítems debe evitar términos técnicos, acrónimos y lenguaje especializado que no sean familiares para la población general. Las preguntas compuestas o de doble sentido, que solicitan comentarios sobre dos temas distintos en una misma frase, deben eliminarse, ya que imposibilitan determinar qué aspecto

de la pregunta se está respondiendo realmente. La precisión semántica es vital: evitar adverbios vagos como frecuentemente, mucho o bastante y sustituirlos por unidades de medida concretas o por frecuencias específicas mejora drásticamente la fiabilidad de los resultados.

Fundamentos psicométricos y validación

Para que un instrumento de recolección de datos sea científicamente válido, debe cumplir con los requisitos de confiabilidad y validez. La confiabilidad garantiza que el instrumento produzca resultados consistentes en condiciones similares, mientras que la validez asegura que mida realmente el constructo teórico que pretende medir.

Metodologías de validación técnica

El proceso de validación es multifacético y debe iniciarse en etapas tempranas del desarrollo del cuestionario. La validez de contenido suele establecerse mediante paneles de expertos que evalúan la representatividad de los ítems respecto del dominio del conocimiento estudiado. Sin embargo, la validez de constructo es la más exigente, pues requiere un análisis de la relación teórica entre conceptos y la contrastación empírica mediante técnicas estadísticas, como el análisis factorial.

La confiabilidad se mide comúnmente a través de coeficientes de consistencia interna, como el alfa de Cronbach, o mediante métodos de estabilidad como el test-retest. El método de formas alternativas o paralelas también se emplea para asegurar que las variaciones menores en la redacción no alteren el núcleo de la medición. En investigaciones de alta precisión, la triangulación de métodos y la validación interna, mediante la devolución de los resultados a los informantes para su corroboración, refuerzan la credibilidad de los hallazgos.

El rol de las entrevistas cognitivas

Las entrevistas cognitivas han emergido como una herramienta cualitativa esencial para la validación de cuestionarios, proporcionando evidencia sobre el proceso de respuesta que las métricas estadísticas a menudo ignoran. El objetivo fundamental es identificar fallos en la arquitectura mental del encuestado. Basándose en el modelo de Tourangeau, este método analiza cuatro etapas críticas:

1. **Comprensión:** Se evalúa si el encuestado interpreta el significado de las palabras y la intención de la pregunta de la manera esperada por el investigador.
2. **Recuperación:** Se indaga cómo el sujeto recupera en su memoria la información relevante para responder, detectando posibles sesgos de memoria o dificultades de acceso a dicha información.
3. **Juicio:** Se analiza el proceso de decisión en el que el encuestado evalúa la relevancia de la información recuperada y decide si emitirá una respuesta honesta o socialmente aceptable.
4. **Respuesta:** Se observa la etapa final de la traducción del juicio interno a las opciones de respuesta del cuestionario.

Las técnicas aplicadas durante estas entrevistas incluyen el pensar en voz alta (*think-aloud*), en el que el participante verbaliza su flujo de conciencia mientras completa el cuestionario, y el sondeo verbal (*verbal probing*), en el que el entrevistador formula preguntas inquisitivas específicas tras la respuesta del participante. Este último permite profundizar en conceptos ambiguos; por ejemplo, preguntar: ¿Qué significa para usted 'calidad de servicio'? Revela si los encuestados comparten el mismo marco conceptual que el equipo de diseño.

Diseño de experiencia de usuario (UX) y optimización digital

En el entorno digital contemporáneo, la interfaz de la encuesta es el canal que media la interacción entre el investigador y el dato. Un diseño UX deficiente actúa como una barrera que genera frustración, incrementa la tasa de rebote y degrada la calidad de la información recopilada. Invertir en optimización de UX no es una cuestión estética, sino una decisión estratégica que impacta directamente en el retorno de la inversión en investigación.

Usabilidad, accesibilidad y diseño responsivo

La usabilidad implica que el sistema sea intuitivo y que el usuario no deba dedicar esfuerzo cognitivo adicional a comprender su funcionamiento. La arquitectura de la información debe organizar los elementos de forma lógica para que todo esté donde el encuestado espera encontrarlo. La accesibilidad es otro pilar fundamental: diseñar para personas con discapacidades físicas o cognitivas no solo es un imperativo ético, sino que también amplía la representatividad de la muestra.

Dado que más del 50% de las interacciones digitales ocurren en dispositivos móviles, el diseño responsivo es obligatorio. Una encuesta que no se visualiza correctamente en un teléfono inteligente sufrirá un sesgo de selección involuntario, excluyendo segmentos de la población que utilizan predominantemente estos dispositivos. La velocidad de carga también es crítica; los retrasos técnicos son uno de los principales desencadenantes del abandono del sitio.

El enfoque Lean y encuestas in situ

Para mejorar la experiencia y reducir la tasa de abandono, se recomienda aplicar un enfoque Lean adaptado a la fase de desarrollo de la investigación. En etapas de ideación, el uso de preguntas abiertas permite identificar puntos de dolor, mientras que en etapas de validación, las escalas cerradas proporcionan la métrica necesaria para priorizar cambios. Las encuestas in situ, que se deslizan o aparecen como pop-ups durante la navegación, deben ser extremadamente breves y relevantes al contexto de la página en la que se activan para no interferir con la experiencia del usuario.

La frecuencia de las encuestas debe gestionarse con prudencia, mediante períodos de descanso, para evitar que un mismo usuario sea bombardeado con solicitudes repetitivas. Configurar las herramientas para que la encuesta desaparezca tras ser respondida o ignorada una vez es una práctica esencial para mantener la buena voluntad del usuario hacia la marca o la institución.

Análisis avanzado de respuestas abiertas mediante IA y NLP

Las respuestas abiertas constituyen el núcleo cualitativo de una encuesta, proporcionando que hay detrás de las cifras. Sin embargo, procesar este volumen de información de forma manual resulta ineficiente y propenso a sesgos interpretativos. El procesamiento de lenguaje natural (NLP) y el aprendizaje automático (*Machine Learning*) han revolucionado este campo, permitiendo un análisis sistemático y escalable de datos no estructurados.

Pipeline técnico del procesamiento de texto

El análisis automatizado de respuestas abiertas sigue un proceso estructurado

que transforma el lenguaje humano en vectores numéricos procesables por algoritmos. Este flujo comienza con la ingesta de datos y continúa con el preprocesamiento, una etapa crucial para eliminar el ruido informativo. El preprocesamiento incluye la tokenización, la normalización a minúsculas, la eliminación de *stop words* (palabras sin valor semántico, como preposiciones) y la lematización, que reduce las palabras a su raíz común (Shehzadi et al., 2025).

Una vez procesado, el texto se somete a la extracción de características mediante modelos como TF-IDF o representaciones densas (embeddings) como Word2Vec o BERT. Estos modelos permiten capturar no solo la frecuencia de las palabras, sino también el contexto semántico y las relaciones entre ellas, lo que facilita la identificación de temas complejos y sentimientos matizados.

Tipologías de análisis de sentimientos

El análisis de sentimientos moderno va más allá de la simple clasificación binaria (positivo/negativo). Se han desarrollado técnicas de alta granularidad que permiten una comprensión multidimensional:

- **Análisis de grano fino:** Utiliza escalas graduadas (muy positivo, positivo, neutral, negativo, muy negativo), útiles para industrias que requieren comprender niveles precisos de satisfacción.
- **Análisis basado en aspectos (ABSA):** Desglosa la opinión en características específicas. Por ejemplo, en una reseña de auriculares, el sistema puede identificar sentimientos positivos hacia el diseño, pero negativos hacia la calidad del sonido.
- **Detección de emociones:** Identifica estados específicos como ira, frustración, alegría o tristeza, utilizando léxicos especializados y modelos neuronales para evaluar el lenguaje subjetivo.

- **Análisis basado en la intención:** Detecta propósitos subyacentes, como la intención de compra, la urgencia de una queja o el deseo de cancelar un servicio.

La inteligencia artificial permite, además, una mayor objetividad. Mientras que un evaluador humano puede verse influenciado por una frase negativa inicial e ignorar matices positivos posteriores, los sistemas de IA están entrenados para sopesar todas las partes del texto de manera equitativa, ofreciendo una visión más equilibrada de la opinión del cliente.

Integración metodológica y triangulación de datos

La optimización de encuestas alcanza su máxima madurez cuando se integra en marcos de investigación mixta. Este enfoque no solo suma métodos cualitativos y cuantitativos, sino que también busca su convergencia para proporcionar una comprensión holística de fenómenos complejos que ningún solo método podría capturar por completo.

Funciones de la triangulación

La triangulación actúa como un mecanismo de control de calidad y de validación cruzada. Al utilizar múltiples puntos de referencia, se minimizan los sesgos inherentes a cualquier técnica aislada y se aumenta la credibilidad de los hallazgos. Los principales tipos de triangulación aplicables a la optimización de encuestas son:

1. **Triangulación Metodológica:** Emplear diferentes técnicas para estudiar un mismo objeto, como combinar una encuesta masiva (cuantitativa) con grupos focales o con la observación directa (cualitativa) para validar si lo declarado por los sujetos coincide con su comportamiento observable.
2. **Triangulación de Datos:** Utilizar diversas fuentes de información en distintos momentos o espacios. Por ejemplo, contrastar las respuestas de una encuesta de

satisfacción con registros de ventas, logs de actividad web y quejas en redes sociales.

3. **Triangulación de Investigadores:** Involucrar a múltiples analistas con trayectorias disciplinares distintas para interpretar el mismo conjunto de datos, lo que reduce el sesgo de confirmación individual y enriquece la profundidad del análisis.

Diseños de investigación mixta

La integración de datos puede seguir diferentes estructuras temporales y lógicas. El enfoque convergente busca la igualdad de importancia entre los datos cuantitativos y cualitativos, integrándolos en la etapa de análisis para obtener una visión equilibrada. El enfoque secuencial explicativo utiliza una fase cualitativa posterior para profundizar en los resultados cuantitativos atípicos. Por el contrario, el enfoque secuencial exploratorio utiliza hallazgos cualitativos iniciales para orientar y diseñar el instrumento cuantitativo final, asegurando que las preguntas sean culturalmente pertinentes y semánticamente precisas.

En la práctica contemporánea, se utilizan joint displays o visualizaciones conjuntas para presentar datos integrados, lo que permite a los investigadores y tomadores de decisiones visualizar cómo las estadísticas cuantitativas se correlacionan con las narrativas cualitativas en un único plano analítico.

El ecosistema tecnológico

En la actualidad, el panorama del software de análisis de encuestas ha evolucionado hacia la integración total de la inteligencia artificial, democratizando el acceso a análisis estadísticos y cualitativos que antes requerían conocimientos técnicos profundos. Las herramientas actuales no solo recolectan datos, sino que también

actúan como socios analíticos capaces de generar visualizaciones y detectar patrones de forma autónoma.

Comparativa de herramientas líderes

La elección del software debe basarse en la complejidad del proyecto y en los objetivos estratégicos de la organización. Las plataformas modernas pueden categorizarse según su fortaleza principal (véase la Tabla 17).

Tabla 17: Comparativa de software de análisis de datos

Herramienta	Especialidad Principal	Funcionalidad Clave de IA
MonkeyLearn	Análisis de datos no estructurados	Modelos personalizados de clasificación de texto y sentimiento.
Power BI + Copilot	Business Intelligence y dashboards corporativos	Generación de reportes mediante lenguaje natural y conectividad masiva
Tableau + Einstein GPT	Storytelling visual avanzado	Consultas predictivas y análisis exploratorio intuitivo
ATLAS.ti	Investigación cualitativa profunda	Codificación automática y análisis semántico de redes
Excelmatic	Limpieza y preparación de datos	Transformación de hojas de cálculo mediante prompts conversacionales
SurveyMonkey	Plantillas y distribución masiva	Análisis de sentimientos integrado y bancos de preguntas validados

Innovaciones en automatización y visualización

Las herramientas no-code permiten que equipos de marketing y analistas de negocios realicen tareas complejas, como la detección de valores atípicos, el tratamiento de valores faltantes y la elaboración de estadísticas descriptivas, sin escribir una sola línea de código (Bailón & Pico, 2025). Plataformas como Polymer Search o ClickUp destacan por su capacidad de convertir datos crudos de encuestas en dashboards interactivos en cuestión de segundos, lo que facilita una respuesta ágil a las tendencias del mercado.

Además, la integración con asistentes de inteligencia artificial conversacional (como ChatGPT y Code Interpreter) permite a los investigadores cargar archivos de datos y solicitar análisis exploratorios complejos, la limpieza de datos y la generación de hipótesis de manera casi instantánea, acelerando drásticamente el ciclo de vida de la investigación.

Ética, privacidad y mitigación de sesgos algorítmicos

El uso masivo de la inteligencia artificial en la optimización y el análisis de encuestas plantea dilemas éticos y legales sin precedentes. La automatización no es inherentemente neutral; los algoritmos pueden replicar y amplificar sesgos presentes en los datos de entrenamiento, lo que puede llevar a decisiones discriminatorias en ámbitos tan sensibles como la contratación, la educación o la salud.

Sesgo algorítmico y Diseño Sensible al Valor (VSD)

El sesgo algorítmico a menudo se debe a la falta de representatividad de los datos iniciales o a la opacidad de los modelos de caja negra. El enfoque del Diseño

Sensible al Valor (VSD) aboga por una postura proactiva en la que la tecnología se diseña de manera moral y ética desde su origen, reconociendo y reparando las desigualdades sociales existentes en lugar de simplemente automatizarlas.

Para mitigar estos riesgos, las organizaciones deben implementar estrategias de diseño ético y transparente que incluyan:

- **Explicabilidad e interpretabilidad:** Asegurar que los sistemas de IA permitan comprender cómo se ponderan las variables y cómo se llega a una conclusión específica, lo que facilita la supervisión humana.
- **Auditorías de equidad (Fairness):** Evaluar los resultados del modelo para detectar discriminación sistemática contra grupos específicos por motivos de género, etnia o edad. Esto implica el preprocesamiento de datos para balancear las muestras y la validación de variables significativas, de modo que las decisiones se basen en criterios legítimos.
- **Gobernanza y Responsabilidad:** Establecer evaluaciones de impacto y mecanismos de rendición de cuentas que identifiquen responsabilidades claras ante fallos algorítmicos o decisiones injustas.

Desafíos de privacidad y marcos regulatorios

El entrenamiento de modelos de IA requiere volúmenes masivos de datos, lo que pone en riesgo la privacidad personal. Los datos recolectados en encuestas pueden ser vulnerables a sustracciones o a su uso indebido para el perfilamiento no consentido de individuos. Es imperativo que el diseño de las encuestas incorpore principios de privacidad por diseño, utilizando técnicas de anonimización y cumpliendo rigurosamente con normativas como el GDPR o la HIPAA (Sánchez, 2024).

La transparencia hacia el encuestado sobre quién es el investigador, para qué

se utilizará su información y cómo se protegerá su anonimato es fundamental para mantener la integridad ética de la investigación y la confianza del público en las herramientas digitales de recolección de datos.

La optimización de encuestas representa la culminación de un proceso de integración entre la ciencia del comportamiento y la computación avanzada. La mejora del diseño de los instrumentos ya no puede entenderse únicamente como una tarea de redacción literaria; es una ingeniería de la información que prioriza la experiencia del usuario para salvaguardar la pureza del dato inicial. La implementación de metodologías de validación rigurosas, como las entrevistas cognitivas, garantiza que el puente entre el constructo teórico y la respuesta del participante sea sólido y libre de distorsiones.

En el plano analítico, la transición del procesamiento manual al análisis automatizado mediante NLP e IA permite a las organizaciones capturar la riqueza de las respuestas abiertas a escala masiva. Sin embargo, esta capacidad técnica conlleva una responsabilidad ética profunda: la necesidad de auditar y mitigar los sesgos algorítmicos para asegurar que la inteligencia artificial sea una herramienta de inclusión y no de discriminación.

Finalmente, el futuro de la recolección de datos reside en la flexibilidad y la integración. Las plataformas tecnológicas continuarán evolucionando hacia sistemas más intuitivos y conversacionales, permitiendo que la investigación mixta y la triangulación de datos se conviertan en la norma y no en la excepción. Al combinar la precisión de las métricas cuantitativas con la profundidad narrativa de las respuestas abiertas, los investigadores pueden ofrecer una visión verdaderamente holística que transforme los datos brutos en sabiduría estratégica y en decisiones informadas para el bienestar social y el éxito organizacional.

Conclusión

La productividad en el sector público se potencia mediante la automatización de tareas mundanas y repetitivas, lo que permite a los servidores públicos centrarse en actividades de mayor valor estratégico y humano. En América Latina, se estima que la IAGen podría impactar hasta el 40% del total de horas de trabajo, ya sea mediante la automatización completa o el aumento de las capacidades humanas. En América Latina, específicamente, se proyecta que el impacto en las horas laborales sea uno de los más altos de la región, lo que subraya la urgencia de integrar estas herramientas de manera planificada para evitar disrupciones laborales negativas y fomentar el reskilling de la fuerza laboral pública.

Un ejemplo concreto de este aumento de productividad se observa en la administración tributaria y en la gestión de finanzas públicas. La IA permite mejorar el diseño de la regulación, detectar transacciones indebidas y evaluar riesgos de integridad con una precisión y una velocidad superiores a los métodos tradicionales.

La IAGen redefine la relación Estado-ciudadano al permitir una comunicación multicanal efectiva y respuestas en tiempo real. Los asistentes virtuales de nueva generación van más allá de los simples árboles de decisión de los chatbots antiguos; ahora son capaces de comprender contextos complejos y asistir a los ciudadanos en la cumplimentación de trámites, la búsqueda de servicios y la interpretación de normativas.

Esta capacidad de respuesta va más allá de la atención al cliente. En casos de crisis o desastres naturales, la IAGen facilita la creación de escenarios de evacuación, reduce los impactos y activa los planes de mitigación con mayor rapidez y precisión. La opción de realizar simulaciones en un entorno sandbox virtual permite a los gobiernos experimentar con cambios en la infraestructura o en las políticas de transporte antes de implementarlos físicamente, analizando posibles consecuencias

no deseadas con un alto nivel de detalle.

El impacto de la IAGen en las ciencias sociales es igualmente profundo y ambivalente. Por un lado, ofrece herramientas sin precedentes para el análisis de grandes volúmenes de datos cualitativos; por otro, plantea interrogantes sobre la validez de las conclusiones cuando el investigador delega parte de su labor interpretativa en un algoritmo. La integración de sistemas de IA no solo actúa como herramienta auxiliar, sino también como un actante que influye activamente en la construcción de sentido y en la validación del conocimiento social.

Tradicionalmente, la investigación cualitativa ha sido un proceso intensivo en mano de obra humana, dependiente de la interpretación hermenéutica de entrevistas y grupos focales. La integración de LLMs en herramientas como Social Verbatim permite realizar codificaciones inductivas y deductivas a gran escala, identificando temas y patrones con una consistencia que desafía las limitaciones de las capacidades humanas, ya sea por la fatiga o por sesgos cognitivos iniciales. Se ha observado que el razonamiento por cadena de pensamiento en modelos avanzados puede igualar el rendimiento humano en tareas interpretativas complejas, lo que permite a los investigadores manejar cadenas de texto antes inabordables.

En la investigación cuantitativa orientada a las ciencias sociales, la IAGen asiste en la generación y depuración de código en lenguajes como R, Python, SPSS o STATA, facilita la limpieza de datos y ayuda en la interpretación de resultados estadísticos complejos. Esta democratización de las capacidades técnicas permite que investigadores de diversas disciplinas sociales realicen análisis de datos sofisticados sin necesidad de una formación previa exhaustiva en programación.

En conclusión, uno de los principios centrales de la gobernanza de la IA es que los ciudadanos tienen derecho a comprender cómo se toman las decisiones administrativas que los afectan. No obstante, la naturaleza probabilística de los LLMs dificulta la trazabilidad exacta de cómo el sistema llega a una respuesta específica ante

un prompt determinado. Para mitigar este riesgo, las normativas emergentes exigen transparencia algorítmica y la implementación de mecanismos de explicabilidad que permitan a los funcionarios públicos auditar y, si es necesario, anular las recomendaciones del sistema. El concepto de supervisión humana significativa y competente se vuelve central: el agente público debe contar con la formación técnica y ética para intervenir de forma crítica en el uso de la IA, manteniendo la conciencia del sesgo de automatización o del exceso de confianza en el sistema.

Bibliografía

Alawida, M., Mejri, S., Mehmood, A., Chikhaoui, B. e Isaac Abiodun, O. (2023). Un estudio exhaustivo de ChatGPT: Avances, limitaciones y consideraciones éticas en el procesamiento del lenguaje natural y la ciberseguridad. *Información* , 14 (8), 462. <https://doi.org/10.3390/info14080462>

Ato Alvarado, M.E. (2021). El lenguaje claro y la transparencia de las decisiones judiciales. *Revista Oficial Del Poder Judicial*, 13(16), 61-76. <https://doi.org/10.35292/ropj.v13i16.450>

Bailón Sánchez, L.E., & Pico Bazurto, S.P. (2025). Impacto de la inteligencia artificial en el marketing digital: Análisis de tendencias y percepción empresarial. *Código Científico Revista De Investigación*, 6(E1), 748-767. <https://doi.org/10.55813/gaea/ccri/v6/nE1/716>

Bakker, T. (2019). *Social Thinking and Interpersonal Behavior*. Tucumán: Bibliotex

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., Ferrero, E., Agapow, P. M., Zietz, M., Hoffman, M. M., Xie, W., Rosen, G. L., Lengerich, B. J., Israeli, J., Lanchantin, J., Woloszynek, S., Carpenter, A. E., Shrikumar, A., Xu, J., Cofer, E.M., & Greene, C.S. (2018). Opportunities and obstacles for deep learning in biology and medicine. *Journal of the Royal Society, Interface*, 15(141), 20170387. <https://doi.org/10.1098/rsif.2017.0387>

Cina, E., Elbasi, E., Elmazi, G. y AlArnaout, Z. (2025). El rol de la IA en la modelización predictiva para el desarrollo urbano sostenible: desafíos y oportunidades. *Sustainability* , 17 (11), 5148. <https://doi.org/10.3390/su17115148>

Corvalán, J.G. (2018). Inteligencia artificial: retos, desafíos y oportunidades - Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la Justicia. *Revista De Investigações Constitucionais*, 5(1), 295–316. <https://doi.org/10.5380/rinc.v5i1.55334>

Díaz Díaz, C.V., Robles Mendoza, A.L., Rocha Sánchez, T.E., & García Castillo, Z. (2025). Construcción de la sexualidad femenina con enfoque de género: una aproximación al delito sexual. *Revista jurídica Peruana Desafíos En Derecho*, 2(1), 28-40. <https://doi.org/10.37711/RJPDD.2025.2.1.5>

Facultad Latinoamericana de Ciencias Sociales. (2020). *Estado de la Educación Secundaria en América Latina y el Caribe. Aportes para una mirada regional*. San José: FLACSO. https://eulacfoundation.org/system/files/digital_library/2023-07/1_informe_regional.pdf

Ginocchi, M., Ponci, F., & Monti, A. (2021). Sensitivity Analysis and Power Systems: Can We Bridge the Gap? A Review and a Guide to Getting Started. *Energies*, 14(24), 8274. <https://doi.org/10.3390/en14248274>

Hauthal, E., Dunkel, A. y Burghardt, D. (2021). Emojis como indicadores contextuales en publicaciones de redes sociales basadas en la ubicación. *ISPRS, Revista Internacional de Geoinformación*, 10 (6), 407. <https://doi.org/10.3390/ijgi10060407>

Leyva Vázquez, M.Y., Batista Hernández, N., Estupiñan Ricardo, J., & Goyes García, J.F. (2023). Aplicación de análisis de sentimientos y enfoque neutrosófico para la comprensión de información textual en la investigación. *Conrado*, 19(94), 294-300

Mamani, V., & Arisaca, W. (2021). Transparencia en la gestión pública y su incidencia en el derecho al acceso a la información. *Gestionar: Revista De Empresa Y Gobierno*, 1(3), 40-51. <https://doi.org/10.35622/j.rg.2021.03.003>

Matas, A. (2018). Diseño del formato de escalas tipo Likert: un estado de la cuestión. *Revista electrónica de investigación educativa*, 20(1), 38-47

McCoy, J. (2022). Reflexiones sobre el populismo y la polarización en América Latina y sus consecuencias para la democracia. *Desafíos*, 34(2), 1-19. <https://doi.org/10.12804/revistas.urosario.edu.co/desafios/a.11307>

Norabuena Mendoza, C.H., Huamán Osorio, A.P., & Ramirez Asis, E.H. (2020). Modelo de Ecuaciones Estructurales (Con estimación PLS) basado en calidad de servicio y lealtad del Cliente de las Cajas Rurales Peruanas. *Ciencias Administrativas*, 18, 081. <https://doi.org/10.24215/23143738e081>

OECD. (2025). Gobernar con la inteligencia artificial. *Panorama actual y hoja de ruta en las funciones centrales de gobierno*. París: OECD Publishing. <https://doi.org/10.1787/dc00e56a-es>.

Ortíz Palma, L.A., & Mosquera Angulo, H. (2020). Herramienta para la extracción y análisis de información obtenida de la red social Twitter, como apoyo a los procedimientos: nuevo registro calificado y renovación de registros. *INVENTUM*, 15(28), 51-63. <https://doi.org/10.26620/uniminuto.inventum.15.28.2020.51-63>

Peña Carlos, J. S. (2022). Inteligencia artificial para la seguridad jurídica. Superando el problema de la cognoscibilidad del derecho. *Revista Oficial Del Poder Judicial*, 14(17), 55-117. <https://doi.org/10.35292/ropj.v14i17.568>

Pepe-Oliva, R. (2025). Inteligencia Artificial y periodismo: entre la innovación y el desafío ético. *Cuadernos del Audiovisual del Consejo Audiovisual de Andalucía*, (14). <https://dx.doi.org/10.62269/cavcaa.XX>

Pinto da Costa, J.F., y Cabral, M. (2022). Métodos estadísticos con aplicaciones en minería de datos: Una revisión de los trabajos más recientes. *Matemáticas* , 10 (6), 993. <https://doi.org/10.3390/math10060993>

Reyes Palau, N.C., & Rodríguez Caballero, G.A. (2025). Integración de Inteligencia artificial (IA) en la didáctica de las Ciencias Sociales para la formación docente. *European Public & Social Innovation Review*, 10, 1–12. <https://doi.org/10.31637/epsir-2025-2409>

Rodríguez-Martínez, E.A., Flores-Fuentes, W., Achakir, F., Sergiyenko, O. y Murrieta-Rico, F.N. (2025). Navegación y percepción basadas en visión para robots autónomos: Sensores, SLAM, estrategias de control y aplicaciones multidominio: Una revisión. *Eng* , 6 (7), 153. <https://doi.org/10.3390/eng6070153>

Romero Llerena, M.A., & Pandia Yañez, E.J. (2025). La inteligencia artificial en la salud pública: Mejorando la atención médica y previniendo enfermedades. *Aula Virtual*, 6(13), e493. <https://doi.org/10.5281/zenodo.17247148>

Sánchez Díaz, M.F. (2024). Inteligencia artificial generativa y los retos en la protección de los datos personales. *Estudios en derecho a la información*, (18), 179-205. <https://doi.org/10.22201/ijj.25940082e.2024.18.18852>

Shehzadi, N., Ferzund, J., Fatima, R. & Riaz, A. (2025). Análisis automático de complejidad de diagramas de clases UML mediante técnicas de preguntas y respuestas visuales (VQA). *Software*, 4 (4), 22. <https://doi.org/10.3390/software4040022>

Thanasas, G., Kampiotis, G. y Halkiopoulos, C. (2026). Transformando la contabilidad digital: Big Data, IoT y tecnologías de la Industria 4.0: Un estudio exhaustivo. *Journal of Risk and Financial Management* , 19 (1), 92. <https://doi.org/10.3390/jrfm19010092>

Ushiñahua Serrano, M.F., & Contreras Rivera, R.J. (2026). Revisión sistemática sobre inteligencia artificial generativa en creación de contenidos mediáticos para gestión pública y satisfacción de usuarios. *Revista InveCom*, 6(2), e602138. <https://doi.org/10.5281/zenodo.17664203>

Yasaca Pilco, N.E., Alcívar Sosa, G.E., Pérez Loor, J.A., Malavé Beltran, L.A., & Almeida Marin, M.F. (2025). El Impacto de las Redes Sociales en la Evolución del Lenguaje. *Ciencia Y Reflexión*, 4(3), 31–46. <https://doi.org/10.70747/cr.v4i3.396>

De esta edición de *“Inteligencia artificial generativa aplicada a las ciencias sociales y gestión pública”*, se terminó de editar en la ciudad de Colonia del Sacramento en la República Oriental del Uruguay el 25 de febrero de 2026

INTELIGENCIA ARTIFICIAL GENERATIVA APLICADA A LAS CIENCIAS SOCIALES Y GESTIÓN PÚBLICA

ISBN: 978-9915-698-68-7



Escrito por:

Frida Mereyda López Córdova
Rusvelth Paima Paredes
Daniel Alberto Oswaldo Valenzuela Narváez
Timoteo Solano Armas
Carlos Máximo Gonzáles Añorga
Viviana Inés Vellón Flores de Solano