GZ

GIZARTERATUZ
monografikoak

BEHATUZ
OBSERVATORIO VASCO
DE SERVICIOS SOCIALES



Inteligencia artificial y servicios sociales

2026 | NOVIEMBRE

Inteligencia artificial y servicios sociales

La inteligencia artificial (IA) se está consolidando como una herramienta poderosa que está transformando numerosos sectores, y el ámbito de la intervención social no es una excepción. Su capacidad para procesar grandes volúmenes de datos, identificar patrones y aprender de la experiencia abre nuevas posibilidades para mejorar los servicios sociales bajo los principios de universalidad, equidad, adecuación y eficiencia que los sustentan.

La mayoría de los estudios sobre la aplicación de la inteligencia artificial en los servicios sociales se enfocan en optimizar la organización y la eficiencia de los procesos, con el propósito, entre otros, de hacer frente a la creciente sobrecarga que afrontan los y las profesionales del sector (Vallejo-Andrada *et al.,* 2024). En esta línea, la digitalización de los procesos se plantea como una vía para transformar y optimizar las áreas de gestión, prevención, detección e intervención, con el fin de mejorar tanto la eficiencia como la efectividad en la atención a las personas usuarias. La incorporación de sistemas inteligentes puede agilizar trámites, reforzar la coordinación interna y favorecer una asignación más estratégica de los recursos disponibles.

Asimismo, las capacidades analíticas de la inteligencia artificial proporcionan un apoyo significativo en la evaluación de necesidades, la detección temprana de situaciones de riesgo y la implementación de estrategias preventivas más precisas. A ello se suma el potencial de la IA para asistir y acompañar a los equipos profesionales mediante sistemas de apoyo a la toma de decisiones. En conjunto, estos avances contribuyen a configurar un modelo de servicios sociales más ágil, proactivo y centrado en las necesidades de las personas.

Ética y gobernanza pública

La inteligencia artificial puede convertirse en un gran apoyo para el trabajo de los y las profesionales de los servicios sociales, pero en ningún caso está llamada a sustituirlos. Su valor radica en complementar su labor, ampliar capacidades y ofrecer herramientas que permitan intervenir antes, mejor y de forma más equitativa. La tecnología, bien utilizada, puede reforzar el acompañamiento humano, no desplazarlo.

No obstante, toda innovación exige responsabilidad. La literatura científica subraya que la inteligencia artificial no es neutral y que, sin una gobernanza adecuada, puede reproducir o incluso intensificar sesgos existentes. Hay ejemplos en los que sistemas automatizados más diseñados han perjudicado a las mismas personas a las que debían proteger, generando decisiones injustas o distribuciones desequilibradas de recursos.

A ello se suma la posibilidad de deshumanizar procesos o perpetuar patrones de exclusión ya presentes en la sociedad. Por esta razón, es imprescindible que las tecnologías digitales no solo contribuyan a optimizar la gestión, sino que respeten y promuevan principios éticos de justicia y equidad. La incorporación de la IA debe garantizar intervenciones adecuadas, beneficiosas y centradas en las personas, preservando siempre su dignidad y derechos.

Desafíos técnicos y prácticos

Más allá de su potencial, la incorporación de la IA en los servicios sociales implica afrontar diversos desafíos técnicos y prácticos. Para empezar, solo pueden desarrollarse adecuadamente si se apoyan en sistemas de información sólidos y bien estructurados (Minguijón y Serrano-Martínez, 2022). Sin embargo, este es uno de los puntos débiles del sector, caracterizado por la fragmentación entre administraciones y por la falta de integración de datos, lo que dificulta la creación de modelos fiables y útiles. De ahí que el reto de la interoperabilidad siga siendo una prioridad ineludible.

Aun así, cuanto las bases de datos son consistentes, la IA ofrece un valor añadido importante: la posibilidad de impulsar intervenciones más personalizadas. A partir de la información ya disponibles en los sistemas, los algoritmos pueden contribuir a elaborar diagnósticos más amplios y precisos, así como planes de apoyo más ajustados a las necesidades, aspiraciones y capacidades de cada persona. Esto permite avanzar hacia una atención verdaderamente personalizada y centrada en la persona.

No obstante, persisten problemas relevantes. Muchas soluciones tecnológicas no se adaptan bien a determinados entornos o realidades sociales, ya sea por limitaciones de los algoritmos, por dificultades para reconocer situaciones complejas o por la rigidez de algunos sistemas (Vallejo-Andrada *et al.,* 2024). Es imprescindible, por tanto, diseñar modelos más flexibles y validar su funcionamiento en distintos contextos geográficos, institucionales y demográficos para asegurar que sean aplicables y efectivos allí donde se utilicen. También existe el riesgo de incrementar la brecha digital, especialmente si los equipos profesionales no cuentan con la formación suficiente para comprender, usar e interpretar estas herramientas.

Por otro lado, la falta de integración de distintos tipos de datos y la opacidad de ciertos modelos de IA complican su interpretación y dificultan su validación. Cuando no se conoce con claridad cómo se generan las recomendaciones o decisiones automatizadas, disminuye la confianza en estas herramientas y, por ende, su adopción. Por ello, avanzar hacia sistemas más transparentes, auditables y comprensibles resulta esencial para que la inteligencia artificial pueda desplegar todo su potencial en los servicios sociales, sin comprometer su calidad.

Compromisos para una integración adecuada

Tal y como se destacó en la apertura institucional del encuentro Behatuz, dedicado específicamente a esta temática, el Gobierno Vasco ha definido una serie de compromisos para garantizar que la incorporación de la inteligencia artificial en los servicios sociales se lleve a cabo de manera segura, útil y justa:

1. **La persona siempre en el centro**. La IA debe funcionar como un apoyo: acompaña, orienta y agiliza procesos, pero no decide. En ningún caso puede sustituir la valoración profesional ni la autonomía de las personas usuarias en la toma de decisiones sobre su propio proyecto vital.
2. **Ética y rendición de cuentas**. Todo proyecto piloto o servicio que incorpore IA deberá incluir una evaluación de impacto, un análisis de sesgos potenciales, mecanismos que garanticen la explicabilidad y canales accesibles para presentar reclamaciones o solicitar revisiones.
3. **Datos con garantías y sentido**. Se avanzará hacia sistemas interoperables y hacia modelos de plataformas federadas que permitan conectar información sin necesidad de centralizarla, aprovechando así su valor sin comprometer la privacidad ni la protección de datos.
4. **Capacidad pública y talento interdisciplinar**. Es necesario reforzar los equipos incorporando perfiles diversos —profesionales de lo social, especialistas tecnológicos, juristas, economistas, entre otros— y garantizar la participación de las entidades del tercer sector. Sin capacidades propias, no es posible ejercer una verdadera soberanía que permita definir el "para qué" antes que negociar el "cómo".
5. **Evidencia y escalabilidad**. Los proyectos piloto deberán probarse con rigor, medir sus resultados —incluida la experiencias y percepción de las personas usuarias— y escalar únicamente aquellos que demuestren mejorar la calidad, la equidad y la eficiencia del sistema.

REFERENCIAS

Melgosa, N. **Encuentro Behatuz sobre inteligencia artificial y servicios sociales: apertura institucional.** Donostia-San Sebastián, 08-10-2025.

Minguijón, J. y Serrano-Martínez, C. **La Inteligencia Artificial en los Servicios Sociales: estado de la cuestión y posibles desarrollos futuros.** *Cuadernos de Trabajo Social,* vol. 35, n.º 2, 2022, págs. 319-329.

Vallejo-Andrada, A., *et al.* **Aportaciones y limitaciones de la incorporación de la inteligencia artificial a los servicios sociales: una revisión sistematizada.** *Methaodos. Revista de Ciencias Sociales,* vol. 12, n.º 2, 2024, m241202a04.

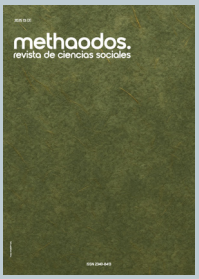
EN DETALLE



Implementación de la inteligencia artificial en las políticas sociales



Los sistemas autonómicos de servicios sociales ante la innovación tecnológica



Ventajas y retos de la inteligencia artificial en los servicios sociales

OCDE
Modernising access to social protection. Strategies, technologies and data advances in OECD countries. París, OECD Publishing, 2024, 131 págs. Ref. 594226.

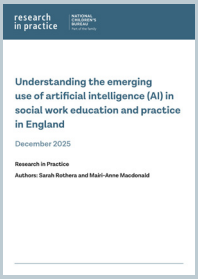
Este informe expone cómo los Gobiernos de los países de la OCDE están aplicando tecnologías avanzadas y el uso de macrodatos para mejorar la eficiencia y cobertura de los programas sociales. A pesar de ello, están avanzando con cautela en la implementación de la inteligencia artificial, debido a los riesgos que implica. Esta tecnología se usa principalmente para apoyar a las personas usuarias, automatizar procesos administrativos y para detectar el fraude. Sin embargo, presenta un gran potencial para el análisis predictivo de demandas y crisis, identificar mejor a las personas beneficiarias o para optimizar la intervención temprana. En este contexto, los Gobiernos están llevando a cabo pruebas a pequeña escala. Un desafío importante es el alto grado de escepticismo público, ya que solo el 40 % de las personas encuestadas en los 27 países de la OCDE creen que el uso de la inteligencia artificial en la aprobación de solicitudes sociales es beneficioso para las personas usuarias. Por ello, este informe subraya la necesidad de un mayor compromiso público para fomentar la transparencia y generar confianza en el uso adecuado y sostenible de esta herramienta para la protección social.

Minguijón, J. *et al.*
La innovación tecnológica en los servicios sociales públicos españoles. *Reforma y Democracia,* vol. 86, 2023, págs. 229-256. Ref. 607138.

Este trabajo es, probablemente, el primer análisis cualitativo que se centra en la percepción de los estamentos directivos de los servicios sociales públicos españoles respecto a su desarrollo tecnológico. Las personas entrevistadas — responsables de ordenación, planificación o innovación de los servicios sociales de cada una de las diecisiete comunidades autónomas— consideran la entrada en la sociedad digital como dificultosa y transmiten el temor que el cambio tecnológico suscita en sus organizaciones. Pese a ello, se encontró una relación positiva entre la percepción del nivel de desarrollo tecnológico de un sistema y la existencia de buenas infraestructuras y apoyo técnico: así, las organizaciones que tienen una mejor percepción del desarrollo tecnológico son también las que demuestran una mayor voluntad política para traspasar nuevas fronteras tecnológicas, como la inteligencia artificial. Esto sugiere que las inversiones tecnológicas previas colocan a los sistemas en una mejor posición de partida para enfrentar los retos futuros. Sin embargo, el análisis también evidencia un desarrollo tecnológico desigual entre las distintas comunidades autónomas, lo que implica una desigual respuesta a las necesidades de la ciudadanía.

Vallejo, A. *et al.*
Aportaciones y limitaciones de la incorporación de la inteligencia artificial a los servicios sociales: una revisión sistematizada. *Methaodos. Revista de Ciencias Sociales,* vol. 12, n.º 2, 2024, 16 págs. Ref. 607137.

El creciente interés científico en torno a la relación entre los procesos automatizados derivados de la inteligencia artificial y los servicios sociales ha impulsado una incipiente producción de conocimiento. Esta revisión sistemática de la literatura publicada entre 2009 y 2024 analiza dicha intersección. En primer lugar, la inteligencia artificial puede optimizar la gestión administrativa mediante la automatización robótica de procesos y modelos matemáticos para mejorar la asignación de recursos. En segundo lugar, facilita la detección temprana y la prevención. Un tercer enfoque se centra en el desarrollo de capacidades y la reducción del aislamiento social a través de *chatbots* o "gemelos digitales". Finalmente, la inteligencia artificial ofrece apoyo directo a las y los profesionales, agilizando el análisis que realizan en casos complejos. A pesar de estos beneficios, su implementación enfrenta resistencias y riesgos. Persisten dilemas éticos por el temor a la sustitución humana, la brecha digital o la pérdida de privacidad. Además, existen problemas técnicos de implementación relacionados con la falta de datos de calidad y la complejidad de los modelos. Para que sea una herramienta exitosa, el estudio concluye que debe superar esos desafíos, desarrollando mediante un diseño participativo modelos transparentes y explicables.



Inteligencia artificial y trabajo social en Inglaterra



La inteligencia artificial en el ecosistema profesional del trabajo social



Algoritmos que discriminan

Rothera, S. y Macdonald, M.
Understanding the emerging use of artificial intelligence in social work education and practice in England: research report.
Londres, Research in Practice, 2025, 73 págs.
Ref. 607380.

El presente informe, encargado por Social Work England, combina una encuesta a 203 participantes, grupos focales y entrevistas con personas expertas, para cartografiar el uso emergente de la inteligencia artificial en el trabajo social en Inglaterra. El hallazgo más relevante es que la inteligencia artificial generativa —aquella capaz de producir texto, imágenes o resúmenes a partir de instrucciones en lenguaje natural— constituye el tipo de inteligencia artificial más extendido. Su adopción se concentra en tareas administrativas: transcripción de visitas, redacción de notas de caso, gestión de correo y minutas de reuniones. El potencial liberador es innegable; sin embargo, su adopción es profundamente desigual: algunas entidades la han implementado con marcos de gobernanza; otras la toleran sin orientación alguna, y un 24 % del personal reconoce usarla sin autorización explícita. El informe es especialmente valioso en su análisis de los riesgos. La protección de datos emerge como la preocupación más acuciante. A ello se suman los riesgos de sesgo algorítmico, la pérdida de la voz propia en los registros de casos, la erosión del pensamiento crítico y los problemas en torno al consentimiento informado. El documento concluye con recomendaciones para el uso ético y responsable.

Loy, A. y Bantry-White, E.
Evaluating AI in social programs: reframing complex intervention as socio-technical intervention. *Journal of Evidence-Based Social Work,* vol. 23, n.º 1, 2026, págs. 94-122.
Ref. 607383.

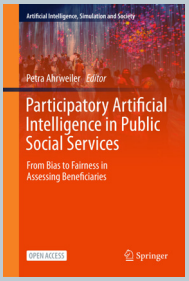
¿Puede la inteligencia artificial integrarse éticamente en los servicios sociales sin transformar —o distorsionar— las relaciones profesionales y las estructuras de poder que los sostienen? Esta pregunta vertebra un artículo conceptual de notable ambición teórica, que propone un reencuadre metodológico para evaluar intervenciones que incorporan sistemas de inteligencia artificial en el ámbito de los servicios sociales. Su argumento central es que los marcos de evaluación existentes resultan insuficientes frente a la inteligencia artificial. Las autoras sostienen que esta nunca es neutral: actúa como agente co-constituyente dentro del ecosistema profesional, reconfigurando roles, relaciones institucionales y dinámicas de poder. A partir del paradigma sociotécnico —enfoque que estudia la mutua configuración entre sistemas humanos y tecnológicos—, se identifican seis dimensiones procedimentales para enriquecer el marco Medical Research Council/National Institute for Health and Care Research (MRC/NIHR). Para ilustrar la aplicabilidad del modelo, se desarrolla un caso práctico centrado en la implantación de un *chatbot* de apoyo psicosocial en una organización juvenil. A pesar de sus limitaciones teóricas, la propuesta de tratar la resistencia de las personas usuarias no como un fracaso, sino como información valiosa para la innovación, resulta especialmente sugerente para la práctica profesional.

Bias in algorithms: artificial intelligence and discrimination. Viena, European Union Agency for Fundamental Rights, 2022, 106 págs. Ref. 607381.

Partiendo de dos estudios de caso, la Agencia de los Derechos Fundamentales de la Unión Europea elabora en este informe un retrato sólido y empíricamente respaldado de los mecanismos mediante los cuales el sesgo algorítmico se traduce en discriminación. El primer estudio simula algoritmos de policía predictiva —sistemas que determinan dónde desplegar patrullas policiales en función de datos históricos de criminalidad— para analizar situaciones en que las predicciones de un sistema condicionan los datos con los que ese mismo sistema se actualiza, amplificando sesgos de forma acumulativa. Los resultados son perturbadores: incluso cuando las tasas de criminalidad reales son uniformes, los modelos de aprendizaje automático concentran progresivamente todos los recursos policiales en determinados barrios. El segundo caso analiza algoritmos de detección de discurso ofensivo, concluyendo que términos como "musulmán", "judío" o "gay" disparan de forma desproporcionada las predicciones de ofensividad. Este fenómeno se produce porque esos términos aparecen profusamente en los datos de entrenamiento etiquetados como discurso de odio, lo que convierte a los propios marcadores identitarios en indicadores de toxicidad. Si bien el alcance restringido de las simulaciones dificulta la extrapolación de los resultados, la lógica estructural de los sesgos identificados sigue siendo válida.



¿Hasta qué punto protegen los algoritmos los derechos sociales?



¿Quién merece este apoyo? Pregunta a la inteligencia artificial



Inteligencia artificial en la gestión de ayudas económicas

Adán, R. *et al.*
Radar de algoritmos y procesos de decisión automatizada para el acceso a los derechos sociales de la ciudadanía. Barcelona, Taula d'Entitats del Tercer Sector Social de Catalunya, 2024, 112 págs. Ref. 593647.

La Taula d'Entitats del Tercer Sector Social de Catalunya editó en un solo volumen dos documentos complementarios: el primero presenta la inteligencia artificial con un enfoque divulgativo —historia, gobernanza y aplicación en el sector público—, y el segundo recopila sistemáticamente sistemas algorítmicos y procesos de decisión automatizada en el ámbito social de las administraciones públicas catalanas. La mayoría de los sistemas examinados están orientados a la eficiencia interna, y no a la toma de decisiones que afectan directamente a la ciudadanía, aunque el sistema de evaluación de riesgo RisCanvi —aplicado a la población reclusa desde 2008— muestra indicios de discriminación racial y socioeconómica que el propio informe documenta sin desarrollar con la profundidad que el caso requeriría. Aquí reside la principal limitación de este este trabajo: el "radar" prioriza la catalogación sistemática sobre el análisis crítico. La ausencia de una perspectiva externa e independiente en la valoración de riesgos, unida al escaso desarrollo de las implicaciones éticas de sistemas como RisCanvi, debilita la utilidad del repositorio como herramienta de control público. Con todo, el informe cumple su propósito declarado: ofrecer al tercer sector una cartografía inicial del ecosistema algorítmico en el ámbito social catalán.

Ahrweiler, P.
Participatory artificial intelligence in public social services: from bias to fairness in assessing beneficiaries. Serie: Artificial Intelligence, Simulation and Society, Cham, Springer Cham, 2025, 312 págs. Ref. 607140.

Este libro es el primer fruto de un proyecto de investigación global e interdisciplinario para acercar el desarrollo de la inteligencia artificial a la sociedad. La iniciativa reúne a personas investigadoras en el ámbito social con especialistas en informática para examinar conjuntamente cómo se utiliza la inteligencia artificial en la selección de personas beneficiarias para servicios de bienestar y cómo se garantiza la equidad en ese proceso. El trabajo señala dos pilares críticos para una inteligencia artificial justa en los servicios sociales: los valores y el contexto. Las decisiones de justicia social son inherentemente decisiones de valor, que la inteligencia artificial debe procesar dentro de marcos dinámicos que distinguen categorías como "merecedor" y "no merecedor". Además, la inteligencia artificial debe ser sensible a la cultura, ya que los criterios de evaluación y el sesgo que discrimina a ciertos grupos son altamente dependientes del contexto cultural específico. El volumen contiene nueve estudios de caso realizados en otros tantos países utilizando métodos participativos con grupos de interés clave, incluyendo grupos vulnerables. Los capítulos exploran las ventajas y los problemas de la IA en la evaluación social.

Using AI to manage minimum income benefits and unemployment assistance: opportunities, risks and possible policy directions. Serie: OECD Artificial Inteligence Papers, n.º 21, París, OECD Publishing, 2024, 54 págs. Ref. 607379.

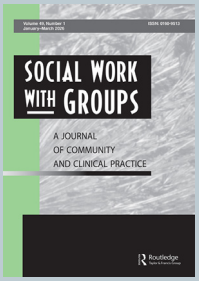
Las elevadas tasas de no utilización de prestaciones sociales *(non take-up)* constituyen uno de los grandes fracasos silenciosos de los sistemas de bienestar modernos. Este documento cartografía con detalle las posibilidades de la inteligencia artificial para la gestión de rentas mínimas y ayudas al desempleo: desde el uso de *chatbots* con procesamiento de lenguaje natural para orientar a las personas solicitantes, pasando por la cumplimentación automática de formularios y la detección de pagos indebidos mediante aprendizaje automático, hasta la evaluación predictiva de elegibilidad basada en criterios normativos predefinidos. Los casos de Australia, Canadá, Dinamarca, Países Bajos y Costa Rica ilustran tanto las implementaciones más avanzadas como los fracasos más documentados. Frente a estas evidencias, el informe propone ocho orientaciones de política pública que incluyen la supervisión humana obligatoria, la gestión estratégica de los umbrales de falsos positivos y negativos, la transparencia algorítmica y la capacitación en alfabetización en inteligencia artificial tanto del personal técnico como de la ciudadanía. La propuesta es sólida, aunque se echa en falta un análisis más detallado sobre los costes reales de implementación y sobre los mecanismos de participación de las personas usuarias en el diseño de estos sistemas.



La inteligencia artificial y evaluación social



La inteligencia artificial, ¿una herramienta frente a la soledad?



Inteligencia artificial e intervención grupal

Besaldi
Exploración, análisis y prospección en la aplicación de la IA en los procesos de evaluación de las políticas públicas. Informe de resultados. Vitoria-Gasteiz, Gobierno Vasco-Eusko Jaurlaritza, 2025, 69 págs. Ref. 603595

Este informe constituye una de las primeras aproximaciones al uso de la inteligencia artificial en la evaluación de las políticas sociales que se edita en nuestro entorno más cercano. Encargado por Besaldi, el órgano responsable de evaluar las políticas de empleo e inclusión del País Vasco, analiza 53 estudios, seleccionados tras una revisión exploratoria de más de 9800 documentos. Los hallazgos se articulan en torno a dos grandes ejes: la aplicación de la inteligencia artificial en la fase de implementación de las políticas, y su incorporación a los procesos evaluativos propiamente dichos. Más allá del catálogo de casos (ref. 603594), destaca la distancia crítica con que se analizan. Frente a narrativas tecno-optimistas frecuentes en este campo, el informe señala con nitidez los riesgos del sesgo algorítmico, la opacidad de algunos modelos, el riesgo moral derivado de la abdicación de responsabilidad profesional ante las decisiones automatizadas y las dificultades de interoperabilidad entre sistemas y administraciones. Se señala también, con pertinencia, que el paradigma de la supervisión humana *(human-in-the-loop),* presentado habitualmente como solución a estos problemas, puede convertirse en un "cascarón procedimental vacío" si no va acompañado de capacitación real y marcos de gobernanza robustos.

Pinazo-Hernandis, S. y Redolat, R.
Artificial intelligence-based technologies to reduce loneliness and improve social connectedness in older people: a systematic review. *The Gerontologist,* vol. 66, n.º 1, 2026, gnaf267, 31 págs. Ref. 605468.

La soledad en la vejez constituye uno de los desafíos de salud pública más urgentes de nuestro siglo. En este contexto, la pregunta que formula este estudio es tan sencilla como relevante: ¿puede la inteligencia artificial contribuir a reducir la soledad y mejorar la conexión social de las personas mayores? La revisión sistemática, elaborada siguiendo la metodología PRISMA, seleccionó 19 estudios publicados entre 2020 y 2024. Los hallazgos más sólidos apuntan a que los asistentes de voz personales y los *chatbots* pueden ofrecer compañía, estimulación cognitiva y apoyo emocional en contextos donde los servicios profesionales resultan insuficientes. Los modelos de aprendizaje automático *(machine learning)* demuestran capacidad para identificar a las personas con mayor probabilidad de participar en intervenciones sociales, lo que permite optimizar recursos y personalizar apoyos. Los resultados sobre robots sociales son, en cambio, más ambiguos: aunque generan actitudes positivas entre las personas usuarias, su eficacia sobre la soledad no queda demostrada de forma concluyente. El artículo es claro respecto a las limitaciones del corpus documental examinado y apenas profundiza en las implicaciones éticas de las tecnologías analizadas, pese a su relevancia.

Wong, J. M. S.
Artificial intelligence-supported group facilitation: emerging potentials, cautious approaches, and ethical considerations in social work practice. *Social Work with Groups,* vol. 49, n.º 1, 2026, págs. 77-90. Ref. 607382.

https://www.behatuz.eus/es/biblioteca-virtual/biblioteca/Record/607382
Frente al optimismo de buena parte de la literatura sobre tecnología y servicios sociales, este artículo adopta una postura deliberadamente cautelosa: reconoce el potencial teórico de la inteligencia artificial en la facilitación grupal, pero advierte con insistencia de que dicho potencial carece aún de respaldo empírico en entornos grupales presenciales. Se trata de una revisión teórica que apela a tres marcos conceptuales clásicos del trabajo social (teoría de sistemas, perspectiva basada en fortalezas y competencia cultural) para estructurar tanto las oportunidades como los riesgos de integrar sistemas de inteligencia artificial en el trabajo con grupos. El artículo identifica cuatro ámbitos donde esta herramienta podría aportar valor. Sin embargo, la mayor parte de la evidencia disponible procede de terapias individuales o de entornos digitales, lo que hace incierta su transferibilidad a la dinámica viva del grupo. El apartado sobre ética resulta lo más sólido del trabajo: la propuesta de una hoja de ruta para la integración ética, articulada en torno a estudios piloto, participación de las partes implicadas, acceso equitativo y evaluación continua, aporta utilidad práctica, aunque su nivel de concreción es limitado.

¿Cómo garantizar un uso responsable de la inteligencia artificial?

González, F. *et al.*
Uso responsable de la IA para las políticas públicas: manual de ciencia de datos. Nueva York, Banco Interamericano de Desarrollo, 2020, 102 págs. Ref. 607473.



El debate sobre la ética de la inteligencia artificial lleva años generando declaraciones de principios que, sin embargo, rara vez descienden al plano operativo. Este manual colma precisamente ese vacío: ofrece a los equipos técnicos que trabajan en el diseño e implementación de sistemas de aprendizaje automático *(machine learning)* un conjunto estructurado de herramientas, listas de verificación y cuadernillos de trabajo replicables para detectar y mitigar los riesgos más comunes a lo largo de todo el ciclo de vida de un sistema de inteligencia artificial orientado a la toma de decisiones de política pública.
La propuesta se organiza en torno a cinco dimensiones: conceptualización y diseño, recolección y procesamiento de datos, desarrollo y validación del modelo, uso y monitoreo, y rendición de cuentas. Esta última actúa como eje transversal, no como etapa final. El texto insiste en una premisa metodológica fundamental: la inteligencia artificial no sustituye a la política pública, sino que la asiste. Destaca el riguroso tratamiento del sesgo algorítmico y el apartado sobre equidad algorítmica. Como limitación, cabe señalar que el asunto de la participación ciudadana y la consulta a las comunidades afectadas está prácticamente ausente del documento.

Ética y normativa aplicada a la inteligencia artificial

Lillo, A.
El Model PIO (principis, indicadors i observables) sobre obligacions i drets: un enfocament d'autoavaluació per al compliment de la regulació en dades i sistemes d'intel·ligència artificial en el marc de la Unió Europea. Girona, Universitat de Girona, 2025, 228 págs. Ref. 607474.



En 2025, el Observatori d'Ètica en Intel·ligència Artificial de Catalunya presentó la versión actualizada de su herramienta de autoevaluación para organizaciones que diseñan, despliegan o contratan sistemas de inteligencia artificial. El modelo propuesto integra el nuevo marco normativo de la Unión Europea (Reglamento 2024/1689) con los principios éticos reconocidos internacionalmente y los instrumentos de derechos humanos aplicables. El núcleo de la herramienta es un cuestionario estructurado en torno a siete principios éticos —transparencia y explicabilidad, justicia y equidad, seguridad y no maleficencia, responsabilidad y rendición de cuentas, privacidad, autonomía y sostenibilidad— que vertebran la evaluación. Cada pregunta incorpora una doble fundamentación: por un lado, las fuentes académicas e institucionales que la inspiran; por otro, los artículos concretos de la normativa complementaria que obligan a o recomiendan su cumplimiento. Ello permite que cada organización no solo conozca qué debe hacer, sino por qué y con qué base normativa. La herramienta distingue, además, entre sistemas de riesgo alto, limitado y mínimo, adaptando en consecuencia la carga evaluativa. La guía contractual que cierra el cuestionario principal ofrece cláusulas-tipo organizadas por los mismos siete principios.

¿Cómo poner en marcha un 'chatbot' con inteligencia artificial?

Administració Oberta de Catalunya
Guía de buenas prácticas para la implantación de un chatbot con IA generativa.

Los *chatbots* son una de las aplicaciones de la inteligencia artificial que más rápida acogida están teniendo entre los servicios de atención a la ciudadanía, incluidos los servicios sociales. La finalidad de estas herramientas es proporcionar información actualizada, ajustada a las necesidades individuales y en el mismo momento en que se solicita. Las entidades interesadas en poner en marcha un *chatbot* con inteligencia artificial generativa encontrarán útil esta página web, basada en la propia experiencia en este terreno de la Administració Oberta de Catalunya. Se trata de una guía breve, que abarca todas las fases del proceso, "desde la selección de la tecnología adecuada hasta su optimización y mantenimiento". Como criterios de diseño, tiene en cuenta "la compatibilidad con los sistemas existentes, el cumplimiento normativo y la calidad de las respuestas", así como la flexibilidad, pues se pretende que los dispositivos creados no solo satisfagan las necesidades presentes, sino que "también puedan evolucionar y adaptarse a las futuras exigencias". Entre los mecanismos de protección de datos para este tipo de *chatbots,* la entidad catalana ha elaborado una ficha de transparencia algorítmica, un documento en permanente actualización que puede tomarse como referencia en proyectos similares.



wSocial, una herramienta con IA para mejorar la intervención en los servicios sociales

wSocial es una herramienta digital concebida como un historial social único para las áreas básicas de servicios sociales, orientada a transformar la manera en que se gestiona y comparte la información relativa a la ciudadanía. Impulsada por el Departamento de Derechos Sociales e Inclusión de la Generalitat de Cataluña, esta solución se enmarca en un proceso más amplio de modernización tecnológica que busca ofrecer una atención más eficiente, integrada y centrada en las personas.

Uno de los principales problemas que aborda wSocial es la fragmentación de los sistemas de información. Hasta ahora, las distintas áreas básicas trabajaban con herramientas heterogéneas y poco conectadas entre sí, lo que dificultaba el intercambio de datos y obligaba, en muchos casos, a que la propia ciudadanía tuviera que aportar reiteradamente la misma documentación. Frente a ello, wSocial propone un sistema unificado que conecta y estandariza la información, proporcionando una visión global de cada persona usuaria.

La herramienta incorpora un visor de datos que centraliza la información relevante y permite a los y las profesionales disponer de una perspectiva integral que incluye datos sociales y, de forma progresiva, información procedente de otros ámbitos como salud u otros sistemas administrativos. Esta integración facilita la planificación de intervenciones, mejora la evaluación de los servicios y refuerza la transparencia y el seguimiento de cada caso. Además, la herramienta permite recoger el contenido de las entrevistas en lenguaje natural y ofrece transcripción simultánea y funciones de traducción, lo que mejora la comunicación con la ciudadanía. El objetivo es que se puedan grabar, transcribir y traducir automáticamente las entrevistas para que los y las trabajadoras sociales tengan más tiempo para la escucha activa.

A medida que se introducen los datos, el sistema identifica términos clave mediante un vocabulario controlado y estandarizado, denominado diccionario Intersocial y basado en el estándar internacional SNOMED CT. Gracias a este sistema de codificación y al uso de algoritmos e inteligencia artificial, la herramienta puede detectar indicios de vulnerabilidad o riesgo y sugerir posibles líneas de intervención. De este modo, no sustituye el criterio profesional, pero sí aporta apoyo analítico para la toma de decisiones fundamentadas.

En conjunto, wSocial simplifica trámites, automatiza tareas administrativas y reduce cargas operativas, permitiendo que los equipos profesionales dediquen más tiempo a la atención directa. Más que una aplicación tecnológica, representa un cambio de modelo en la intervención social, orientado a la coordinación, la eficiencia y la mejora de la calidad del servicio público.

REFERENCIAS

Departamento de Derechos Sociales. **Plan de transformación digital: Social herramienta de intervención social y la funcionalidad del visor** [presentación en la jornada Transformación digital para la prestación y gestión de servicios sociales]. Madrid, Generalitat de Cataluña, 25-01-2024.

DIXIT. **Se pone en marcha el Social, una herramienta para la intervención de los servicios sociales con ayuda de inteligencia artificial.** Barcelona, Generalitat de Cataluña, 01-08-2024.

‘Chatbots’ y asistentes virtuales conversacionales no generativos en países europeos

La presencia de *chatsbots* diseñados para simular conversaciones humanas mediante inteligencia artificial —automatizando la interacción con aplicaciones o sitios web—, así como otros asistentes virtuales no generativos, lleva varios años consolidándose en la administración pública. A diferencia de la inteligencia artificial generativa, estos sistemas no crean contenido nuevo, sino que seleccionan respuestas de una base de datos previamente programada en función de la intención detectada en la persona usuaria. Diversos ejemplos de su aplicación pueden encontrarse en Rumanía, Austria, Bélgica y Finlandia.

En Rumanía, el uso de *chatbots* se orienta principalmente a la gestión de recursos humanos y a la mejora de la experiencia del personal. Destaca el asistente virtual IRIS, disponible en cualquier momento a través de la aplicación Employee Self Service. Este sistema reduce la carga de trabajo del departamento de recursos humanos al ofrecer respuestas inmediatas a las consultar internas, proporcionando una experiencia interactiva y contribuyendo, entre otros aspectos, a la identificación de oportunidades de formación. El potencial de esta tecnología radica, además, en su posible aplicación futura a los procesos de selección y contratación de persona.

Austria ha implementado diversos asistentes conversaciones, entre los que destaca Mona, un asistente digital disponible en plataformas oficiales que surgió durante la crisis de la COVID-19 para apoyar al sector empresarial y que posteriormente se integró en distintos portales de gobierno electrónico. Su función principal es responder a preguntas frecuentes sobre trámites y servicios públicos, reduciendo la necesidad de atención directa. Asimismo, el Gobierno austriaco aplica inteligencia artificial en los ámbitos social y sanitario, especialmente en la clasificación automática de documentos. No obstante, la implantación de la IA sigue siendo gradual y basada en proyectos piloto, como el “KAY system”, concebido para la gestión semiautomática de reembolsos de tratamientos médicos.

Por su parte, Bélgica cuenta con ORI, un *chatbot* desarrollado para responder preguntas frecuentes, orientar en la realización de trámites y facilitar el acceso a información oficial. A diferencia del asistente austriaco, ORI ofrece una asistencia más amplia, dirigida a la ciudadanía en general y no a un sector específico como el empresarial.

Finalmente, en Finlandia destaca el *chatbot* Kela-Kelpo/FPAFolke, cuyo propósito principal es asesorar sobre prestaciones y ayudas sociales, así como guiar en los procesos de solicitud. Se trata de un asistente multilingüe, con versiones en finlandés y sueco, lo que permite atender a las personas usuarias en su propio idioma en las distintas regiones del país.

REFERENCIAS

Berta, L. y Frătică-Dragomir, A. **Artificial intelligence in the public sector using chatbot: Romania vs Austria, Belgium, Bulgaria and Czech Republic.** *Internacional Journal of Applied Science and Research,* vol. 8, n.º 3, 2025, págs. 127-138.

Väänänen, N. **The digital transition of social security in Finland. Frontrunner experiencing headwinds?** *Social Insurance. Theory and Practice,* vol. 151, n.º 4, 2021, págs. 71-86.

Teleasistencia predictiva

La Fundación Adinberri, organismo dedicado a la innovación en el ámbito del envejecimiento y dependiente de la Diputación Foral de Gipuzkoa, ha impulsado, junto al Gobierno Vasco, una iniciativa pionera de teleasistencia predictiva. Este proyecto tiene como finalidad reforzar la calidad de vida, la autonomía personal y la seguridad de las personas mayores que desean continuar viviendo en sus propios hogares. Su propósito principal consiste en identificar de manera temprana posibles situaciones de riesgo, como la fragilidad o el inicio de procesos de dependencia, con el fin de pasar desde un modelo de atención reactivo hacia otro sustentado en la prevención, la predicción, la personalización de los cuidados y la participación activa de las personas usuarias, conocido como modelo de atención 4P.

El funcionamiento del sistema se apoya en la integración de tecnologías avanzadas como la inteligencia artificial, el internet de las cosas *(internet of things)* y el análisis de grandes volúmenes de datos *(big data),* lo que permite desarrollar modelos predictivos sólidos y eficaces. En particular, se emplean herramientas de procesamiento del habla capaces de detectar de forma precoz signos de deterioro cognitivo o situaciones relacionadas con la soledad no deseada.

Gracias a esta tecnología, es posible analizar tanto hábitos individuales como dinámicas colectivas, definir patrones de comportamientos habituales y reconocer desviaciones significativas en tiempo real. Por el ejemplo, el sistema puede registrar cuánto tiempo permanece una persona en determinadas estancias del hogar, como la cocina, el salón o el dormitorio, con el objetivo de identificar cambios relevantes en sus rutinas que podrían estar asociados a un deterioro progresivo de su estado de salud. Asimismo, el uso inusual de electrodomésticos o los desplazamientos en horarios poco habituales pueden alertar sobre alteraciones conductuales o fisiológicas. De forma complementaria, se contempla la instalación de sensores ambientales y dispositivos portátiles, como relojes inteligentes, que monitorizan indicadores de salud y comportamiento. Esto permitiría identificar con antelación riesgos de caídas y otros incidentes domésticos que pueden comprometer la autonomía personal.

Este proyecto piloto, con una duración prevista de 18 meses, se ha desarrollado dentro del servicio de teleasistencia betiON del Gobierno Vasco en un centenar de hogares de la comarca de Oarsoaldea (Gipuzkoa). Su carácter replicable permitirá, en caso de obtener resultados satisfactorios, su futura integración en betiON y su extensión progresiva al conjunto de Euskadi. La iniciativa ha sido presentada como buena práctica en el proyecto europeo KORALE, orientado a promover respuestas innovadoras frente a la soledad no deseada.

REFERENCIAS

Fundación Adinberri. **Teleasistencia predictiva**.

Fundación Adinberri (2024). **Adinberri lanza un pionero proyecto para identificar de forma precoz situaciones de riesgo de personas mayores que viven en su hogar**. 12-12-2024.

Diputación Foral de Gipuzkoa (2025). **La Diputación Foral de Gipuzkoa, junto a Adinberri, presenta ante Europa su proyecto pionero de teleasistencia predictiva**. 27-05-2025.

Sistema de analítica predictiva y alertas tempranas para abordar el sinhogarismo

El Consejo Municipal de Maidstone, en el Reino Unido, ha implementado un enfoque innovador y preventivo en la prestación de servicios sociales mediante el uso de análisis de datos predictivos para hacer frente al aumento de la pobreza y el sinhogarismo. En colaboración con las empresas EY y Xantura, se desarrolló la herramienta One View, cuyo objetivo principal es integrar la información dispersa en más de quince bases de datos con el fin de identificar de manera proactiva a personas en riesgo de perder su vivienda. El sistema combina analítica predictiva y técnicas de procesamiento del lenguaje natural para generar alertas tempranas a partir de indicadores como impagos de suministros o incidencias en ayudas de vivienda. Esto permite a los equipos profesionales intervenir antes de que la situación derive en una crisis.

Este modelo aspira a sustituir sistemas tradicionales, más manuales y fragmentados, para una gestión basada en la anticipación. Su principal fortaleza reside en la capacidad de actuar de forma preventiva: únicamente el 0,4 % de los casos de alto riesgo que recibieron intervención temprana acabaron en situación de sinhogarismo, frente al 40 % de aquellos que siguieron los procedimientos habituales. Durante su implementación piloto en el contexto de la pandemia de la COVID-19, One View contribuyó a evitar que cien hogares perdieran su vivienda y a reducir en un 40 % la tasa general de sinhogarismo en el municipio. Además, el programa generó un gran ahorro económico y permitió disminuir significativamente el tiempo dedicado a tareas administrativas.

Para garantizar la viabilidad del proyecto, se ha prestado especial atención a la protección de datos personales. One View anonimiza la información y restringe el acceso a los datos individuales exclusivamente al personal profesional responsable de cada caso. No obstante, su éxito no depende únicamente de la tecnología, sino también de la adaptación del modelo de atención, que implica formar al personal para aprovechar la visión integral que ofrece la herramienta y mejorar la interacción con la población vulnerable.

REFERENCIAS

Maidstone Borough Concil. **The One View Project. Using data to identify households at risk of homelessness.**

Proyecto Paloma: inteligencia artificial para detectar la soledad no deseada

El Ayuntamiento de Madrid llevó a cabo el proyecto piloto Paloma, una iniciativa innovadora orientada a detectar situaciones de soledad no deseada entre personas mayores mediante el uso de tecnologías de procesamiento del lenguaje en tiempo real. Este sistema permite a una inteligencia artificial realizar llamadas telefónicas que simulan de forma realista la voz humana, manteniendo conversaciones fluidas y formulando preguntas específicas destinadas a identificar posibles indicadores de riesgo asociados a la soledad no deseada.

El objetivo principal de esta herramienta es complementar la labor de los equipos profesionales de los servicios sociales, ofreciendo un primer contacto a través de un asistente virtual conversacional dirigido especialmente a personas mayores que viven solas. A partir de la información obtenida durante la conversación, el sistema puede recomendar actividades, recursos o servicios adaptados a las necesidades detectadas, contribuyendo así a mejorar el acompañamiento y la atención preventiva. Las llamadas siguen una metodología estructurada que incluye un saludo inicial y la presentación del servicio, seguido de cinco preguntas orientadas a evaluar el posible riesgo de soledad no deseada. Posteriormente, se ofrece a la persona la posibilidad de recibir una llamada de seguimiento por parte del personal municipal y se facilita información sobre recursos o servicios que puedan resultar de interés, así como la derivación a aquellos cuando procede.

El programa piloto se desarrolló entre noviembre y diciembre de 2023 e incluyó llamadas a 5163 personas mayores de 75 años que vivían solas, con una edad media de 81 años y una presencia mayoritaria de mujeres (70 %). Se logró contactar efectivamente con 2937 personas, completándose el 71 % de las conversaciones iniciadas. Los resultados obtenidos revelaron que una de cada tres personas contactadas manifestó sentirse sola. Asimismo, 646 personas expresaron interés en recibir seguimiento por parte del ayuntamiento. En algunos casos, estas ya formaban parte del programa municipal de acompañamiento a la integración social de personas mayores, mientras que otras fueron derivadas a talleres específicos para el abordaje de la soledad o a programas de voluntariado que ofrecen acompañamiento domiciliario. De este modo, el proyecto Paloma se configura como una herramienta de apoyo para la detección temprana y la intervención preventiva en situaciones de aislamiento social.

REFERENCIAS

Ayuntamiento de Madrid. **Paloma: inteligencia artificial para detectar la soledad no deseada en personas mayores.** 2024.

IA, robótica y analítica avanzada para agilizar la valoración de la discapacidad

Actualmente, la administración pública se enfrenta a una presión creciente para ofrecer servicios sociales de calidad frente a recursos limitados y la complejidad de la normativa y los casos que deben atender. En el contexto de Canarias, esta situación se evidenció de manera particular en el proceso de valoración de la discapacidad, donde la elevada demanda y los procedimientos burocráticos derivaban en acumulación de solicitudes y retrasos significativos en los tiempos de resolución. Para afrontar esta ineficiencia, el Gobierno de Canarias implementó una estrategia integral que combinaba ajustes normativos, reorganización interna y, de manera clave, la incorporación de nuevas tecnologías como la inteligencia artificial, la analítica avanzada y la robótica. El objetivo principal fue automatizar, simplificar y agilizar los trámites, para así liberar al personal de tareas rutinarias y repetitivas que no aportaban eficiencia.

Las herramientas de inteligencia artificial se despliegan mediante "agentes especializados", cada uno centrado en funciones concretas dentro del tratamiento y clasificación de solicitudes. Entre ellos, se incluyen agentes organizadores, procesadores, validadores y generadores, que analizan grandes volúmenes de datos procedentes de distintas fuentes para clasificar y priorizar los expedientes. De este modo, se identifican automáticamente los casos urgentes, se realizan comprobaciones rutinarias de la documentación y se solicitan automáticamente los datos faltantes a los solicitantes. Además, los agentes pueden evaluar aspectos relacionados con los grados de discapacidad y redirigir los expedientes a los equipos profesionales correspondientes para su verificación final. Este enfoque reduce considerablemente la carga administrativa, tanto para la ciudadanía como para el personal de los servicios sociales, y acelera la resolución de los casos.

La analítica avanzada complementa este proceso al transformar el volumen de datos administrativos en conocimiento útil. Mediante técnicas de análisis predictivo y aprendizaje automático, la Administración puede anticipar tendencias de demanda, prever la saturación de servicios y comprender mejor los patrones de las solicitudes. Esta información permite diseñar políticas públicas más efectivas y asignar recursos de manera estratégica, orientándolos hacia donde generarán un mayor impacto social.

Por su parte, la robótica automatiza tareas manuales y repetitivas mediante agentes asistenciales y dispositivos inteligentes, liberando al personal para actividades de mayor valor. Por ejemplo, se han desarrollado sistemas que monitorizan las entradas de solicitudes en el registro electrónico las 24 horas del día y ejecutan acciones iniciales de procesamiento sin intervención humana.

La integración de estas tecnologías representa una apuesta estratégica del Gobierno de Canarias para modernizar los servicios sociales. Combinada con medidas normativas y organizativas, la tecnología permite anticipar necesidades, optimizar recursos y ofrecer una atención más personalizada, equitativa y eficiente, eliminando retrasos y deficiencias que afectaban negativamente a la ciudadanía.

REFERENCIAS

Morín, C. **Canarias utilizará Inteligencia Artificial para reducir la burocracia en discapacidad.** *El Día,* 17-03-2025.



2026 | NOVIEMBRE

GZ

GIZARTERATUZ monografikoak

Contenidos

SIIS SERVICIO DE INFORMACIÓN E INVESTIGACIÓN SOCIAL

Edita

BEHATUZ. OBSERVATORIO VASCO DE SERVICIOS SOCIALES

Diseño, ilustración y maqueta

Concetta Probanza

Licencia

Este documento está bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivadas 4.0 Internacional

BEHATUZ
OBSERVATORIO VASCO
DE SERVICIOS SOCIALES

EUSKO JAURLARITZA
GOBIERNO VASCO

BERDINTASUN, JUSTIZIA
ETA GIZARTE POLITIKETAKO SAILA

DEPARTAMENTO DE IGUALDAD,
JUSTICIA Y POLÍTICAS SOCIALES