

Aplicación de Inteligencia Artificial a servicios de atención ciudadana

Informe Gobe sobre Espacio de Innovación de
Atención Ciudadana

Junio 2026 — Fernando Fernández-Monge¹



Índice

Índice.....	0
Objetivo.....	2
Un servicio público clave en profunda transformación.....	3
El stack tecnológico de la IA en atención ciudadana.....	4
Un mapa cambiante de proveedores.....	5
Los 5 grados de autonomía de los agentes de IA de atención ciudadana.....	6
Guía de decisiones para las Administraciones públicas.....	7
Conclusión.....	9

Objetivo

El objetivo de este *Informe Gobe* sobre innovación en servicios de atención ciudadana es ofrecer una guía para directores de transformación e innovación en la Administración pública. La atención ciudadana es uno de los servicios públicos con mayor impacto en la percepción de calidad de las Administraciones públicas y, al mismo tiempo, uno de los espacios donde la incorporación de IA conversacional y agéntica está produciendo cambios estructurales en tres frentes simultáneos: el grado de autonomía de los sistemas, la madurez de los modelos de voz y la transformación del mapa de proveedores.

El informe identifica las diferentes capas del *stack* tecnológico que componen la categoría de la atención ciudadana y los proveedores asociados a cada una de ellas, describe los diferentes niveles de autonomía en herramientas de IA disponibles en el mercado, y presenta preguntas guía para ayudar a tomar decisiones de compra de innovación tecnológica en servicios de atención ciudadana.

Un servicio público clave en profunda transformación

La atención ciudadana es una función fundamental de muchas Administraciones públicas. Por ejemplo, en 2024, el servicio municipal del Ayuntamiento de Madrid, [Línea Madrid](#), reportó 10,45 millones de servicios, de los cuales 5,54 se prestaron a través de la línea telefónica y 1,22 por vía telemática.² Este volumen refleja la importante dedicación de recursos por parte de muchas Administraciones públicas, y lo convierte en un servicio con gran impacto en la percepción sobre la calidad de los servicios públicos y la satisfacción ciudadana.

Este espacio está experimentando tres transformaciones simultáneas:

1. **Profundización de la autonomía mediante IA.**

Los sistemas ya no son únicamente chatbots de respuesta a preguntas frecuentes, sino agentes capaces de ejecutar acciones (citás, inscripciones, pagos, devoluciones) y de orquestar trámites de extremo a extremo. La frontera de innovación se está desplazando del lenguaje natural a la integración con sistemas de registro y a la trazabilidad de la cadena de acciones.³

2. **Mejora de los modelos de voz.**

El canal telefónico – mayoritario en muchos servicios de atención ciudadana y esencial para colectivos vulnerables – se vuelve viable para la automatización y mejora la experiencia con el servicio, con implicaciones directas sobre la mitigación de la brecha digital y la accesibilidad cognitiva.

3. **Reconfiguración del mapa de proveedores.**

La entrada de plataformas nativas en IA con respaldo financiero excepcional (Sierra, Parloa, Fin, Decagon, PolyAI, ElevenLabs) está erosionando el modelo de negocio de *incumbents* (call centers, CRMs). La adquisición de Fin (Intercom) por Salesforce en junio de 2026, de Cognigy por NICE en septiembre de 2025 y las valoraciones alcanzadas por estas nuevas compañías en 2025–2026 confirman la velocidad de la consolidación. En este mapa cambiante también están surgiendo múltiples empresas que buscan consolidarse a través de la especialización vertical en algunos sectores de la Administración como salud, transporte o servicios municipales.

Estos cambios e innovaciones provocan la necesidad para la Administración pública de entender el sector y las tendencias del mercado para comprender las funcionalidades disponibles y evaluar proveedores. Este conocimiento les permitirá identificar, evaluar y comprar innovaciones para la mejora de la atención ciudadana.

El *stack* tecnológico de la IA en atención ciudadana

Entender las cinco capas del *stack* es condición previa para evaluar proveedores y descomponer la decisión de compra. Cada capa tiene dinámicas competitivas, ritmos de cambio y costes de portabilidad distintos:

Capa 1 — Modelos fundacionales.

Los modelos de lenguaje a gran escala (LLMs) de propósito general (GPT, Claude, Gemini, Mistral, ALIA) proporcionan la comprensión y generación del lenguaje subyacente. Suelen consumirse a través de los proveedores de las capas superiores, y las decisiones más relevantes tienen que ver con la soberanía sobre los modelos, los costes de consumo de *tokens* y la localización de los datos transferidos.

Capa 2 — Integración con sistemas de información interno.

Para funcionar – sobre todo los sistemas de atención ciudadana con capacidades de gestión y resolución de trámites y prestación de servicios – las funcionalidades de atención tienen que conectarse con los sistemas municipales subyacentes: padrón, gestión tributaria, plataformas de administración electrónica, gestión documental, sistema de tickets, y GIS (sistemas de información geográfica). Esta es la capa con el ritmo de cambio más lento y los costes de cambio más elevados.

Capa 3 — Orquestación y resolución.

La compone el *middleware* que combina los modelos fundacionales (capa 1) con la recuperación de información (capa 2), la lógica de flujos de trabajo, los adaptadores de integración, las reglas de escalado y los registros de auditoría (*audit logs*). Aquí es donde se concentra actualmente la mayor parte de la competencia entre proveedores.

Capa 4 — Canales e interacción.

Las interfaces a través de las cuales interactúan los ciudadanos: telefonía de voz, chat web, API de WhatsApp, mensajería de aplicaciones móviles,

mensajes directos en redes sociales y correo electrónico. Los proveedores pueden centrarse en un solo canal o ser multicanal.

Capa 5 — Analítica, calidad e inteligencia.

Analítica de conversaciones, analítica de la "voz del ciudadano", automatización del control de calidad, previsión de la demanda, monitorización de equidad y sesgos, y análisis de causa raíz. En esta capa hay un ecosistema de proveedores especializados, además de los módulos analíticos integrados en otras plataformas o proveedores.

Capa	Ritmo de cambio	Coste de portabilidad
1. Modelos fundacionales	Muy rápido	Bajo si la orquestación es agnóstica al modelo
2. Sistemas de registro	Lento	Muy alto (lock-in estructural) ⁴
3. Orquestación y resolución	Muy rápido	Medio-alto
4. Canales e interacción	Rápido (voz) / medio (resto)	Medio
5. Analítica e inteligencia	Medio	Bajo-medio

Un mapa cambiante de proveedores

Pueden distinguirse cuatro tipologías de actores en el espacio de innovación de la atención ciudadana, con coincidencias parciales y movimientos cruzados entre capas:

1. CRM (Customer Relationship Management).

Controlan los sistemas de registro y operan principalmente en la capa 2. Buscan incorporar agentes de IA en sus *suites* (Salesforce, Microsoft Dynamics, o Berger-Levrault y GTT en el ámbito local).

2. Contact centers.

Coordinan la interacción en tiempo real entre Administración y ciudadanía. Operan principalmente en capa 3 y 4 (Genesys, NICE, Five9).

3. Asistentes digitales y chatbots.

Capa conversacional ciudadano-directa, operan principalmente en capa 4 (1MillionBot, Aunoa, CivitPhone o Tucuvi).

4. Plataformas nativas en IA agéntica.

Diseñadas desde cero en torno a LLM y flujos agénticos multipaso (Sierra, Parloa, Decagon, PolyAI). Operan transversalmente entre capas 3 y 4 y compiten por el segmento más alto del mercado.

La mayoría de compradores públicos en España contratan a varios proveedores que cubren diferentes capas. El enfoque de proveedor único es viable principalmente para ayuntamientos pequeños con necesidades sencillas o para compradores plenamente comprometidos con un ecosistema CRM o *contact center*. Los proveedores consolidados de CRM/Contact Center aumentarán probablemente cuota en orquestación y canales, impulsados por sus inversiones en IA, relaciones existentes y adquisición de empresas emergentes.⁵ Los especialistas pueden mantener su posición donde el conocimiento especializado del sector o el marco regulatorio sean determinantes. Por su parte, los nuevos proveedores como Sierra, Parloa, Decagon, PolyAI y ElevenLabs se están posicionando para captar el segmento alto: grandes administraciones con despliegues multicanal y alto volumen de voz. Su capitalización extraordinaria⁶ los convierte en opciones especializadas más estables a largo plazo, pero su penetración actual en el sector público español es limitada.

Algunas capas, como la de *middleware*, también se desarrolla internamente – o con un desarrollo a medida – por la Administración pública. Esto permite mayor control y flexibilidad, aunque requiere un equipo dedicado. Otra ventaja es que los modelos de orquestación definidos de manera customizada permiten también usar los desarrollos para otros servicios fuera de la atención ciudadana.

Los 5 grados de autonomía de los agentes de IA de atención ciudadana

La taxonomía de cinco niveles de autonomía permite ajustar la ambición tecnológica al perfil de riesgo y capacidades de la Administración:

Nivel 1 – Únicamente informativo.

El sistema responde preguntas a partir de una base de conocimiento y deriva a humanos para cualquier otra cosa. Este es el estado actual de la mayoría de los chatbots de las Administraciones públicas españolas.

Nivel 2 – Acciones en un solo sistema.

El sistema puede ejecutar acciones individuales y de bajo riesgo en un único

sistema de *back-office* (reservar una cita, solicitar un documento, registrar una queja).⁷ Los despliegues más innovadores de 2025 se encuentran principalmente en este nivel. Por ejemplo, el asistente de Línea Madrid en el Ayuntamiento no solo informa, sino que permite o canaliza gestiones como solicitud de autorizaciones para el Distrito Centro, gestión de cita previa, alta y consulta de avisos de la ciudad, y pagos de multas, el IVTM y tasas de gestión de residuos por Bizum.

Nivel 3 — Orquestación entre sistemas.

El sistema puede encadenar acciones a través de múltiples sistemas de *back-office* para completar un trámite de extremo a extremo (verificación de elegibilidad, generación de documentos, presentación, notificación). Esta es la vanguardia de los despliegues actuales, aunque la mayoría se centran en automatización administrativa interna y no en asistentes conversacionales ciudadanos.⁸

Nivel 4 — Resolución autónoma con discrecionalidad.

El sistema puede gestionar casos no rutinarios, incluyendo cierto ejercicio de discrecionalidad administrativa dentro de parámetros acotados, con revisión humana en las excepciones. Esto es en gran medida aspiracional en 2026 y plantea dilemas jurídicos sustantivos.

Nivel 5 — Interacción proactiva.

El sistema anticipa las necesidades del ciudadano e inicia el contacto (recordatorios de renovación, notificaciones de elegibilidad, recuperación de trámites abandonados) de manera proactiva. Es el futuro que algunos anticipan pero que aún no está desplegado.

A día de hoy, los niveles de madurez óptimos (desde un punto de vista de innovación versus riesgo) para la contratación pública se encuentran en los niveles 2 y 3: agentes que guían a los ciudadanos, gestionan la demanda repetitiva de la atención telefónica, reservan citas, inician trámites y escalan de forma segura.

Guía de decisiones para las Administraciones públicas

Las nueve preguntas siguientes sintetizan las dimensiones más relevantes en una evaluación de un proveedor en este espacio:

1. *¿Es un proveedor especializado o un módulo de un CRM o contact center existente?*

Un interlocutor único simplifica la contratación pero a costa de precios más altos y mayor dependencia. Los enfoques multi-proveedor *best-of-breed* aportan flexibilidad y precios competitivos a cambio de mayor complejidad de integración.

2. *¿Cuántos canales cubre? ¿Cuál es la calidad real del canal de voz?*

Los proveedores omnicanal eliminan silos pero pueden sacrificar profundidad. Los especialistas en voz hiperrealista ofrecen calidad superior pero exigen integración telefónica compleja. Desde la perspectiva del ciudadano, los canales no son intercambiables: la voz es esencial para personas mayores, colectivos con necesidades de accesibilidad o situaciones complejas, y constituye un requisito de equidad.

3. *¿Qué soporte ofrece a lenguas cooficiales y accesibilidad?*

Catalán, gallego, euskera, valenciano y aranés pueden representar un requisito funcional, pero la calidad de los modelos de voz varía notablemente por lengua.

4. *¿Cuál es el grado de especialización vertical y regulatoria?*

La especialización vertical profunda (modelos de datos, integraciones y flujos sectoriales propios) ofrece ventajas reales y barreras de entrada elevadas. La especialización regulatoria (certificaciones específicas como MDR Clase IIb en sanidad, en el caso de Tucuvi) es difícil de replicar en plazos cortos y constituye un activo decisivo en sectores regulados.

5. *¿Cuál es el modelo de pago y el CTA real?*

Hay al menos tres modelos: *pago por resolución* (entre 0,50 y 2,00 EUR por interacción resuelta), *suscripción anual fija* y *add-on de IA* sobre CRM/Contact center existente. El coste total de adquisición (CTA) real incluye integraciones (que puede llegar hasta el 60-70% del coste total del primer año), formación, mantenimiento evolutivo y soporte. El precio nominal por interacción raramente refleja el coste total.

6. *¿Cómo se gestiona la trazabilidad de la cadena de acciones?*

Es exigible que el sistema registre de forma auditable el porqué de cada respuesta o trámite automatizado, de modo que ante cualquier reclamación pueda reconstruirse el camino lógico de la decisión.

7. *¿Cómo es el protocolo Human-in-the-Loop?*

La derivación a funcionario debe transferir contexto (no obligar al ciudadano a repetir su problema) y debe existir un protocolo claro que permita al empleado público corregir o validar respuestas del sistema. La derivación es indicador clave de equidad: tasas de derivación significativamente más altas en colectivos vulnerables son una señal de alerta.

8. *¿Cuál es la postura del proveedor sobre soberanía de datos y modelos?*

Conviene comprobar la localización de los datos, las transferencias internacionales, el régimen aplicable a terceros, la portabilidad y la salida, la conformidad con el Esquema Nacional de Seguridad y la opción real de ejecutar sobre modelos abiertos (Mistral, ALIA, Llama) si la criticidad del servicio lo recomienda.

9. *¿La capa de orquestación de procesos, me permite abordar más casos de uso que solamente el de atención ciudadana? ¿La podré reutilizar para flujos internos o para que las aplicaciones corporativas puedan usarla también?*

La posibilidad de reutilizar la orquestación de procesos puede determinar la decisión de si invertir en un desarrollo propio que la Administración pueda escalar horizontalmente (a otros procesos) controlando los costes. Si la orquestación comprada a un tercero se enfoca únicamente a atención ciudadana y no puede transferirse fácilmente, puede perder utilidad si la voluntad es replicarla y conectarla a otros procesos, flujos y aplicaciones.

Conclusión

La atención ciudadana es uno de los servicios públicos con mayor impacto en la percepción de calidad de las Administraciones públicas y, al mismo tiempo, un ámbito donde la incorporación de IA está produciendo cambios profundos. Este documento resume algunas de las tendencias clave en este espacio de innovación, y aporta una serie de preguntas para informar la toma de decisiones de compra, que se resumen en tres principios clave finales:

1. Analizar innovaciones por capas del *stack* y grados de autonomía ayuda a identificar proveedores, priorizar innovación controlando riesgo, y diseñar arquitecturas para mantener opcionalidad evitando sobrecostes.

2. Las decisiones variarán en función de la tipología de la Administración: ayuntamiento pequeño/mediano/grande, administración autonómica y AGE tienen problemas, capacidades, y restricciones distintas. Por ejemplo, mantener arquitecturas multi-proveedor (*best-of-breed*) con cláusulas de portabilidad y salida explícitas puede ser la opción óptima en entidades grandes y con capacidad, pero no en ayuntamientos pequeños o cuando se opte deliberadamente por un ecosistema único.
3. Las decisiones de compra han de incorporar elementos clave en un contexto público como trazabilidad auditable de cada acción automatizada y un protocolo *Human-in-the-Loop* documentado. Además, para los niveles superiores de autonomía, los términos contractuales deben incluir indicadores de resultado ligados a tasa de resolución autónoma, satisfacción ciudadana, equidad por colectivo y derivaciones por motivo.

¹ Gracias a Guillermo Areán, de Attico, por sus comentarios a una versión anterior de este documento. El contenido final es responsabilidad única del autor. Desde Gobe nos encantaría recibir cualquier comentario o sugerencia para mejorar el informe.

² El resto, 3,69 millones, se prestaron a través de servicios presenciales en oficinas de atención ciudadana.

³ Distinguimos entre “acciones” y “decisiones”. Las acciones no implican una decisión discrecional por parte de la Inteligencia Artificial, sino que se limita a aplicar criterios predeterminados. Las decisiones implican un juicio ante una circunstancia no predeterminada, donde puede haber un elemento de ambigüedad y de discrecionalidad. A día de hoy, las herramientas de IA tienen limitaciones jurídicas para tomar decisiones discrecionales, por lo que principalmente se centran en ejecutar acciones de acuerdo a criterios determinados de antemano.

⁴ La complejidad técnica de la integración no siempre es elevada. La mayoría de las entidades locales cuentan con sedes electrónicas maduras que resuelven trámites habituales (como la emisión de certificados de empadronamiento o la cita previa) mediante servicios de *backend*. Desde una perspectiva arquitectónica, incorporar un agente de IA implica simplemente habilitar un nuevo canal de entrada sobre APIs que ya existen, por lo que el coste de desarrollo suele ser bajo. El desafío crítico es, en realidad, la gobernanza de la identidad y la seguridad. Mientras que la sede electrónica descansa sobre sistemas de autenticación robustos (como certificado digital o Cl@ve), el despliegue de agentes de voz o texto plantea una brecha: ¿cómo verificar la identidad del ciudadano de forma automatizada y segura en un canal abierto? ¿o cómo verificar la identidad de los agentes de IA? Al igual que el servicio 060 recurre al cotejo de variables específicas (datos del DNI o del registro de vehículos), los ayuntamientos necesitan definir qué fuentes de datos de confianza pueden explotar en tiempo real para articular un sistema de autenticación alternativo, seguro y plenamente automatizado en sistemas de atención ciudadana con herramientas de IA.

⁵ La adquisición de Intercom y su agente Fin por parte de Salesforce es un claro ejemplo de estas dinámicas.

⁶ Sierra alcanzó una valoración de 15.800 M USD en mayo de 2026, Parloa triplicó valoración hasta 3.000 M USD en enero de 2026, Decagon alcanzó 4.500 M USD tras rondas sucesivas, PolyAI cerró Series D a 750 M USD,

ElevenLabs ascendió a 11.000 M USD y la adquisición de Intercom por Salesforce se cerró en 3.600 M USD en junio de 2026 y la de Cognigy por NICE se cerró en ~955 M USD en septiembre de 2025.

⁷ En estos casos, la IA puede reducir los costes administrativos, permitiendo determinar si la documentación aportada para un trámite es correcta, evitando subsanaciones posteriores al permitir el envío sólo cuando cumple con los requisitos (esto requiere un entrenamiento previo para validar si un documento es un DNI, o si un contrato de compraventa y un recibo telefónico contienen los datos correctos). La IA puede extraer toda esta información de forma automática para el funcionario, lo que reduce notablemente los tiempos de tramitación tanto para la Administración como para el ciudadano.

⁸ Un ejemplo es el uso de robotización e IA en la tramitación y justificación de ayudas y subvenciones por Red.es, incluyendo *Kit Digital* y *Kit Consulting*. Para *Kit Consulting* Red.es [estimó](#) una reducción del 80% en tiempos de justificación de expedientes, de unas 5 horas a 1,5 horas, y de comparaciones masivas entre documentos para detectar duplicidades o plagio.