

# Desafía Madrid

Informe Final del Piloto en Madrid: Tecnología  
GreenView® para la aplicación de inteligencia artificial en  
el mantenimiento de zonas verdes

## Acerca de este documento

El mantenimiento de las grandes zonas verdes de Madrid, un servicio con un coste anual superior a los 250 millones de euros, se ha basado históricamente en un modelo reactivo y con menor capacidad de anticipación. Para dar respuesta a este desafío, se ha desarrollado y validado con éxito en condiciones reales la tecnología GreenView®, una solución integral que transforma la gestión de las infraestructuras urbanas.

El piloto ha demostrado la capacidad de GreenView® para digitalizar la ciudad a gran escala mediante la captura masiva de imágenes 360° y analizarlas automáticamente con Inteligencia Artificial. La tecnología ha probado su fiabilidad en la detección de una amplia gama de incidencias (obstrucción de señales, estado de alcorques, etc.), alcanzando los objetivos de rendimiento (precisión y exhaustividad) fijados al inicio del proyecto.

La clave del éxito de GreenView® reside en dos pilares:

1. Una tecnología eficiente: Su motor de análisis híbrido optimiza los costes al combinar una IA de cobertura amplia con otra de alta precisión que actúa de forma selectiva.
2. Una arquitectura estratégica y abierta: Su diseño agnóstico y preparado para el futuro garantiza la independencia tecnológica del Ayuntamiento, permitiendo la integración de modelos opensource y aprendizaje continuo en los modelos de IA propios.

Este informe no solo detalla los resultados positivos del piloto, sino que establece las bases técnicas y estratégicas para que el Ayuntamiento pueda incorporar en futuros pliegos de contratación la exigencia de usar tecnologías abiertas, interoperables y con capacidad de aprendizaje como GreenView®.

# Índice

Índice.....	2
0. Resumen Ejecutivo.....	6
El piloto de GreenView® funciona y es viable.....	6
Lo que hemos validado .....	6
Los aprendizajes clave.....	6
La viabilidad económica .....	7
Cómo está construido el sistema .....	8
Los resultados del piloto .....	8
Qué puede venir después.....	9
Conclusiones .....	10
Estructura del documento.....	10
1. Introducción: Del reto a la solución tecnológica .....	11
1.1. El desafío del mantenimiento urbano: Un reto de escala y eficiencia .....	11
1.1.1. El patrimonio verde de Madrid: Magnitud del desafío.....	11
1.1.2. El modelo tradicional y las limitaciones de un enfoque reactivo .....	11
1.1.3. Oportunidad de la transformación .....	12
1.2. GreenView® como solución estratégica .....	13
1.2.1. Los tres pilares tecnológicos de GreenView® .....	13
1.2.2. Los dos pilares de valor estratégico.....	15
1.3. Marco del proyecto: Desafía Madrid y el compromiso con la innovación pública .....	16
1.3.1. El programa Desafía Madrid.....	16
1.3.2. Colaboración entre actores .....	16
1.3.3. Aprendizaje como objetivo secundario .....	16
2. La tecnología GreenView®: Arquitectura y metodología .....	17
2.1. Arquitectura agnóstica y abierta.....	17
2.1.1. Principios arquitectónicos fundamentales.....	17
2.1.2. Arquitectura y flujo de datos del sistema .....	17
2.1.3. Independencia tecnológica: diseño para el futuro .....	20
2.1.4. Escalabilidad y rendimiento.....	21
2.1.5. Plataforma tecnológica base: SIGO.....	21
2.1.6. La interfaz de trabajo: cuadro de mandos y gestión de incidencias .....	23
2.2. Proceso de captura masiva con datos reales .....	25

2.2.1. Especificaciones técnicas del sistema de captura .....	26
2.2.2. Flujo operativo de captura .....	26
<b>2.2.3. Aspectos legales y de protección de datos .....</b>	<b>27</b>
2.2.4. Condicionantes técnicos y factores a considerar .....	28
2.3. Pipeline de procesamiento de imágenes: De vídeo 360° a datos estructurados .....	29
2.3.1. Extracción y transformación de imágenes 360° .....	29
2.3.2. Detección de objetos con YOLO .....	30
2.3.3. Asociación con el inventario GIS municipal .....	31
2.3.4. Selección inteligente de imágenes por activo .....	32
2.3.5. Recorte (crop) y preparación para análisis de IA .....	32
2.3.6. Integración con el flujo de análisis de IA .....	33
2.4. El motor de análisis de GreenView®: un modelo híbrido único .....	34
2.4.1. El desafío: equilibrar cobertura, precisión y coste .....	34
2.4.2. Arquitectura del modelo híbrido .....	34
2.4.3. Fase 1: Agentes de vistas (modelo de cobertura) .....	35
2.4.4. Fase 2: Agentes de activos (modelo de precisión) .....	37
2.4.5. Agente de priorización (scoring) .....	39
2.4.6. Flujo completo del proceso de análisis .....	41
2.4.7. Adaptabilidad y escalabilidad del modelo .....	42
2.5. Garantía de calidad: El flujo de validación asistida (Human-in-the-Loop) .....	42
2.5.1. Filosofía del Human-in-the-Loop .....	42
2.5.2. Interfaz de validación: diseño para la eficiencia .....	43
2.5.3. Mejora inmediata con recuperación de ejemplos (RAG) .....	43
2.5.4. Potencial de reentrenamiento futuro .....	46
3. Resultados del piloto y validación de GreenView® .....	46
3.1. ¿Cómo hemos validado el sistema? .....	47
3.1.1. Comprobación con incidencias reales del Ayuntamiento .....	47
3.1.2. Medición con métricas técnicas .....	51
3.1.3. Protocolo técnico de validación .....	52
3.2. Métricas de rendimiento .....	52
3.2.1. Importancia del volumen de incidencias detectadas .....	52
3.2.2. Evolución de las métricas con el entrenamiento de los agentes .....	52
3.2.3. Resultados consolidados .....	53
3.2.4. Importancia de la medición rigurosa .....	54
3.2.5. Modelos comerciales vs open-source .....	55

3.2.6. Modelos open-source en detección sobre activos.....	56
3.2.7. Proyección estratégica de los resultados.....	56
3.3. Resultados por caso de uso .....	57
3.3.1. CU-01: Interferencias con señales de tráfico .....	57
3.3.2. CU-02: Interferencias con fachadas .....	62
3.3.3. CU-03: Arbolado en mal estado .....	64
3.3.4. CU-05: Alcorques vacíos .....	68
3.3.5. CU-06: Obstrucción de vías peatonales.....	69
3.3.6. CU-08: Nidos de gran tamaño (cotorras).....	72
3.3.7. CU-09: Tocones (Exploratorio) .....	75
3.3.8. CU-10: Desbroce/escarda (Exploratorio) .....	75
3.3.9. CU-12: Papeleras llenas (Exploratorio) .....	75
3.4. Comparativa: Análisis de vistas vs análisis de activos.....	77
Ejemplo comparativo: Mismo punto, diferentes enfoques.....	77
3.5. Sistema de scoring y priorización .....	79
3.5.1. Configurabilidad según prioridades municipales.....	79
3.5.2. Proceso de co-diseño del scoring.....	80
3.5.3. Valor estratégico del enfoque preventivo .....	81
3.6. Limitaciones y próximos pasos .....	82
3.6.1. Demostrado con datos.....	82
3.6.2. Hipótesis a validar en siguiente fase.....	82
4. Viabilidad y valor .....	83
4.1. Análisis de costes de despliegue.....	83
4.1.1. Evolución del modelo de costes: de unitario a desglosado.....	83
4.1.2. Componentes del modelo de costes .....	84
4.1.3. Estrategia híbrida: optimización coste-precisión.....	89
4.1.4. Escenario de referencia: análisis de viabilidad .....	90
4.1.5. Factores de optimización dinámica.....	93
4.1.6. Proyección de costes y economía de escala .....	93
4.1.7. Capacidades de evolución y escalabilidad del sistema .....	93
4.1.8. Justificación de la inversión: valor estratégico vs. coste operativo .....	95
4.2. Independencia tecnológica y evolución continua .....	98
4.2.1. Arquitectura agnóstica .....	98
4.2.2. Interoperabilidad con sistemas municipales.....	98
4.2.3. Preparación para modelos open-source.....	98

4.2.4. Capitalización del conocimiento interno .....	98
4.3. Gestión de riesgos y estrategias de mitigación .....	98
4.3.1. Riesgo: variabilidad en el rendimiento de los modelos de IA.....	98
4.3.2. Riesgo: incremento de costes de proveedores de IA.....	99
4.3.3. Riesgo: Continuidad del Proveedor de Plataforma.....	100
4.3.4. Riesgo: Escalabilidad técnica .....	100
4.4. Especificaciones técnicas recomendadas para futuros pliegos.....	101
<b>Anexo Técnico: Especificaciones para Sistemas de Inspección Inteligente de Arbolado Urbano</b> .....	102
5. Conclusiones.....	109
5.1. Lo que hemos aprendido .....	109
5.1.1. La captura no es un problema técnico, es un problema operativo .....	109
5.1.2. El análisis en dos niveles hace viable el proyecto .....	110
5.1.3. Detectar no es suficiente, hay que priorizar .....	110
5.1.4. El valor está en prevenir, no en detectarlo todo .....	111
5.1.5. Cada validación humana mejora el sistema.....	111
5.1.6. La tecnología tiene que ser fácil de usar .....	112
5.1.7. Tres cambios importantes surgieron del feedback .....	112
5.1.8. Centrarse en lo importante primero .....	112
5.1.9. Lo que quedó fuera del piloto respecto al DDE.....	113
5.2. Qué podría venir después .....	114
5.3. Conclusión final.....	114

## 0. Resumen Ejecutivo

### **El piloto de GreenView® funciona y es viable**

El piloto de GreenView® ha validado que el sistema funciona. Hemos conseguido detectar incidencias en el arbolado de forma automática con suficiente precisión como para que sea útil operativamente. Esto lo hemos comprobado con casos reales documentados por el Ayuntamiento.

El objetivo fundamental es cambiar cómo gestionamos el arbolado. Tradicionalmente, actuamos cuando alguien reporta un problema: una rama caída, una señal tapada, un árbol que preocupa a los vecinos. GreenView® permite hacer las cosas al revés, revisar sistemáticamente el estado del arbolado y detectar situaciones que requieren atención antes de que nadie las reporte. Esto no sustituye los mecanismos actuales (avisos del 010, inspecciones, emergencias), los complementa. Añade una capa de anticipación que no teníamos, basada en evidencia visual geolocalizada que permite actuar preventivamente cuando tiene más sentido.

El piloto ha demostrado que esto es viable técnica, económica y operativamente.

### **Lo que hemos validado**

Durante el piloto hemos probado y medido tres aspectos fundamentales del sistema:

**La tecnología funciona.** Los agentes de IA detectan incidencias con precisión suficiente. El sistema alcanza un F1-Score superior al 60% en casos prioritarios como árboles en mal estado. En interferencias con seguridad vial, detecta correctamente 7 de cada 10 casos reales. Hemos validado las detecciones con avisos ciudadanos reales que el Ayuntamiento nos facilitó, comprobando que el sistema encuentra lo que sabemos que existe.

**El proceso es viable operativamente.** Los técnicos pueden validar incidencias en poco tiempo gracias a una interfaz ágil que presenta toda la información necesaria en una pantalla. El sistema de scoring ordena automáticamente las detecciones por criticidad, permitiendo concentrar esfuerzos en lo urgente. Y cada validación mejora el sistema inmediatamente mediante RAG, confirmando que aprende de la experiencia municipal.

**La arquitectura es independiente.** Hemos probado que podemos cambiar de proveedor de IA (OpenAI, Google, Meta) con total facilidad. La plataforma está preparada para trabajar con modelos open-source, que podrían reducir costes sustancialmente en el futuro. El Ayuntamiento mantiene propiedad total de los datos y control sobre la evolución del sistema.

### **Los aprendizajes clave**

Durante el piloto descubrimos cosas que no eran evidentes al principio. Cuatro cambios importantes surgieron del feedback continuo con los técnicos del Ayuntamiento:

- **El modelo híbrido de dos niveles.** No estaba en el diseño inicial. Surgió al buscar cómo hacer el proyecto económicamente más viable e independiente, a la vez que pudiera detectar un mayor número de incidencias. El sistema analiza primero todas las vistas panorámicas con modelos económicos buscando indicios de incidencias. Solo cuando

detecta algo sospechoso, activa un análisis detallado con modelos caros. Esto reduce el coste de IA en aproximadamente un 85% manteniendo alta efectividad.

- **La arquitectura agnóstica.** Aunque estaba prevista, no con el alcance final. El equipo planteó una preocupación legítima: ¿qué pasa si los proveedores suben los precios? Además de diseñar el sistema para que fuera completamente independiente, capaz de trabajar con cualquier proveedor de IA, se preparó para trabajar con modelos open-source. Esto garantiza flexibilidad y control.
- **El sistema de scoring configurable.** Al principio pensábamos que nuestro trabajo era detectar incidencias y entregarlas todas. Los técnicos plantearon su inquietud en base al volumen de incidencias detectadas y la viabilidad de su gestión. Así propusieron un sistema de priorización y desarrollamos un sistema que ordena automáticamente por criticidad, y lo que es más importante, configurable por el propio Ayuntamiento sin dependencia del proveedor. Si cambian las prioridades municipales, se modifican los criterios.
- **El valor está en prevenir, no en detectarlo todo.** Este fue un aprendizaje fundamental. No necesitamos un sistema que encuentre el 100% de las incidencias. Necesitamos un sistema que encuentre suficientes problemas críticos a tiempo para justificar la inversión. El Ayuntamiento ya gestiona decenas de miles de incidencias reactivamente. GreenView® añade anticipación. Una rama peligrosa detectada a tiempo, una señal obstruida liberada antes de causar un accidente, un árbol en mal estado intervenido preventivamente. Esas detecciones tienen incalculable.

### **La viabilidad económica**

Durante el piloto hemos trabajado en desarrollar un modelo de costes transparente y realista. El objetivo era entender qué componentes determinan el coste total de la solución y cómo se comporta al escalar.

El modelo identifica cuatro componentes principales: la plataforma tecnológica (que incluye infraestructura, procesamiento, motor de IA e interfaz de validación), la captación de imágenes (equipos, personal y operativa), el procesamiento de IA (que varía según volumen y modelos utilizados), y la asistencia técnica (para evolución continua del sistema).

Hemos analizado diferentes escenarios de despliegue considerando distintos alcances de cobertura, frecuencias de captura según jerarquía viaria, y combinaciones de modelos de IA (comerciales vs. open-source). El modelo híbrido de dos niveles reduce significativamente el coste de procesamiento respecto a un análisis exhaustivo. Los modelos open-source pueden representar un ahorro considerable frente a los comerciales, abriendo una ruta de optimización futura.

La justificación económica no se basa en un ahorro directo euro por euro frente al modelo actual, sino en el valor de la anticipación. Prevenir incidencias graves, planificar actuaciones de forma eficiente y reducir la exposición a riesgos tiene un valor que puede superar ampliamente el coste de operación del sistema.

**La justificación no está en ahorro directo.** El valor está en la anticipación proactiva. Prevenir la caída de una rama, evitar un accidente por una señal oculta, planificar las podas de forma eficiente tiene un valor incalculable en términos de seguridad, eficiencia y servicio público.

### **Cómo está construido el sistema**

GreenView® se basa en una arquitectura de cuatro capas diseñada para ser flexible, eficiente, independiente y capaz de evolucionar:

**Captura y almacenamiento.** Vehículos equipados con cámaras 360° y GPS de precisión recorren las calles generando aproximadamente 300 GB de datos por día. Las imágenes se almacenan en la nube con gestión optimizada de costes y trazabilidad completa.

**Procesamiento y preparación.** Un pipeline automatizado extrae imágenes de los vídeos, las transforma en vistas direccionales (Norte, Sur, Este, Oeste) mediante proyección Cubemap para eliminar distorsiones, y detecta automáticamente los activos (árboles, alcorques, señales) usando modelos YOLO.

**Motor de inteligencia artificial.** Este es el núcleo del sistema. Opera en dos niveles. El Nivel 1 analiza todas las vistas panorámicas con modelos económicos buscando zonas con potenciales incidencias. El Nivel 2 entra solo cuando el primer nivel detecta algo, usando modelos de alta precisión para diagnosticar árboles específicos. Esta arquitectura híbrida permite cobertura total con coste optimizado.

El sistema cuenta con tres agentes especializados: Inspector de Arbolado (detecta interferencias con señales, fachadas, mal estado del árbol, obstrucciones, nidos), Inspector de Alcorques (identifica alcorques vacíos, tocones, necesidad de desbroce), e Inspector de Limpieza (detecta papeleras llenas o residuos). Cada agente genera detecciones con justificación técnica explicando qué observó y por qué lo considera una incidencia.

**Plataforma de validación y gestión.** Los técnicos municipales revisan las propuestas de la IA en una interfaz ágil. Pueden confirmar, corregir o descartar detecciones en segundos. Cada validación alimenta inmediatamente el aprendizaje del sistema mediante RAG (mejora instantánea) y reentrenamiento periódico (mejora estructural). El sistema también incluye un módulo de scoring que ordena automáticamente las incidencias por criticidad según parámetros configurables por el Ayuntamiento.

### **Los resultados del piloto**

Hemos validado el sistema con dos enfoques complementarios: comprobación práctica con avisos ciudadanos reales y medición técnica rigurosa.

**Validación con casos reales del Ayuntamiento.** Partimos de avisos ciudadanos reportados en 2025 (señales tapadas, farolas ocultas, árboles que necesitaban poda). Pasamos las imágenes por GreenView® sin decirle qué esperábamos encontrar. El sistema detectó correctamente la mayoría de las incidencias que los vecinos habían reportado. Esto nos dio confianza operativa: el sistema encuentra problemas reales, no solo artefactos de laboratorio.

**Medición con métricas técnicas estándar.** Creamos datasets de más de 1.000 imágenes validadas manualmente y medimos el rendimiento con métricas estándar de la industria

(Precisión, Recall, F1-Score). Los resultados consolidados muestran que el sistema funciona especialmente bien en casos prioritarios:

- **CU-03 (Arbolado en mal estado):** F1-Score del 75% con GPT-4O, detectando 9 de cada 10 casos reales. Es el caso de uso más maduro del sistema.
- **CU-01 (Interferencias con señales):** Gemini alcanza 80% de recall, detectando 8 de cada 10 casos reales de obstrucción de señales o farolas.
- **CU-02 (Interferencias con fachadas):** Gemini detecta el 100% de los casos reales, aunque genera más falsos positivos que se validan rápidamente.

**Los modelos open-source ya son competitivos.** Hemos probado Llama-4 con resultados prometedores. En algunos casos de uso, su rendimiento es comparable a los modelos comerciales. La diferencia de coste puede ser de 20x menos. Esto confirma que la ruta hacia independencia total es viable.

**El sistema de scoring funciona.** Hemos desarrollado un algoritmo que ordena las miles de incidencias detectadas por criticidad. No todas las detecciones son igual de urgentes. Una señal de STOP tapada cerca de un colegio es más crítica que una rama baja en una calle residencial. El scoring permite gestionar eficientemente el volumen de alertas concentrando recursos en lo verdaderamente urgente.

**Comparativa: análisis de vistas vs. activos.** Descubrimos que ambos enfoques son complementarios. El análisis de vistas (contextual) detecta bien elementos pequeños en el entorno urbano (tocones, alcorques, papeleras). El análisis de activos (centrado en árboles específicos) logra mayor precisión en diagnósticos del arbolado. Lo que uno pierde, el otro probablemente lo detecta.

### **Qué puede venir después**

El piloto ha demostrado que el sistema es suficientemente eficiente para su despliegue operativo. La capacidad de detección, el proceso de validación y la arquitectura técnica están validados y son funcionales. Hay camino de mejora, como en cualquier sistema, pero no son necesidades críticas.

**El sistema mejorará naturalmente con el uso.** Cada validación de un técnico enriquece el conocimiento del sistema mediante RAG. Las capturas repetidas acumulan información temporal. Con datos de 2-3 años, se podrían identificar patrones estacionales y tendencias. Pero estas son evoluciones deseables, no requisitos previos. El sistema ya es operativamente viable.

**La migración a open-source es realista.** Los modelos open-source están evolucionando rápidamente. Estimamos que en 1-2 años alcanzarán el rendimiento de los modelos Pro actuales. La arquitectura agnóstica permite al Ayuntamiento hacer esta transición cuando lo considere oportuno, optimizando costes sin perder calidad.

**Las posibilidades de evolución son amplias.** Se pueden incorporar nuevos sistemas de captura (bicicletas, mochilas) para llegar a zonas donde los vehículos no acceden. El sistema puede ampliarse a otros casos de uso (pavimentos, señalización) siguiendo la misma arquitectura validada. La integración con el gemelo digital puede enriquecer bidireccionalmente el inventario municipal con datos actualizados automáticamente.

## **Conclusiones**

El piloto ha puesto los cimientos sólidos. GreenView® funciona, es viable, y está listo para construir sobre lo aprendido. La tecnología está validada con evidencia medible. El modelo económico es transparente y sostenible. La independencia estratégica está garantizada mediante diseño agnóstico. Los técnicos pueden usar el sistema en su trabajo diario de forma ágil.

El valor está en la anticipación. Poder detectar problemas antes de que alguien los reporte. Priorizar actuaciones con datos objetivos. Actuar preventivamente cuando tiene más sentido. Eso es lo que hemos probado que funciona.

Este documento aporta evidencia y análisis sobre lo que funciona y lo que falta por hacer. Pero la valoración estratégica sobre si el cambio hacia gestión proactiva justifica la inversión. Lo que podemos afirmar con certeza, respaldado por datos del piloto: el sistema funciona, genera valor, y está preparado para la siguiente fase.

## **Estructura del documento**

**Capítulo 1 - Introducción:** Presenta el contexto del desafío (modelo reactivo tradicional) y la propuesta de GreenView® como solución basada en captura masiva 360°, análisis con IA y arquitectura agnóstica. Explica los tres pilares de valor: precisión del diagnóstico, eficiencia operativa e independencia estratégica.

**Capítulo 2 - La tecnología GreenView®:** Describe cómo está construido el sistema. Explica la arquitectura de cuatro capas, el proceso de captura y procesamiento de imágenes, el motor de análisis híbrido de dos niveles, los tres agentes especializados, el sistema de scoring configurable, y el proceso de validación Human-in-the-Loop que genera aprendizaje continuo mediante RAG y reentrenamiento.

**Capítulo 3 - Resultados del piloto:** Presenta los resultados de la validación con datos reales. Explica los dos enfoques de validación (práctica y técnica), detalla las métricas de rendimiento por caso de uso con ejemplos documentados, analiza la comparativa entre análisis de vistas y activos, y reconoce transparentemente las limitaciones y aspectos pendientes de validación futura.

**Capítulo 4 - Viabilidad y valor:** Demuestra la sostenibilidad económica del sistema. Presenta el modelo de costes desglosado por componentes, explica cómo el modelo híbrido optimiza la relación coste-precisión, analiza las economías de escala, detalla las estrategias de independencia tecnológica, describe la gestión de riesgos, y proporciona especificaciones técnicas recomendadas para futuros pliegos.

**Capítulo 5 - Conclusiones:** Recoge los nueve aprendizajes clave del piloto, reconoce lo que quedó fuera del alcance, explica las posibilidades de evolución futura, y confirma que las bases están validadas para la continuidad del proyecto. El sistema funciona, es viable, y está listo para seguir construyendo sobre lo aprendido.

# 1. Introducción: Del reto a la solución tecnológica

Este capítulo presenta el contexto y el desafío que motiva GreenView®. Describe la magnitud del patrimonio verde de Madrid (3.400 hectáreas, 660.000 árboles, 250 millones de euros anuales) y las limitaciones del modelo tradicional de gestión reactiva. Se explican los tres pilares tecnológicos de la solución: captura masiva de datos mediante cámaras 360°, análisis con inteligencia artificial usando agentes especializados en modelo híbrido de dos niveles, y arquitectura agnóstica e independiente. El capítulo también posiciona GreenView® como plataforma evolutiva que genera valor estratégico mediante precisión del diagnóstico, eficiencia operativa e independencia estratégica, contextualizándolo dentro del marco Desafía Madrid.

El mantenimiento de las zonas verdes de Madrid, con un coste anual superior a 250 millones de euros, se ha basado históricamente en un modelo reactivo. GreenView® propone un enfoque proactivo: mediante captura masiva de imágenes 360° e inteligencia artificial, permite detectar incidencias antes de que se reporten, priorizar actuaciones y optimizar recursos. El piloto ha validado la viabilidad técnica, económica y operativa de este sistema.

## 1.1. El desafío del mantenimiento urbano: Un reto de escala y eficiencia

### **1.1.1. El patrimonio verde de Madrid: Magnitud del desafío**

El Ayuntamiento de Madrid gestiona uno de los patrimonios verdes urbanos más extensos de Europa:

- **Magnitud del activo:** La magnitud del activo que gestiona la Subdirección General de Conservación de Zonas Verdes y Arbolado Urbano es considerable: **3.400 hectáreas de zonas verdes y parques, más de 660.000 árboles** distribuidos por el viario urbano, múltiples especies arbóreas con diferentes necesidades de mantenimiento, y millones de alcorques, papeleras y elementos de mobiliario urbano asociados.
- **Inversión económica:** La inversión económica asociada supera los **250 millones de euros anuales** en mantenimiento y conservación, uno de los presupuestos de zonas verdes más importantes entre las capitales europeas. En este contexto, resulta imperativa la necesidad de optimizar recursos y justificar de forma transparente cada euro del gasto público.
- **Complejidad operativa:** La complejidad operativa se manifiesta en múltiples dimensiones: intervienen diversos actores (personal municipal, empresas contratistas, agentes ambientales y ciudadanía), se realizan actuaciones muy variadas (poda, tala, plantación, desbroce, riego, tratamientos fitosanitarios) y existe una marcada variabilidad estacional que genera picos de actividad y necesidades cambiantes según la época del año.

### **1.1.2. El modelo tradicional y las limitaciones de un enfoque reactivo**

Históricamente, la gestión del arbolado urbano de Madrid se ha basado en un **modelo de gestión reactivo**. Los mecanismos tradicionales de detección de incidencias se basan

fundamentalmente en tres vías. Los **avisos ciudadanos** a través del 010 o el portal de participación representan la primera línea, aunque presentan limitaciones significativas: el tiempo de detección puede extenderse durante semanas o incluso meses desde que aparece el problema, y solo se reportan incidencias muy evidentes o que causan molestias directas a los vecinos. Las **inspecciones manuales programadas**, realizadas por rutas de personal municipal y contratistas, enfrentan el desafío de una cobertura necesariamente limitada (resulta imposible revisar con frecuencia la totalidad del arbolado urbano), además de depender del criterio individual de cada inspector y carecer de un registro fotográfico sistemático del estado previo de cada activo. Finalmente, las **actuaciones tras incidentes** como la caída de ramas o árboles obligan a intervenciones de emergencia con costes elevados y exponen al Ayuntamiento a responsabilidades por incidentes que podrían haberse prevenido.

Este modelo reactivo genera importantes consecuencias operativas que afectan tanto a la eficiencia de la gestión como a la capacidad de respuesta del Ayuntamiento:

- **Ineficacia en la planificación:** La falta de información proactiva dificulta la previsión de cargas de trabajo y la asignación óptima de recursos. Los picos de demanda, especialmente tras eventos climáticos, resultan difíciles de gestionar cuando la planificación se basa en estimaciones históricas en lugar de en el estado real de los activos.
- **Aumento de incidencias evitables:** El modelo reactivo propicia que problemas menores evolucionen hasta convertirse en críticos por falta de detección temprana. Ramas que obstruyen señales de tráfico pueden permanecer durante meses sin ser detectadas, y árboles con deterioro progresivo no se identifican hasta que representan un riesgo grave para la seguridad pública.
- **Optimización limitada de recursos:** La ausencia de datos objetivos impide priorizar actuaciones por criticidad real, dificulta demostrar la diligencia en el mantenimiento preventivo y consume tiempo valioso de los técnicos municipales en desplazamientos para realizar inspecciones puntuales.
- **Falta de evidencia objetiva:** La ausencia de un registro histórico del estado de cada árbol complica demostrar su evolución temporal y justificar las actuaciones realizadas, además de dificultar la gestión de reclamaciones ciudadanas al carecer de documentación de respaldo.

### 1.1.3. Oportunidad de la transformación

En este contexto de limitaciones del modelo actual, convergen varios factores que hacen viable y necesaria una transformación hacia un enfoque basado en datos:

- En un contexto de presupuestos municipales ajustados, es imperativo demostrar la eficiencia del gasto en mantenimiento de zonas verdes. Cada euro invertido debe generar valor medible y auditable.
- Los ciudadanos de Madrid exigen mayor rapidez en la resolución de incidencias, transparencia en la gestión del espacio público y servicios municipales que incorporen tecnología de vanguardia.

- El Ayuntamiento asume responsabilidad legal sobre la seguridad de las infraestructuras verdes (prevención de caídas de ramas), la accesibilidad universal (itinerarios peatonales libres de obstrucciones) y la seguridad vial (señalización visible). Los incidentes derivados de un mantenimiento insuficiente pueden derivar en responsabilidad patrimonial, reclamaciones ciudadanas y costes de intervención urgente que superan significativamente el coste de un mantenimiento preventivo planificado.
- La madurez de tecnologías como las cámaras 360º de bajo coste, la inteligencia artificial de visión artificial, el almacenamiento cloud escalable y los modelos cada vez más potentes (incluyendo modelos open-source) hace viable técnica y económicamente una transformación que hace apenas cinco años resultaba impensable.
- Evolucionar de un **modelo reactivo** (actuar cuando el problema ya existe) a un **modelo proactivo** (detectar y anticipar antes de que se convierta en problema), y eventualmente a un **modelo predictivo** (prever basándose en patrones históricos).

## **1.2. GreenView® como solución estratégica**

Para abordar este desafío, se ha desarrollado **GreenView®**, una solución diseñada específicamente para la gestión de las infraestructuras verdes urbanas de Madrid.

GreenView® combina tecnología de captura, análisis mediante inteligencia artificial y metodología de priorización para mejorar la gestión municipal.

### **1.2.1. Los tres pilares tecnológicos de GreenView®**

#### **Pilar 1: Captura masiva de datos - El gemelo digital de Madrid**

- **Tecnología:** El sistema de captura se basa en cámaras 360º de alta resolución montadas en vehículos, equipadas con GPS de precisión para geolocalización exacta (menos de 5 metros de error) y configuradas para captura continua durante las rutas operativas normales.
- **Capacidad estimada:** Se estima una capacidad operativa de entre **55 y 82 km lineales por jornada** por cada vehículo equipado, generando aproximadamente **300 GB de datos diarios** (imágenes, metadatos, geolocalización) que conforman un registro visual completo y datado de la ciudad. Además, esta tecnología no se limita únicamente a vehículos motorizados: también puede adaptarse para su uso en bicicletas, patinetes o incluso a pie, mediante el empleo de una mochila equipada con el sistema de captura. Esta versatilidad resulta especialmente relevante para el mapeo y monitorización en el interior de parques y zonas peatonales donde el acceso con vehículos es limitado o inviable.
- **Valor diferencial:** En lugar de hacer inspecciones aisladas, el objetivo es mantener un registro visual actualizado de la ciudad que permita ver cómo está cualquier zona en distintas fechas, seguir la evolución de elementos concretos y poder revisar cambios ocurridos en el tiempo de manera sencilla.

#### **Pilar 2: Análisis con inteligencia artificial - El inspector virtual**

- **Arquitectura de agentes especializados:** GreenView® no es un único algoritmo, sino un **ecosistema de 3 agentes de IA especializados:**
  1. **Inspector de arbolado:** Experto en salud, seguridad y estado de los árboles
  2. **Inspector de Alcorques:** Especialista en el entorno urbano inmediato al árbol
  3. **Inspector de Limpieza Viaria:** Detección de incidencias de limpieza
- **Modelo híbrido:** El sistema opera en dos fases complementarias: la **Fase 1 (Cobertura)** analiza vistas direccionales con modelos de bajo coste, mientras que la **Fase 2 (Precisión)** realiza análisis detallado de activos con modelos especializados. El resultado es máxima cobertura con coste optimizado, logrando un 85% de ahorro en procesamiento.
- **Capacidad validada:** El sistema es capaz de detectar automáticamente **7 casos de uso** con métricas de rendimiento robustas: **Precision entre el 50 y 60%, Recall superior al 60% y F1-Score superior a 0.50** en casos prioritarios. Incorpora un **sistema de scoring** para priorización automática por criticidad y un **sistema RAG** (Recuperación de Ejemplos) que mejora el rendimiento en tiempo real con cada validación humana (ver sección 2.5.3).
- **IA Explicativa para transparencia y autonomía:** Cada detección incluye una justificación técnica que explica qué observó la IA y por qué considera que existe una incidencia. Esta capacidad explicativa cumple dos funciones esenciales: permite a los técnicos validar las detecciones de forma informada y fundamenta la parametrización autónoma del sistema de scoring por parte del Ayuntamiento. Al comprender qué factores evalúa el sistema (tipo de elemento obstruido, porcentaje de obstrucción, contexto urbano), el personal municipal puede ajustar ponderaciones según su experiencia operativa. Sin IA explicativa, el scoring sería más opaco y dependiente del proveedor; con ella, el Ayuntamiento mantiene mayor control sobre los criterios de priorización.

### **Pilar 3: Independencia tecnológica - Arquitectura agnóstica y abierta**

- **Agnóstico respecto a proveedores de IA:** La plataforma ha demostrado capacidad de operar con múltiples proveedores (OpenAI, Google, Meta), permitiendo cambiar de proveedor en cuestión de horas en lugar de meses, gracias a una arquitectura sin dependencias hard-coded.
- **Ruta hacia open-source:** El diseño contempla la integración activa de modelos open-source (Llama 4) y está preparado para alojar modelos propios en infraestructura municipal, lo que permitirá una reducción sustancial de costes en el futuro.
- **Interoperabilidad nativa:** La arquitectura incluye una API REST documentada y funcional, emplea formatos estándar (GeoJSON, JSON, CSV), contempla la integración con la plataforma MINT (arquitectura validada, pendiente implementación operativa completa - ver Cap. 2 sección 2.1.2) y garantiza la exportación de datos sin restricciones.
- **Human-in-the-Loop para capitalización del conocimiento:** La interfaz de validación ágil permite a los técnicos municipales validar detecciones en segundos. Cada

validación alimenta el sistema mediante recuperación de ejemplos (RAG), lo que busca generar una mejora continua en dos niveles: instantánea (RAG) y periódica (reentrenamiento). Este proceso busca crear un activo de conocimiento propio: modelos que puedan adaptarse progresivamente al contexto local, sobre los que el Ayuntamiento mantiene control. El principio observado en el piloto sugiere que el sistema puede mejorar con el uso.

## **1.2.2. Los dos pilares de valor estratégico**

Más allá de la tecnología, GreenView® se fundamenta en dos pilares de valor que justifican la inversión:

### **1.2.2.1. Precisión del diagnóstico**

- **Datos objetivos, no estimaciones:** El sistema proporciona evidencia visual geolocalizada de cada incidencia, un registro histórico completo de la evolución de los activos, métricas cuantificables del rendimiento de la IA y trazabilidad completa de todas las decisiones y actuaciones realizadas.
- **Valor para el Ayuntamiento:** Esto se traduce en una justificación auditable de cada intervención, la capacidad de defensa objetiva ante reclamaciones ciudadanas con evidencia visual geolocalizada y datada, y la demostración fehaciente de diligencia en el mantenimiento preventivo. El registro histórico permite documentar el estado previo de cada activo, facilitando la gestión de expedientes de responsabilidad patrimonial y reduciendo la exposición legal del Ayuntamiento.

### **1.2.2.2. Eficiencia operativa**

- **Optimización de recursos humanos y materiales:** El sistema reduce el tiempo de decisión de horas a minutos, realiza priorización automática por criticidad, permite la agrupación inteligente de actuaciones para optimizar rutas y posibilita la planificación anticipada (1-3 meses) frente a la tradicional actuación reactiva.
- **Impacto económico:** El sistema permite reducir costes de intervenciones de emergencia mediante la detección temprana de incidencias, optimizar recursos de contratas a través de planificación basada en datos objetivos y prevenir incidentes que pueden generar responsabilidad patrimonial. La viabilidad económica se fundamenta en el valor de la anticipación proactiva frente al coste de la gestión reactiva tradicional.

### **1.3. Marco del proyecto: Desafía Madrid y el compromiso con la innovación pública**

#### **1.3.1. El programa Desafía Madrid**

GreenView® se desarrolla en el marco del programa **Desafía Madrid**, una iniciativa del Ayuntamiento de Madrid que busca impulsar la innovación en la gestión municipal mediante la colaboración público-privada.

- **Objetivos del programa:** El programa busca identificar retos reales de gestión municipal que puedan beneficiarse de tecnologías emergentes, conectar a equipos técnicos municipales con empresas innovadoras, validar soluciones tecnológicas en entorno real de operación y generar aprendizajes transferibles a otras áreas municipales.
- **Valor del marco Desafía:** Este marco aporta **co-creación real** mediante iteración continua con técnicos municipales durante todo el desarrollo, **validación práctica** del piloto en condiciones operativas reales (no en laboratorio), **compromiso mutuo** donde el Ayuntamiento aporta conocimiento del dominio y la empresa aporta capacidad tecnológica, y **orientación a resultados** con objetivos medibles y verificables desde el inicio.

#### **1.3.2. Colaboración entre actores**

- **Subdirección General de Zonas Verdes:** La Subdirección ejerce el liderazgo estratégico y la definición de necesidades operativas, valida las funcionalidades y casos de uso propuestos, proporciona acceso a datos, sistemas y conocimiento experto, y mantiene la visión de escalabilidad y sostenibilidad a largo plazo del proyecto.
- **SiGO (proveedor tecnológico):** SiGO asume el desarrollo de la plataforma GreenView®, la integración de tecnologías de IA y visión artificial, el diseño de la arquitectura de sistemas y la experiencia de usuario (UX/UI), así como la gestión de la infraestructura cloud y la operativa técnica.
- **Programa Desafía:** El programa proporciona coordinación y facilitación del proyecto, realiza el seguimiento de hitos y entregables, ofrece apoyo metodológico en el proceso de validación y documenta los aprendizajes y buenas prácticas generadas.

#### **1.3.3. Aprendizaje como objetivo secundario**

Más allá de validar GreenView®, el piloto tiene un **objetivo secundario fundamental**: generar aprendizajes sobre cómo la administración pública puede **colaborar eficazmente con proveedores tecnológicos** en proyectos de IA, **validar rigurosamente** el rendimiento de sistemas de machine learning, **garantizar la independencia tecnológica** mediante arquitectura agnóstica, **establecer requisitos técnicos** para futuros pliegos de contratación y **servir de modelo** para otros proyectos de transformación digital municipal.

Estos aprendizajes, documentados en este informe, tienen valor más allá del caso específico de GreenView® y pueden aplicarse a cualquier proyecto de innovación tecnológica en el sector público.

## **2. La tecnología GreenView®: Arquitectura y metodología**

Este capítulo explica cómo está construido el sistema y cómo funciona técnicamente. Se describe la arquitectura agnóstica y abierta basada en cuatro capas (captura, procesamiento, motor de IA, plataforma de validación), el proceso completo de captura masiva con cámaras 360° (300 GB/día por vehículo), y el pipeline de procesamiento que transforma vídeos panorámicos en vistas direccionales analizables. El núcleo es el motor de análisis híbrido en dos niveles: agentes de vistas para cobertura amplia con modelos económicos, y agentes de activos para diagnósticos precisos con modelos premium solo cuando el primer nivel detecta algo. También se explican los tres agentes especializados (Inspector de Arbolado, Alcorques, Limpieza), el sistema de scoring configurable por el Ayuntamiento, y el proceso de validación Human-in-the-Loop que genera aprendizaje continuo mediante RAG y reentrenamiento.

GreenView® se diseñó desde el inicio con cuatro ideas claras: que fuera flexible, eficiente, independiente y capaz de evolucionar. En este capítulo explicamos cómo está construido el sistema y cómo funciona la metodología para gestionar el arbolado de Madrid a escala.

### **2.1. Arquitectura agnóstica y abierta**

#### **2.1.1. Principios arquitectónicos fundamentales**

El núcleo de GreenView® es una arquitectura agnóstica respecto al proveedor de IA. El sistema no depende de OpenAI, Google o Microsoft de forma exclusiva: cambiamos entre ellos sin tener que reescribir código. De hecho, durante el piloto validamos esto pasando de un proveedor a otro en cuestión de horas.

Esta flexibilidad abre una posibilidad estratégica importante: integrar modelos open-source alojados en infraestructura propia. Si el Ayuntamiento decide avanzar en esa dirección, puede eliminar dependencias externas y reducir significativamente los costes operativos conforme crece el volumen de análisis.

#### **2.1.2. Arquitectura y flujo de datos del sistema**

Para entender cómo funciona GreenView®, lo más útil es seguir el recorrido de los datos desde que se capturan en la calle hasta que se convierten en actuaciones concretas. El sistema está organizado en cuatro capas que se activan secuencialmente, cada una con su función específica pero diseñadas para trabajar juntas sin crear dependencias rígidas.

##### **Capa 1: Captura y almacenamiento de datos**

Todo empieza en la calle. Un vehículo equipado con una cámara 360° y GPS de precisión recorre su ruta mientras graba continuamente y registra la posición exacta en cada momento. El volumen de datos que genera es considerable: aproximadamente 300 GB por día de captura.

Los vídeos en bruto se suben a Azure, donde se almacenan usando tiers económicos (Archive/Cool) para optimizar costes. Al mismo tiempo, una base de datos geoespacial guarda los metadatos fundamentales de cada captura: coordenadas GPS, hora y condiciones de

captura (luz, clima, etc.). Esta separación entre datos brutos (los vídeos) y metadatos (la información estructurada) permite gestionar eficientemente el volumen sin perder trazabilidad.

### **Capa 2: Procesamiento y preparación**

Una vez los vídeos están almacenados, comienza el procesamiento. Un pipeline automatizado extrae imágenes discretas de los vídeos 360°, típicamente una cada 8-10 metros de recorrido. Cada imagen extraída se georeferencia con sus coordenadas GPS correspondientes.

Después viene una transformación importante, las imágenes equirectangulares (el formato nativo de las cámaras 360°) se convierten en proyecciones cúbicas mediante transformación Cubemap. ¿Por qué? Porque las imágenes equirectangulares tienen mucha distorsión, especialmente en los polos, y los modelos de IA funcionan mucho mejor con proyecciones lineales como las del Cubemap. El resultado son cuatro vistas direccionales por cada punto (Norte, Sur, Este, Oeste) que los modelos de detección pueden analizar con precisión.

A partir de ahí, el sistema identifica qué activos (árboles) aparecen en cada imagen y selecciona las vistas para cada uno.

**Mejora prevista para futuras fases:** Está planificado desarrollar un módulo de calibración GPS que corrija imprecisiones del posicionamiento mediante ajuste al eje de la vía, triangulación con elementos fijos y consolidación de detecciones entre imágenes consecutivas.

### **Capa 3: Motor de inteligencia artificial**

Con las imágenes ya preparadas, entra en acción el núcleo analítico del sistema. Aquí es donde los agentes de IA analizan lo que ven y detectan incidencias.

El proceso funciona en dos niveles. Primero, los **agentes de vistas (Nivel 1)** hacen un análisis de cobertura amplia: miran las cuatro vistas direccionales de cada punto de muestreo buscando zonas que puedan tener algún problema. Son rápidos y económicos, y su objetivo es no perderse nada. Si detectan algo sospechoso, activan el segundo nivel.

Los **agentes de activos (Nivel 2)** entran solo cuando el primer nivel ha visto algo. Estos se centran en analizar árboles específicos con mucho más detalle y precisión. Son más lentos y caros, pero hacen diagnósticos fiables. Esto permite concentrar el esfuerzo (y el presupuesto) solo donde hace falta.

Una vez detectada una incidencia, un **agente de priorización** le asigna un score de criticidad según varios factores: tipo de problema, ubicación, contexto urbano. Esto permite ordenar las actuaciones según urgencia real.

Lo interesante de esta arquitectura es su modularidad: cada agente funciona de forma independiente y puede usar diferentes modelos de IA según lo que necesite. Unos pueden usar GPT, otros Gemini, otros modelos open-source como Llama. El sistema no impone restricciones.

### **Capa 4: Plataforma de validación y gestión**

La última capa es donde trabajan los técnicos municipales. Aquí llegan todas las incidencias detectadas por la IA, pero aún no son definitivas: son propuestas que necesitan validación humana.

La plataforma incluye un **visor 360° interactivo** que permite navegar por las zonas capturadas como si estuvieras allí. Puedes moverte por las calles, girar la vista, ver el contexto completo de cada árbol. Es como Google Street View pero con el histórico temporal de Madrid y las incidencias marcadas.

El **módulo de validación (Human-in-the-Loop)** es el núcleo del sistema de garantía de calidad. Cada técnico puede revisar las incidencias detectadas, confirmarlas si son correctas, corregirlas si necesitan ajustes, o descartarlas directamente si son falsos positivos. Todo esto en segundos, porque la interfaz presenta toda la información necesaria en una única pantalla: imagen 360° del punto, fotos del árbol desde diferentes ángulos, justificación técnica de la IA, ubicación en mapa, score de criticidad.

Cada validación que hace un técnico se registra al instante y, gracias al sistema RAG (Retrieval-Augmented Generation), mejora inmediatamente el rendimiento del sistema para futuros análisis. Cuanto más se usa la plataforma, mejor funciona.

Finalmente, el **dashboard analítico** muestra KPIs, métricas de rendimiento y el estado actual de las incidencias. Puedes ver cuántas incidencias hay por distrito, cuáles son críticas, qué casos de uso son más frecuentes, cómo evoluciona el estado del arbolado.

La plataforma también ofrece una **API de integración** para conectar con sistemas municipales existentes como MINT. Se podrán generar exportaciones en CSV o Excel con todas las incidencias detectadas y validadas, listas para importar en otros sistemas de gestión municipal.

**Pendiente:** Todavía queda por completar el mapeo de todas las tipologías de incidencias a campos MINT y validar la carga completa en un entorno real de producción.

### Flujo completo de principio a fin

Juntando todas las piezas, el proceso completo funciona así:

1. **Captura en campo:** El vehículo recorre su ruta grabando vídeo 360° con GPS
2. **Almacenamiento:** Los vídeos se suben a Azure, los metadatos a la base de datos geoespacial
3. **Extracción:** Un pipeline extrae imágenes georreferenciadas cada 8-10 metros
4. **Transformación:** Las imágenes equirectangulares se convierten en 4 vistas cubemap (N, S, E, O)
5. **Selección:** El sistema identifica qué árboles aparecen y elige las mejores vistas de cada uno
6. **Análisis Nivel 1:** Los agentes de vistas analizan cada punto buscando potenciales incidencias
7. **Análisis Nivel 2:** Si el Nivel 1 detecta algo, los agentes de activos analizan el árbol en detalle
8. **Priorización:** El agente de priorización asigna un score y clasifica la incidencia (crítica, importante, menor)
9. **Validación humana:** Un técnico municipal revisa la propuesta de la IA y decide si la confirma, corrige o descarta
10. **Registro:** Las incidencias validadas se guardan en el sistema

11. **Mejora continua:** Los datos validados alimentan el sistema RAG para mejorar futuros análisis
12. **Exportación:** Las incidencias se podrán exportar vía API o CSV hacia sistemas municipales como MINT



Este flujo end-to-end transforma datos brutos de la calle en información procesable para la gestión del arbolado, manteniendo siempre el control humano sobre las decisiones finales.

### 2.1.3. Independencia tecnológica: diseño para el futuro

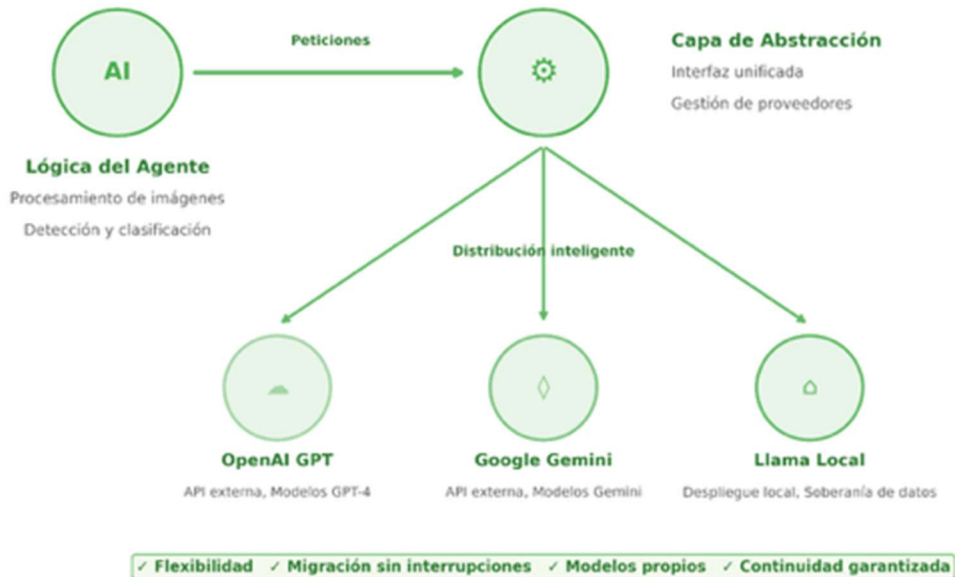
La arquitectura de GreenView® está pensada para garantizar la soberanía del dato y del proceso. Esto lo conseguimos con tres estrategias principales.

#### Abstracción de proveedores de IA

Cada agente de IA se comunica con los modelos a través de una capa de abstracción. ¿Qué significa esto? Que podemos cambiar el proveedor sin tocar la lógica del agente. La lógica permanece constante mientras la capa de abstracción redirige las peticiones al proveedor que queramos usar: OpenAI GPT, Google Gemini o incluso Llama 4 en local.

En la práctica, esto le da al Ayuntamiento flexibilidad para negociar mejores precios, migrar a proveedores más competitivos o desplegar modelos propios sin que nada deje de funcionar.

## Arquitectura de Abstracción GreenView®



### Capitalización del conocimiento interno

El ciclo de validación humana genera un activo estratégico único para el Ayuntamiento. Cada vez que un técnico municipal valida una incidencia, está creando un dato de entrenamiento curado y verificado. Con el tiempo, el sistema se vuelve más preciso y reduce progresivamente la carga de validación.

La herramienta aprende de los expertos municipales y se especializa en las características específicas de Madrid: su arbolado, sus especies y sus problemas típicos.

### 2.1.4. Escalabilidad y rendimiento

La arquitectura puede escalar desde un piloto limitado hasta un despliegue completo que cubra todo el patrimonio verde de la ciudad. Una primera fase de despliegue podría procesar cientos de miles de imágenes anuales, mientras que un despliegue completo manejaría varios millones de imágenes al año.

La estrategia de escalado se basa en procesamiento en batch paralelo con auto-escalado según demanda.

### 2.1.5. Plataforma tecnológica base: SIGO

GreenView® no parte de cero. Se apoya en la plataforma SIGO, que aporta una base sólida y más de diez años de experiencia en gestión de infraestructuras urbanas.

#### 2.1.5.1. La plataforma SIGO como fundamento

SIGO es una empresa tecnológica especializada en soluciones de gestión de infraestructuras urbanas. Su plataforma constituye la base sobre la que construimos GreenView® para el piloto de Madrid, lo que nos permitió avanzar más rápido de lo que habríamos podido partiendo desde cero.

La plataforma incluye varias piezas clave. Por un lado, un gemelo digital urbano con visualización 360º tipo “StreetView” propia, que permite navegar por el histórico temporal de cada punto específico, comparar estados pasados y actuales, y consultar las incidencias detectadas por IA. Por otro, un motor de procesamiento de imágenes con capacidad para manejar volúmenes masivos de vídeo 360º y almacenamiento optimizado.

El sistema de gestión de incidencias ofrece un workflow completo de detección, validación (Human-in-the-Loop), priorización y seguimiento de actuaciones. La arquitectura de integración proporciona un framework de APIs REST para conectar con múltiples proveedores de IA y sistemas GIS. Finalmente, el dashboard analítico incluye herramientas de visualización de KPIs y métricas de rendimiento.

SiGO cuenta con amplia experiencia en el sector público, lo que garantiza un profundo conocimiento de las necesidades, procesos y normativa de las administraciones. Esta trayectoria aporta un nivel de madurez tecnológica TRL 7-8 (Technology Readiness Level), con módulos ya en producción parcial.

La plataforma ha demostrado su fiabilidad en múltiples proyectos, con comprensión profunda del sector público. La reutilización de componentes existentes acelera la implantación, y el soporte continuo permite evolución, mantenimiento y adaptación a nuevos requisitos.

#### **2.1.5.2. Posibilidad de integración con sistemas GIS (ArcGIS/ESRI)**

GreenView® está diseñada con capacidad de integración con sistemas de información geográfica estándar. El Ayuntamiento de Madrid utiliza tecnología ArcGIS (ESRI), y la arquitectura de GreenView® permite una futura integración mediante formatos estándar (GeoJSON, Shapefile y Feature Services compatibles con ArcGIS), APIs REST para intercambio de datos geoespaciales, y consultas espaciales para asociar detecciones con el inventario GIS municipal existente.

Esta compatibilidad garantiza que los datos generados pueden integrarse con el ecosistema tecnológico GIS del Ayuntamiento cuando se considere oportuno, aprovechando la infraestructura existente sin generar silos de información.

#### **2.1.5.3. Exportación y API para terceros**

La plataforma permite exportar los resultados de análisis en formatos estándar:

- Exportación tabular: Las incidencias detectadas podrán exportarse en formato CSV o Excel, con toda la información estructurada: ubicación, tipo de incidencia, nivel de criticidad, fecha de detección, estado de validación. Esto facilita la integración con MINT u otros sistemas de gestión municipal mediante importación de datos.
- API REST: GreenView® expone una API REST que permite a sistemas terceros consultar incidencias, recuperar imágenes y metadatos, y actualizar estados de actuación. Esto

permite integraciones más sofisticadas donde otros sistemas pueden consumir datos de GreenView® de forma automática.

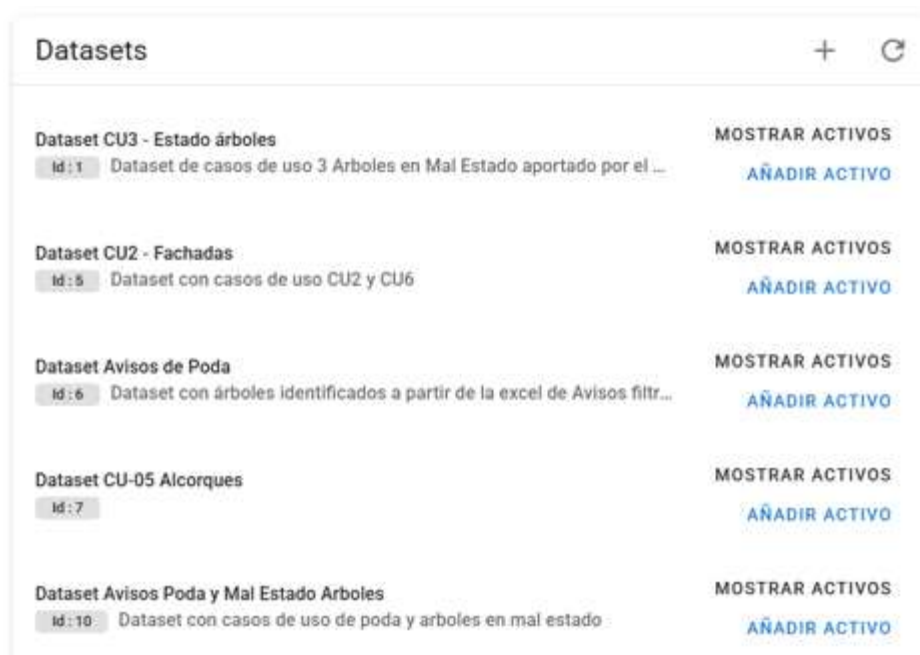
## 2.1.6. La interfaz de trabajo: cuadro de mandos y gestión de incidencias

La plataforma GreenView® incluye una interfaz web que centraliza todo el flujo de trabajo, desde la creación de datasets hasta la gestión de incidencias. Está diseñada para que los técnicos municipales puedan trabajar de forma ágil sin necesidad de formación técnica avanzada.

### 2.1.6.1. Vista principal: gestión de datasets

Al entrar en la plataforma, lo primero que ves es una tabla con todos los datasets de análisis creados hasta el momento. Cada fila representa un dataset y muestra información básica: nombre, ID, descripción y cuántas imágenes incluye. Desde aquí puedes crear un nuevo dataset con el botón “Crear Nuevo Dataset”.

Cuando haces clic en un dataset específico, accedes a su vista de detalle. Desde ahí puedes hacer varias cosas: subir nuevas imágenes al dataset, lanzar uno o varios análisis seleccionando qué agentes quieres usar, y consultar el historial de análisis que ya se han ejecutado sobre ese conjunto de datos.



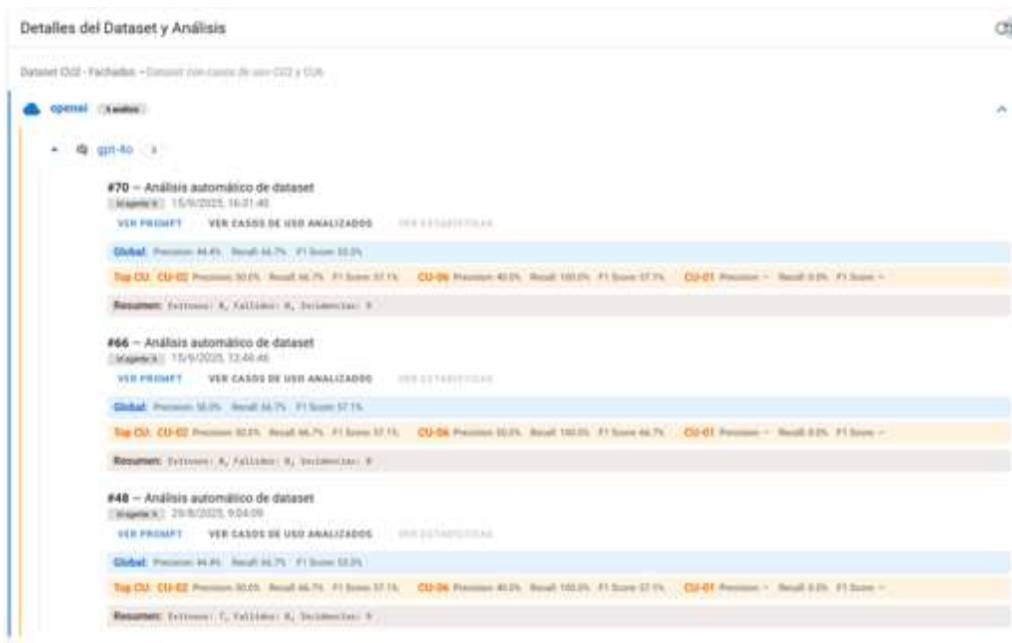
Datasets		+	↻
Dataset CU3 - Estado árboles		MOSTRAR ACTIVOS	
Id: 1	Dataset de casos de uso 3 Arboles en Mal Estado aportado por el ...	AÑADIR ACTIVO	
Dataset CU2 - Fachadas		MOSTRAR ACTIVOS	
Id: 5	Dataset con casos de uso CU2 y CU6	AÑADIR ACTIVO	
Dataset Avisos de Poda		MOSTRAR ACTIVOS	
Id: 6	Dataset con árboles identificados a partir de la excel de Avisos filtr...	AÑADIR ACTIVO	
Dataset CU-05 Alcorques		MOSTRAR ACTIVOS	
Id: 7		AÑADIR ACTIVO	
Dataset Avisos Poda y Mal Estado Arboles		MOSTRAR ACTIVOS	
Id: 10	Dataset con casos de uso de poda y arboles en mal estado	AÑADIR ACTIVO	

### 2.1.6.2. Visualización de resultados

Una vez que un análisis termina, puedes seleccionarlo para ver los resultados. La plataforma te muestra primero un resumen con métricas globales: número total de incidencias detectadas, desglose por casos de uso, y métricas de rendimiento como precisión, recall y F1 score.

Después viene una tabla detallada donde cada fila es un activo analizado (un árbol, por ejemplo). La tabla indica si tiene incidencias y muestra un resumen de los hallazgos. Una

funcionalidad interesante es el comparador de resultados: puedes seleccionar dos o más análisis del mismo dataset y compararlos lado a lado. Esto es útil para evaluar si un cambio en los agentes de IA mejoró los resultados, o para ver cómo ha evolucionado una zona con el tiempo.



### 2.1.6.3. Vista de incidencias

La plataforma tiene una sección dedicada que muestra todas las incidencias detectadas en todos los análisis. Es una tabla global con cada incidencia en una fila, mostrando su scoring de criticidad, fecha de creación, la imagen del activo que la generó, la justificación del agente de IA y su estado de revisión (pendiente, confirmada, descartada).

Esta tabla es interactiva. Puedes filtrar las incidencias por caso de uso, rango de fechas, nivel de scoring, o estado de revisión. Esto facilita la gestión: si quieres ver solo las incidencias críticas pendientes de validar, o todas las relacionadas con interferencias de señales en una zona concreta, lo tienes en dos clics.

Incidencias (1 - 10 de 37)		FIID	Caso de uso	Caso de uso esperado	Categoría	Score	Imagen	Justificación	Estado revisión	Fecha creación	Acciones
822	edab897-008f-4531-8bbe-811118f909c	CU-03	CU-03	Pendiente	22		Localización aproximada: La afectación es generalizada, cubriendo el 100% del edificio, desde la base hasta el ápice de la copa. Cuantificación visual precisa: Se observa una defoliación total del 100%. La totalidad de la copa presenta ramage fino y...	Pendiente	01/09/2025	REVISAR	
2189	e14a2c28-c320-488f-9111-565433476d30	CU-03	CU-03	Pendiente	20		Árbol con más del 70% de la copa seca, evidencia clara en imagen.	Pendiente	01/10/2025	REVISAR	
908	4bc9f08e-f139-48b0-a01a-057a7ee03f7	CU-03	CU-03	Pendiente	19		Localización aproximada: La afectación es generalizada y afecta a la totalidad de la copa del árbol. Cuantificación visual precisa: Se estima una defoliación del 100%. El 100% de la copa se observa seca y sin brotes. Terminología técnica apropiada. E...	Pendiente	16/09/2025	REVISAR	
829	06f9190e-3921-46bc-8553-66203e895a78	CU-01	CU-01, CU-06	Pendiente	19		Localización: Copa del árbol situado en primer plano, interfiriendo con una señal vertical ubicada en la acera. Cuantificación: El follaje provoca un oscurecimiento estimado del 80-90% de la superficie de la señal, impidiendo su correcta lectura. Te...	Pendiente	02/09/2025	REVISAR	

### 2.1.6.4. Módulo de validación (Human-in-the-Loop)

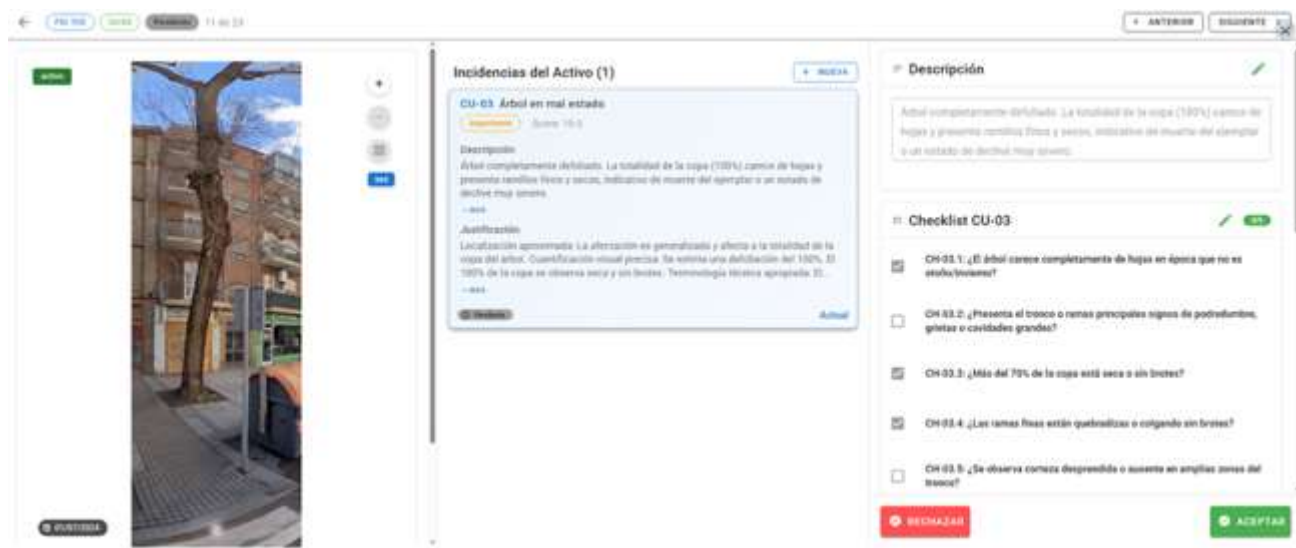
Cuando seleccionas una incidencia para revisarla, entras en el módulo de validación humana. Este es el núcleo del sistema de garantía de calidad: aquí los técnicos municipales pueden confirmar, corregir o descartar cada detección de la IA.

La pantalla de validación reúne toda la información necesaria para tomar una decisión en segundos. Ves la imagen 360º interactiva del punto donde se detectó la incidencia, las fotos de evidencia del árbol desde diferentes ángulos, y la **propuesta de la IA con su justificación técnica detallada** (explicando qué observó y por qué lo considera una incidencia). También aparece el mapa de ubicación, el scoring de criticidad asignado (basado en los factores identificados en la justificación), y toda la información contextual relevante.

El flujo es rápido: puedes confirmar la incidencia tal cual, con un clic, editarla si necesitas ajustar algo (por ejemplo, corregir el tipo de incidencia o modificar el nivel de criticidad), o descartarla directamente si es un falso positivo. Cada validación se registra al instante y, gracias al sistema RAG, mejora inmediatamente el rendimiento del sistema para futuros análisis.

Esta interfaz está diseñada para que un técnico pueda validar incidencias evidentes en cuestión de segundos, mientras que casos más complejos pueden revisarse con calma. El objetivo es mantener el control humano sobre las decisiones finales sin que la validación se convierta en un cuello de botella.

La interfaz está pensada para que el flujo de trabajo sea natural: creas un dataset, subes imágenes, lanzas análisis, revisas resultados, validas incidencias. Todo en un mismo sitio, sin tener que cambiar de herramienta.



## **2.2. Proceso de captura masiva con datos reales**

La generación de un gemelo digital actualizado comienza en la calle. El proceso de captura de GreenView® se basa en un kit de hardware ligero y adaptable, compuesto por una cámara 360º profesional y un módulo GPS de precisión, que puede ser instalado en cualquier vehículo de la flota municipal.

El sistema está diseñado para que cada vehículo equipado pueda generar un volumen de datos de aproximadamente **300 GB por día**. Esta metodología de captura masiva y de bajo coste es el cimiento sobre el que se construye todo el sistema de análisis.

### **2.2.1. Especificaciones técnicas del sistema de captura**

El sistema de captura se basa en un kit modular compuesto por cámara 360º y módulo GPS de precisión, que puede instalarse en vehículos, también en otro tipo de vehículos como bicicletas o patinetes, incluso simplemente con una mochila. El sistema registra vídeo equirectangular georeferenciado durante las rutas de captura, que posteriormente se procesa para generar imágenes discretas. La configuración técnica se optimiza para balance entre calidad de datos, operatividad y costes.

### **2.2.2. Flujo operativo de captura**

**Nota:** Este flujo describe un sistema previsto para futuras fases del proyecto. Actualmente no está en funcionamiento.

El proceso de captura sigue un flujo estructurado que garantiza la cobertura sistemática del territorio:

#### **2.2.2.1. Fase 1: Planificación de rutas**

##### **Criterios de priorización de rutas:**

1. **Densidad de arbolado:** Priorización de calles con mayor número de árboles por kilómetro lineal
2. **Criticidad operativa:** Zonas con mayor número de incidencias históricas o quejas ciudadanas
3. **Ciclos de poda:** Alineación con el calendario de mantenimiento existente
4. **Accesibilidad:** Viabilidad de circulación del vehículo de captura

#### **2.2.2.2. Fase 2: Ejecución de la captura**

La jornada de captura comienza con la preparación del equipo: se monta el kit en el vehículo, se verifica que la cámara y el GPS funcionen correctamente, y se sincroniza la hora entre ambos dispositivos para garantizar la coherencia de los datos. Una vez listo, el conductor sigue la ruta planificada mientras la cámara graba de forma continua y el sistema registra automáticamente las coordenadas GPS. La captura se realiza preferentemente en condiciones de luz natural para optimizar la calidad de las imágenes.

Al finalizar el recorrido, se procede al apagado seguro de la cámara, se extrae la tarjeta SD con los datos capturados y se registra la zona que se ha cubierto durante la jornada. Este protocolo busca mantener un equilibrio entre la rigurosidad técnica necesaria y la simplicidad operativa que permita su ejecución de forma sostenible día a día.

Para zonas donde no es viable la captura vehicular (zonas peatonales, parques, calles estrechas), el sistema contempla métodos alternativos:

- Mochilas con cámara 360°: Permiten capturar zonas peatonales y parques. El operador camina por la zona con el equipo, capturando imágenes con la misma tecnología que en los vehículos. Rendimiento estimado: 2-3 km/día en zonas de alta densidad.
- Bicicletas y patinetes equipados: Para rutas ciclables y zonas verdes amplias. Ofrecen mayor velocidad que la captura a pie manteniendo accesibilidad a zonas sin calzada.
- Captura específica en zonas verdes: En parques y jardines se pueden combinar diferentes métodos según el terreno: vehicular en vías internas, mochila en senderos peatonales.

Todos estos métodos generan el mismo tipo de datos (imágenes 360° + GPS) que se procesan de forma idéntica en la plataforma.

### **2.2.2.3. Fase 3: Volcado y procesamiento inicial**

Al finalizar cada jornada, los datos capturados se transfieren desde la tarjeta SD a la estación de volcado, verificando su integridad. Posteriormente, se realiza una carga automática a la plataforma de almacenamiento en la nube, donde los archivos se organizan de forma estructurada. Finalmente, un pipeline automatizado procesa los vídeos para extraer imágenes discretas a distancia configurable, georreferenciando cada una con sus metadatos GPS correspondientes.

El volumen de almacenamiento proyectado es considerable:

Período	Datos en bruto (vídeo)	Imágenes extraídas	Total
<b>1 día (1 vehículo)</b>	285 GB	15 GB	~300 GB
<b>1 mes (1 vehículo)</b>	6 TB	300 GB	~6.3 TB
<b>1 año (1 vehículo)</b>	70 TB	3.6 TB	~74 TB

Los datos capturados se protegen en todo momento. Solo el personal autorizado del Ayuntamiento tiene acceso a las imágenes. Durante la captura, los datos están en la tarjeta SD que permanece en el equipo hasta el volcado. Una vez volcados a la plataforma cloud, se almacenan en servidores cifrados ubicados en la Unión Europea con controles de acceso estrictos.

Las imágenes se procesan automáticamente para anonimizar rostros y matrículas antes de que cualquier persona las vea. El acceso a la plataforma requiere autenticación y cada acción queda registrada para auditoría. Los datos no se comparten con terceros ni se usan para ningún propósito ajeno a la gestión del patrimonio verde municipal.

### **2.2.3. Aspectos legales y de protección de datos**

La captura masiva de imágenes en el espacio público requiere un tratamiento riguroso de la normativa de protección de datos personales.

#### **2.2.3.1. Marco normativo aplicable**

- **RGPD (Reglamento General de Protección de Datos)**

- **LOPDGDD (Ley Orgánica 3/2018 de Protección de Datos)**
- **Normativa municipal de videovigilancia**

### **2.2.3.2. Medidas de cumplimiento implementadas**

El sistema implementa varias medidas para garantizar el cumplimiento de la normativa de protección de datos:

#### **1. Anonimización automática**

Se aplican técnicas de procesamiento de imagen para proteger la privacidad:

- Desenfoque automático de rostros detectados mediante algoritmos de visión por computador
- Difuminado de matrículas de vehículos
- Procesamiento en tiempo real antes del almacenamiento definitivo

#### **2. Limitación de finalidad**

El uso de las imágenes está estrictamente delimitado:

- Uso exclusivo para la gestión del patrimonio verde municipal
- Prohibición expresa de uso para videovigilancia o identificación de personas
- Acceso restringido únicamente a personal autorizado del Ayuntamiento

#### **3. Minimización de datos**

Se aplican principios de reducción del dato al mínimo necesario:

- Retención limitada: eliminación de imágenes tras el análisis, conservando únicamente metadatos de incidencias
- Almacenamiento cifrado (AES-256) en repositorios seguros
- Sistema de auditoría completa de todos los accesos a los datos

### **2.2.4. Condicionantes técnicos y factores a considerar**

Para un despliegue operativo completo del sistema de captura, es importante tener en cuenta diversos factores que pueden afectar la calidad y eficiencia:

#### **2.2.4.1. Condiciones meteorológicas y de iluminación**

Condición	Impacto en captura	Mitigación
<b>Luz solar directa intensa</b>	Sombras duras, pérdida de detalle en zonas oscuras	Captura preferente en horario 9:00-18:00, evitando mediodía en verano
<b>Lluvia ligera/media</b>	Goteo en lente, reducción de nitidez	Limpieza frecuente de lente, uso de tratamiento hidrofóbico
<b>Niebla/contaminación</b>	Reducción drástica de visibilidad	Suspensión de captura si visibilidad < 50 metros
<b>Noche</b>	Calidad insuficiente incluso con iluminación artificial	No se realizan capturas nocturnas

Condición	Impacto en captura	Mitigación
<b>Nubes densas</b>	Reducción de la luz natural	Captura en días con cielo despejado

### **2.3. Pipeline de procesamiento de imágenes: De vídeo 360° a datos estructurados**

Antes de que los agentes de IA puedan analizar las imágenes, necesitamos un proceso completo de visión por computador que transforme los vídeos 360° en bruto en información estructurada. Este pipeline es clave para que todo el sistema funcione de manera eficiente.

#### **2.3.1. Extracción y transformación de imágenes 360°**

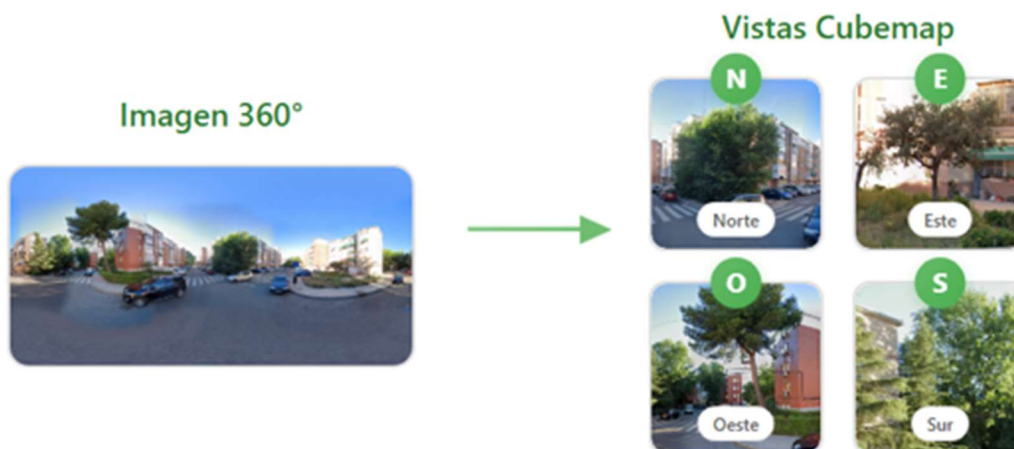
##### **2.3.1.1. Del vídeo equirectangular a vistas direccionales**

Las cámaras 360° capturan el entorno completo en un formato llamado equirectangular, que básicamente es una proyección esférica “desenrollada” en un rectángulo con proporción 2:1. Aunque captura toda la información, este formato tiene varios problemas para el análisis por IA: genera distorsión extrema en los polos (arriba y abajo de la imagen), deforma los objetos según donde estén en la esfera, dificulta que los modelos de IA detecten objetos con precisión, y produce archivos muy grandes (unos 20 MB por imagen en resolución 5.7K).

Para resolver esto, GreenView® transforma cada imagen equirectangular en 4 vistas direccionales (Norte, Sur, Este, Oeste) usando una proyección tipo Cubemap.



Esta transformación aporta varias ventajas significativas para el procesamiento. Las vistas cubemap presentan una proyección lineal sin distorsión, similar a una fotografía convencional, lo que facilita el trabajo de los modelos de detección como YOLO. Además, cada vista mantiene una resolución elevada (1920×1920px), preservando el nivel de detalle necesario para el análisis. Al procesar únicamente las cuatro vistas horizontales relevantes, se optimiza también el volumen de datos a gestionar.



## 2.3.2. Detección de objetos con YOLO

### 2.3.2.1. ¿Qué es YOLO y por qué se usa?

YOLO (You Only Look Once) es un modelo de detección de objetos en tiempo real basado en redes neuronales convolucionales. A diferencia de los modelos de IA generativa como GPT, YOLO está especializado en visión por computador. Lo usamos porque tiene tres ventajas importantes: detecta y localiza múltiples objetos simultáneamente con buena precisión, devuelve un output estructurado (bounding boxes con coordenadas exactas y la clase del objeto), y puede ejecutarse localmente sin coste por llamada a API.

#### Versión utilizada en GreenView®:

- **Modelo base:** YOLOv8 (Ultralytics)
- **Variante:** YOLOv8x (extra-large para máxima precisión)
- **Fine-tuning:** Entrenado adicionalmente con dataset de árboles urbanos

### 2.3.2.2. Clases de objetos detectados

GreenView® ha entrenado YOLO para detectar las siguientes categorías de objetos urbanos:

Categoría	Subcategorías	Uso en GreenView®
<b>Arbolado</b>	árbol, tocón	Identificación del activo principal
<b>Señalización</b>	señal de tráfico, semáforo	Detección de interferencias (CU-01)
<b>Mobiliario urbano</b>	papelera	Contexto urbano
<b>Pavimento</b>	alcorque	Análisis de estado de alcorques (CU-05)

**Nota:** Los modelos de IA en sus análisis también identifican otros elementos del entorno urbano (fachadas, vehículos, farolas, etc.) para contextualizar las incidencias detectadas.

### 2.3.2.3. Proceso de detección por imagen

Para cada vista (N, S, E, O), YOLO ejecuta:



### 2.3.2.4. Filtrado y consolidación de detecciones

**Post-procesamiento de las detecciones de YOLO:**

#### 1. Eliminación de duplicados:

- Dentro de una misma vista, si YOLO detecta el mismo objeto múltiples veces, se consolida en una única detección.
- Algoritmo de Non-Maximum Suppression (NMS) que ayuda a eliminar detecciones redundantes.

**Resultado:** Lista estructurada de todos los objetos detectados en cada vista.

### **2.3.3. Asociación con el inventario GIS municipal**

**Nota:** La asociación automática entre detecciones visuales e inventario GIS constituye una funcionalidad de implementación futura que podría mejorar significativamente la precisión del sistema.

Una de las capacidades previstas para GreenView® es asociar las detecciones visuales con los activos del inventario GIS del Ayuntamiento de Madrid.

#### 2.3.3.1. Enfoque conceptual de matching visual-GIS

El sistema podría vincular las detecciones de árboles en las imágenes con los registros del inventario municipal usando varias estrategias.

La primera sería por **proximidad geográfica**: utilizando las coordenadas GPS de cada imagen capturada, el sistema buscaría árboles del inventario en un radio determinado (por ejemplo, 15-25 metros). Los árboles más cercanos al punto de captura serían los candidatos prioritarios para el matching.

La segunda estrategia sería **triangulación direccional**: si YOLO detecta un árbol en la vista Norte, el sistema buscaría árboles del inventario situados al norte de la posición GPS de captura. Esta información direccional mejoraría la precisión del emparejamiento.

Una tercera opción interesante sería la **integración con datos LIDAR**: si estuvieran disponibles, estos datos podrían complementar las coordenadas GPS planas con información de elevación y estructura tridimensional. Esto permitiría distinguir árboles en situaciones complejas como pendientes, diferentes alturas de terreno o superposición visual.

Finalmente, el sistema también podría hacer **detección de discrepancias**: identificar árboles detectados visualmente que no existen en el inventario (árboles nuevos no catalogados) y alertar sobre árboles del inventario que no aparecen en las imágenes (posibles talas o árboles desaparecidos).

#### **2.3.3.2. Enriquecimiento de detecciones con datos GIS (objetivo futuro)**

Una vez implementado el matching visual-GIS, cada detección podría enriquecerse automáticamente con información del inventario municipal. El sistema tendría acceso al identificador único del árbol, su especie botánica, dimensiones registradas (altura, diámetro), fecha de plantación, historial de actuaciones previas y ubicación administrativa (distrito, calle).

Los beneficios de este enriquecimiento serían importantes. Los agentes de IA tendrían contexto completo del árbol para realizar diagnósticos más precisos, podrían comparar el estado visual actual con los registros históricos, se facilitaría la priorización de árboles con historial de incidencias, y se podrían asociar detecciones a responsables por zona geográfica.

#### **2.3.4. Selección inteligente de imágenes por activo**

**Nota:** Esta funcionalidad representa una optimización planificada para reducir el número de imágenes a analizar por activo, pero aún no está implementada. Parece una mejora interesante para futuras fases.

Un mismo árbol puede aparecer en múltiples imágenes capturadas desde diferentes puntos de la ruta. Un sistema de selección inteligente permitiría identificar automáticamente las mejores imágenes para el análisis de cada activo.

Los criterios propuestos para la selección incluirían varios factores. Por un lado, la distancia al árbol (priorizando imágenes capturadas a distancia óptima de 5-10 metros) y el tamaño de la detección (favoreciendo imágenes donde el árbol ocupe mayor área). Por otro, la ausencia de oclusiones (descartando imágenes donde el árbol esté parcialmente oculto) y la diversidad de perspectiva (seleccionando vistas desde diferentes ángulos). También se tendría en cuenta la calidad de iluminación, evitando imágenes sobreexpuestas o subexpuestas.

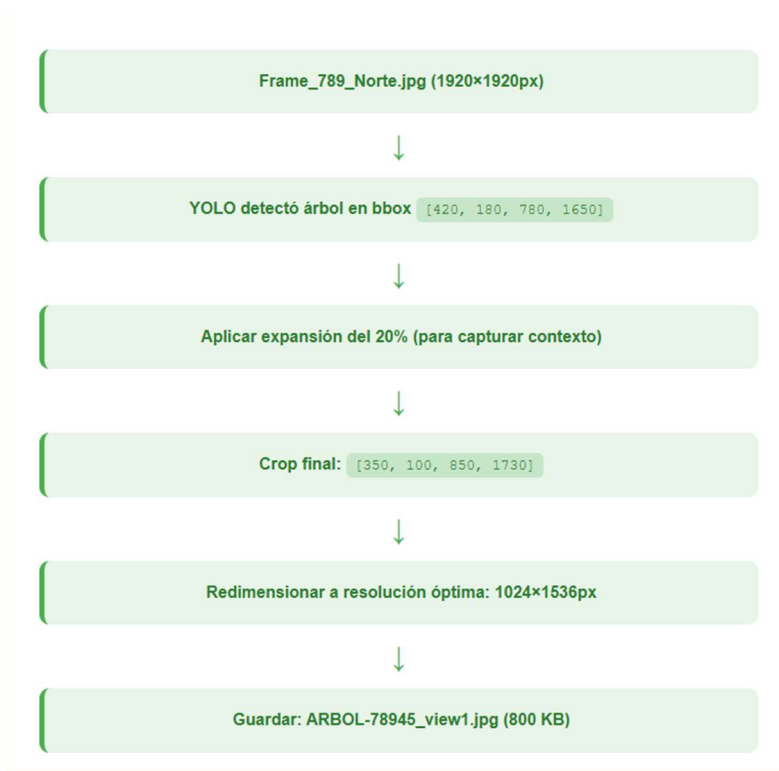
El beneficio esperado es reducir el coste de procesamiento y análisis al seleccionar solo las 3-4 mejores vistas por árbol en lugar de procesar todas las imágenes disponibles.

#### **2.3.5. Recorte (crop) y preparación para análisis de IA**

Una vez obtenidas las imágenes, se preparan para optimizar el análisis de los agentes de IA.

##### **2.3.5.1. Generación de crops de activos**

**Para cada árbol seleccionado:**



El cropping tiene tres ventajas principales: reduce el coste de las APIs (imagen más pequeña significa menos tokens en los agentes de IA), mejora la precisión (la IA se centra solo en el activo relevante), y optimiza el ancho de banda (las imágenes son más livianas para transmitir).

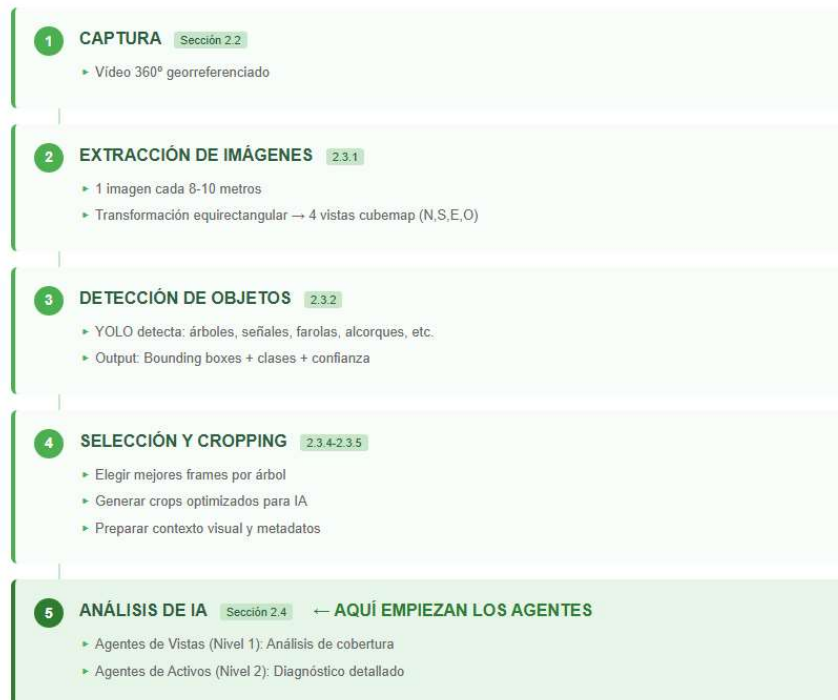
### **2.3.5.2. Preparación del contexto visual**

Además del crop del árbol, se prepara una imagen de contexto amplio: la vista completa (sin crop) para que la IA vea el entorno, y anotaciones visuales como círculos marcando el árbol o flechas señalando incidencias.

### **2.3.6. Integración con el flujo de análisis de IA**

Todo este procesamiento de visión por computador actúa como **capa de preparación** antes de que los agentes de IA entren en acción:

## PIPELINE COMPLETO DE PROCESAMIENTO



Este pipeline de visión por computador transforma los vídeos en bruto en datos estructurados listos para que los agentes de IA los analicen. Este componente busca optimizar el equilibrio entre eficiencia y precisión en el sistema.

### **2.4. El motor de análisis de GreenView®: un modelo híbrido único**

Para analizar el volumen de datos capturados, GreenView® emplea un **modelo de análisis híbrido** que combina dos enfoques en un flujo de trabajo optimizado. Este sistema permite abordar el análisis masivo de infraestructuras verdes de forma sostenible económicamente.

#### **2.4.1. El desafío: equilibrar cobertura, precisión y coste**

Analizar millones de árboles con la máxima precisión en cada uno presenta desafíos computacionales y económicos significativos. Por ejemplo, un análisis exhaustivo de alta precisión de cada árbol en cada pasada representaría un coste considerable: 1.000.000 de árboles multiplicado por 5 imágenes por árbol resulta en 5.000.000 de imágenes. Con un análisis de alta precisión a unos 0,10 € por imagen, el coste anual alcanzaría los 500.000 €, una cifra difícil de justificar operativamente.

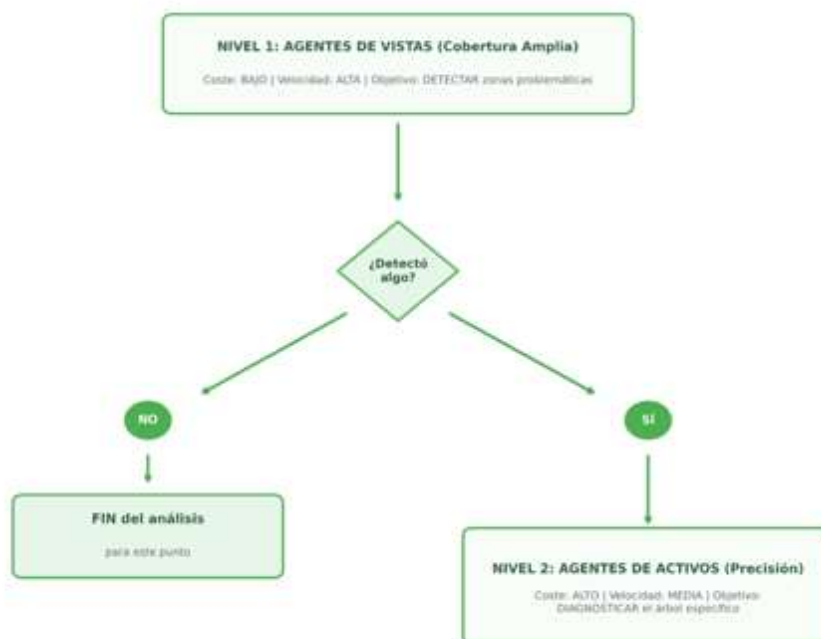
La propuesta de GreenView® es un modelo híbrido de dos niveles que busca reducir significativamente el coste manteniendo la efectividad del análisis.

#### **2.4.2. Arquitectura del modelo híbrido**

El sistema opera en dos fases secuenciales con agentes especializados en cada nivel. Funciona así: primero, los agentes de Nivel 1 (agentes de vistas) realizan un análisis de

cobertura amplia con coste bajo y velocidad alta. Su objetivo es detectar zonas potencialmente problemáticas. Si no detectan nada en un punto de muestreo concreto, el análisis termina ahí y no se gasta más presupuesto. Pero si detectan algo sospechoso, entonces se activa el Nivel 2.

Los agentes de Nivel 2 (agentes de activos) entran en acción solo cuando el primer nivel ha detectado algo. Estos agentes tienen un coste más alto y una velocidad media, pero su misión es hacer un diagnóstico preciso del árbol específico. Esto permite concentrar el esfuerzo computacional (y el presupuesto) donde realmente hay algo que revisar.



La eficiencia del modelo se basa en la activación selectiva del Nivel 2. Según las estimaciones del diseño, si aproximadamente un 30-40% de los puntos de muestreo requieren análisis detallado, el coste total podría reducirse significativamente respecto a un análisis exhaustivo. El objetivo es mantener una alta efectividad en la detección de incidencias reales, aunque estos indicadores deberán validarse durante la fase operativa con un volumen mayor de datos reales del arbolado de Madrid.

### 2.4.3. Fase 1: Agentes de vistas (modelo de cobertura)

#### 2.4.3.1. Objetivo y funcionamiento

Los **agentes de vistas** analizan la escena completa desde un punto de muestreo, observando el entorno panorámico en 4 direcciones cardinales (Norte, Sur, Este, Oeste). Su misión es **detectar rápidamente si hay alguna zona de interés** que requiera un análisis más detallado.

Los agentes de vistas trabajan con modelos rápidos y económicos de OpenAI, Google o Meta (opensource). Reciben como input cuatro imágenes correspondientes a las vistas direccionales (Norte, Sur, Este, Oeste) en resolución media. Su función es hacer una detección

de cobertura amplia, siendo tolerantes a falsos positivos porque el objetivo en esta fase es no perderse nada.

#### **2.4.3.2. Agente de vistas de nivel 1**

El sistema utiliza un **único agente de vistas unificado** que analiza de forma integral todos los casos de uso en un solo análisis. No hay múltiples agentes especializados en este nivel, sino uno solo que busca todo a la vez.

Este agente tiene tres capacidades principales que ejecuta simultáneamente. Por un lado, detecta interferencias y problemas viales como obstrucciones de señales, semáforos, farolas o invasión de aceras. Por otro, analiza el estado del arbolado buscando árboles con signos aparentes de mal estado, sequedad o inclinación peligrosa. Y finalmente, evalúa el entorno urbano inmediato, fijándose en el estado de alcorques, acumulación de residuos y otros elementos de limpieza.

La lógica de funcionamiento es sencilla: el agente analiza simultáneamente las cuatro vistas (N, S, E, O) buscando cualquier tipo de incidencia relacionada con vegetación, seguridad vial, estado del arbolado o entorno urbano. Como output, genera una lista consolidada de todas las zonas con potencial incidencia detectadas, asignando un nivel de confianza (alta, media o baja) a cada una.

#### **2.4.3.3. Caso real: Detección de obstrucción de señal de STOP**

Input:

- Vista Norte: Muestra señal de STOP con rama delante
- Vista Sur: Muestra árbol desde atrás
- Vista Este: Vista lateral del árbol
- Vista Oeste: Vista lateral opuesta

Análisis del Agente de Interferencias y Vías:

1. Identifica señal de STOP en vista Norte
2. Detecta vegetación tapando parcialmente la señal
3. Confirma presencia de árbol en vistas adyacentes
4. RESULTADO: "Potencial interferencia con señal de tráfico - Confianza: Alta"  
→ ACTIVA Nivel 2 para análisis de activos

**Nota técnica sobre activación del Nivel 2:** Cuando el agente de vistas detecta una incidencia, no puede determinar con precisión qué activo específico (árbol) la está causando. Por tanto, se analizan **todos los activos visibles en esa vista** con el modelo de precisión. En promedio, esto representa ~4 activos por vista. Esta limitación podrá optimizarse en el futuro mediante la integración del análisis con el inventario municipal de arbolado, permitiendo identificar directamente qué activo analizar.

#### **2.4.3.4. Ventajas del nivel 1**

Este primer nivel tiene cuatro ventajas fundamentales. Primera, garantiza cobertura total: analiza todo el recorrido sin excepción, ningún punto queda sin revisar. Segunda, compensa las oclusiones: al ver 360°, puede detectar árboles aunque estén parcialmente ocultos por coches, contenedores u otros obstáculos. Tercera, tiene un coste bajo que permite analizar millones de puntos de muestreo de forma económicamente viable. Y cuarta, actúa como filtro eficiente reduciendo en 60-70% el número de análisis detallados de Nivel 2 que hay que ejecutar.

## **2.4.4. Fase 2: Agentes de activos (modelo de precisión)**

### **2.4.4.1. Objetivo y funcionamiento**

Los **agentes de activos** se activan solo cuando el nivel 1 detecta algo. Su misión es realizar un **diagnóstico preciso y fiable** sobre un árbol concreto, determinando si existe realmente una incidencia, su tipo exacto y su nivel de criticidad.

Los agentes de activos trabajan con modelos de alta precisión como de OpenAI o Gemini, aunque en el futuro podrían utilizarse modelos opensource equivalentes. Reciben como input múltiples imágenes del árbol. Su función es hacer un **diagnóstico preciso y justificado** de cada incidencia detectada, explicando las razones técnicas que fundamentan la clasificación. Esta capacidad explicativa de la IA es crucial para que los técnicos municipales puedan validar las detecciones de forma informada. Aunque los falsos positivos pueden aparecer (especialmente con modelos más económicos), el sistema Human-in-the-Loop permite descartarlos rápidamente durante la validación, por lo que el coste operativo de estos errores es bajo.

### **2.4.4.2. Proceso de análisis de Nivel 2**

#### **Paso 1: Selección de imágenes del activo**

Cuando el Nivel 1 detecta una incidencia potencial en una vista, el sistema identifica qué activos (árboles) aparecen en esa vista y recupera sus imágenes cropeadas que ya fueron generadas previamente por YOLO en el pipeline de procesamiento (sección 2.3.5). Estos crops se envían directamente a los agentes de Nivel 2 correspondientes.

**Nota:** Una optimización planificada para futuras fases permitiría vincular cada activo detectado con el inventario GIS municipal. Esto permitiría buscar y seleccionar las mejores imágenes de ese activo no solo del punto de muestreo donde se detectó la incidencia, sino también de otros puntos de captura disponibles, proporcionando múltiples ángulos y perspectivas para un análisis más completo y preciso.

#### **Paso 2: Análisis especializado**

Se ejecuta el agente correspondiente al caso de uso detectado:

##### **A) Agente Inspector de arbolado**

Este inspector se encarga de los casos de uso CU-01 (Interferencias seguridad vial), CU-02 (Interferencias fachadas), CU-03 (Mal estado), CU-06 (Obstrucción vías) y CU-08 (Nidos de cotorra). Trabaja con un checklist de verificación específico para cada caso de uso, y genera un output estructurado en formato JSON que incluye las incidencias detectadas, el nivel de confianza, una justificación técnica y las imágenes de evidencia.

##### **B) Agente Inspector de alcorques**

Este inspector se especializa en CU-05 (Alcorques vacíos), CU-09 (Tocones) y CU-10 (Desbroce). Su enfoque se centra en analizar el entorno inmediato al árbol: la base, el alcorque y el pavimento circundante. También genera un output estructurado en JSON con las incidencias detectadas, confianza, justificación y evidencia.

### C) Agente Inspector de limpieza viaria

Este inspector es responsable del CU-12 (Papeleras). Se enfoca en evaluar el estado de limpieza y orden del espacio público. Al igual que los otros, produce un output estructurado en JSON con incidencias detectadas, confianza, justificación y evidencia.

#### Optimización potencial: Verificación de consistencia

**Nota:** Una mejora futura del sistema podría incluir consolidación de detecciones cuando un mismo árbol ha sido visto desde múltiples puntos de muestreo: - Comparar resultados de análisis desde diferentes ángulos - Descartar detecciones inconsistentes (probables falsos positivos) - Reforzar detecciones confirmadas desde múltiples vistas

#### 2.4.4.3. Ejemplo de análisis de Nivel 2

##### Caso: Árbol con rama que tapa semáforo



#### 2.4.4.4. IA Explicativa: fundamento de la transparencia y parametrización

Un aspecto importante del diseño de GreenView® es que **los agentes buscan generar justificaciones técnicas** de sus detecciones. No solo indican que existe una incidencia, sino que intentan explicar **qué observaron, por qué lo consideran problemático y qué factores influyeron en su evaluación**.

Esta capacidad explicativa busca cumplir tres funciones:

1. **Facilitar la validación humana:** Los técnicos municipales reciben explicaciones comprensibles que les permiten validar las detecciones de forma informada. Por ejemplo: *“Se observa una rama de aproximadamente 3m de longitud que se proyecta sobre el semáforo desde el lado derecho, tapando el círculo rojo superior”*. Este tipo de descripción permite al técnico confirmar o corregir la detección con mayor contexto.
2. **Permitir auditoría y trazabilidad:** Cada incidencia registrada incluye la justificación de por qué la IA la detectó. Esto permite auditar retrospectivamente las decisiones del sistema, entender su comportamiento en diferentes contextos y detectar posibles sesgos o errores sistemáticos.
3. **Fundamentar la parametrización del scoring:** Las justificaciones técnicas facilitan que el Ayuntamiento pueda **parametrizar el sistema de scoring de forma informada**. Al comprender qué factores observa la IA (tipo de elemento obstruido, porcentaje de obstrucción, contexto urbano), el personal municipal puede:
  - Definir qué factores deben tener mayor peso en el score
  - Ajustar umbrales según su experiencia operativa
  - Crear reglas contextuales basadas en casos reales
  - Simular cambios para evaluar su impacto antes de aplicarlos

Sin IA explicativa, el scoring sería más opaco y dependiente del proveedor tecnológico para su interpretación y ajuste. Con ella, se busca que el Ayuntamiento pueda adaptar los criterios de priorización según sus necesidades operativas.

#### **2.4.5. Agente de priorización (scoring)**

Una vez detectada una incidencia por los agentes de nivel 2, el **agente de priorización** asigna una puntuación que permite ordenar las actuaciones según su criticidad. Este componente es esencial para la autonomía operativa del Ayuntamiento.

##### **2.4.5.1. Configurabilidad y autonomía del sistema**

Gracias a la **IA explicativa** descrita anteriormente, el personal municipal comprende qué factores evalúa el sistema en cada detección. Esta comprensión fundamenta la capacidad de parametrización del scoring.

La plataforma proporciona un panel web de configuración que permite:

- **Ajuste de ponderaciones y umbrales:** Modificar los pesos de cada factor según las prioridades municipales
- **Reglas contextuales:** Activar o desactivar bonificaciones según eventos temporales (post-temporales, zonas escolares) o características geográficas

- **Personalización por tipo de incidencia:** Cada caso de uso puede definir sus propios criterios de criticidad
- **Versionado de configuraciones:** Posibilidad de revertir cambios y mantener un historial de las decisiones
- **Modo de simulación:** Probar configuraciones sobre incidencias históricas antes de aplicarlas en el sistema operativo

Este enfoque busca reducir la dependencia del proveedor tecnológico, permitiendo que el Ayuntamiento adapte los criterios de priorización según su experiencia operativa y necesidades cambiantes.

#### **2.4.5.2. Lógica de scoring**

Cada tipo de incidencia tiene factores ponderados definidos por el Ayuntamiento:

##### **Ejemplo: CU-01 (Interferencia con seguridad vial)**

Factor	Peso	Ejemplos
<b>Tipo de elemento obstruido</b>	Alto	Semáforo: +6 pts, Señal STOP: +6 pts, Señal informativa: +3 pts
<b>Nivel de obstrucción</b>	Alto	100% (ilegible): +9 pts, 50-99%: +6 pts, <50%: +3 pts
<b>Contexto adicional</b>	Medio	Cerca de colegio: +3 pts, Cruce peligroso: +3 pts
<b>Farola también obstruida</b>	Bajo	+2 pts

**Score final** = Suma de puntos → Clasificación:

- **Score ≥ 24.0:** Incidencia CRÍTICA (rojo)
- **Score 18.0 - 23.9:** Incidencia IMPORTANTE (naranja)
- **Score 12.0 - 17.9:** Incidencia INTERMEDIA (amarillo-naranja)
- **Score 6.0 - 11.9:** Incidencia MENOR (amarillo)
- **Score < 6.0:** Incidencia OBSERVACIÓN (verde)

¿Qué aporta esto en la práctica? Los técnicos municipales pueden filtrar y priorizar su trabajo viendo primero las incidencias críticas, optimizando así la planificación de las actuaciones según la urgencia real de cada caso.

#### **2.4.5.3. Derivación de casos poco críticos a contratas**

El sistema de scoring no solo sirve para priorizar el trabajo de validación del Ayuntamiento. También permite optimizar el flujo de gestión según la criticidad de cada incidencia.

Para incidencias de baja criticidad o de carácter masivo (por ejemplo, interferencias menores con fachadas, alcorques con vegetación baja, incidencias estéticas sin impacto en seguridad), el sistema puede configurarse para que pasen directamente a las empresas de conservación sin necesidad de validación previa por parte del Ayuntamiento.

Esto funciona mediante criterios configurables:

- Umbral de criticidad: Incidencias por debajo de cierto score (ej. score < 6.0) pueden derivarse automáticamente.
- Tipo de incidencia: Determinados casos de uso de bajo riesgo pueden tener derivación automática activada.
- Volumen: Si se detectan muchas incidencias similares de baja prioridad, pueden agruparse y enviarse en lote.

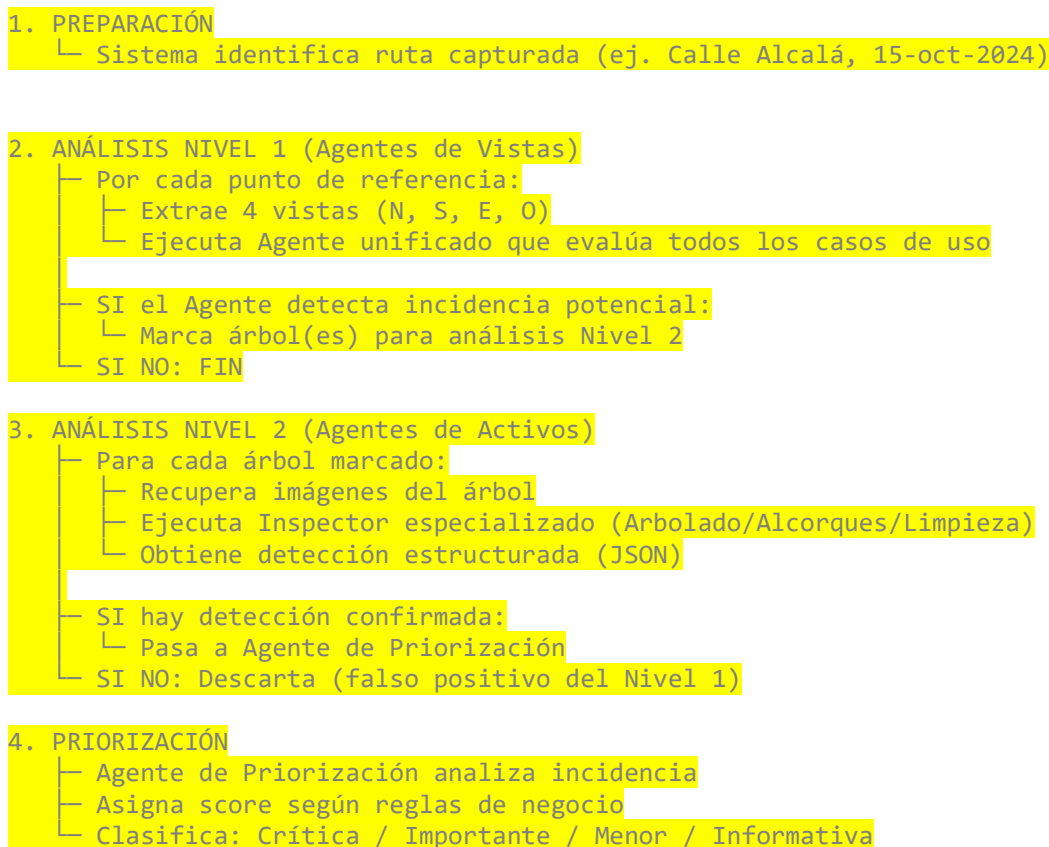
La ventaja es doble. Por un lado, el Ayuntamiento concentra su capacidad de validación en lo realmente crítico (seguridad vial, árboles en riesgo, infraestructuras importantes). Por otro, las empresas de conservación reciben flujos de trabajo continuos y organizados que pueden gestionar de forma eficiente.

En cualquier caso, todas las incidencias quedan registradas en el sistema con trazabilidad completa: qué se detectó, cuándo, qué score recibió, a qué contrata se derivó y cuál fue el resultado de la actuación. El Ayuntamiento mantiene visibilidad total y puede auditar el trabajo realizado.

Este modelo de derivación automática es opcional y configurable. El Ayuntamiento decide qué tipos de incidencias requieren validación humana y cuáles pueden gestionarse de forma directa.

#### 2.4.6. Flujo completo del proceso de análisis

Diagrama de flujo end-to-end:



## 5. REGISTRO PROVISIONAL

- └ Incidencia se guarda en base de datos como "PENDIENTE DE VALIDACIÓN"
- └ Se asocia a: árbol, imágenes de evidencia, score
- └ Queda disponible para validación humana en la plataforma

## 6. VALIDACIÓN HUMANA (ver sección 2.5)

- └ Técnico confirma/corriges/descarta → Estado final

### 2.4.7. Adaptabilidad y escalabilidad del modelo

#### 2.4.7.1. Incorporación de nuevos casos de uso

El sistema está diseñado para permitir la incorporación de nuevos casos de uso. El proceso general incluye definir qué se quiere detectar, desarrollar o adaptar el agente con su lógica de detección específica, entrenar el agente con ejemplos de validación, medir su rendimiento contra un dataset de prueba, y finalmente desplegarlo integrándolo en el pipeline sin afectar a los agentes existentes.

La duración de este proceso dependerá de la complejidad del caso de uso y de la disponibilidad de ejemplos de entrenamiento de calidad.

#### 2.4.7.2. Mejora continua de agentes

Los agentes mejoran constantemente gracias al ciclo de validación humana. Cada corrección que hace un técnico se guarda como ejemplo validado, y el rendimiento del sistema aumenta progresivamente con el uso.

El motor de análisis de GreenView® busca aprovechar las fortalezas de la IA moderna (velocidad y capacidad de procesamiento a gran escala) complementándolas con el criterio experto humano (precisión contextual y capacidad de discernimiento). El objetivo es lograr un sistema híbrido que, en la práctica, pueda ofrecer tanto cobertura como fiabilidad.

### 2.5. Garantía de calidad: El flujo de validación asistida (Human-in-the-Loop)

Toda detección realizada por la IA es considerada una “propuesta” hasta que es validada. GreenView® incorpora una interfaz de **validación asistida** diseñada para que un operador humano pueda revisar, corregir o descartar las incidencias de forma extremadamente ágil. La herramienta presenta al validador todas las evidencias visuales (imagen del activo, vista 360°, mapa) y el contexto de la detección en una única pantalla, permitiendo tomar una decisión en segundos.

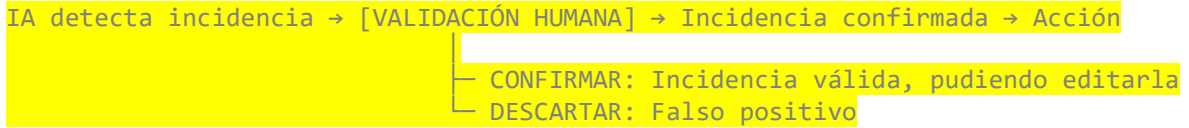
Este proceso de **human-in-the-loop** no solo garantiza la máxima calidad y fiabilidad del dato final que se entrega al gestor, sino que además genera un activo de valor incalculable: un *dataset* de incidencias curado y verificado, listo para ser utilizado en el re-entrenamiento y la mejora continua de los modelos de IA propios.

#### **2.5.1. Filosofía del Human-in-the-Loop**

El sistema GreenView® se fundamenta en un principio clave: **la IA propone, el humano decide.**

Esta filosofía reconoce tres principios fundamentales: que la IA es excelente para detectar patrones a gran escala y de forma sistemática, que el humano es insustituible para el juicio contextual y la decisión final, y que la combinación de ambos genera el mejor resultado posible.

**Flujo de confianza:**



**2.5.2. Interfaz de validación: diseño para la eficiencia**

La interfaz de validación ha sido diseñada para permitir que un técnico pueda revisar, confirmar, corregir o descartar incidencias de forma ágil. El sistema presenta en una única pantalla toda la información necesaria para tomar una decisión informada: una vista 360° interactiva que permite explorar el entorno completo donde se detectó la incidencia, imágenes de evidencia mostrando el activo desde diferentes ángulos, información contextual de la detección (tipo de incidencia, ubicación, score de prioridad), y acciones de validación simples (confirmar, editar o descartar).

El flujo de trabajo permite a los técnicos validar incidencias evidentes en pocos segundos, mientras que casos más complejos pueden editarse o ajustarse según el criterio experto del validador.

**2.5.3. Mejora inmediata con recuperación de ejemplos (RAG)**

Uno de los aspectos más innovadores de GreenView® es su capacidad de **mejora continua en tiempo real** sin necesidad de esperar ciclos de reentrenamiento. Gracias a un sistema de **Retrieval-Augmented Generation (RAG)**, cada validación humana mejora instantáneamente el rendimiento del sistema al enriquecer el contexto que se proporciona a los agentes de IA.

**Concepto clave:** Cuanto más se usa la plataforma, mejor funciona, porque el sistema aprende inmediatamente de cada interacción validada.

**2.5.3.1. Arquitectura del sistema RAG**

El sistema RAG implementado en GreenView® opera en paralelo al análisis de IA, proporcionando contexto relevante basado en casos similares previamente validados:



### 1. DETECCIÓN DE NUEVA INCIDENCIA

- Agente de IA identifica potencial incidencia
- Metadatos: tipo (CU-01), ubicación, características visuales

### 2. RECUPERACIÓN DE CASOS SIMILARES (RAG)

- Sistema busca en base de conocimiento validaciones previas:
  - Mismo caso de uso (CU-01: Interferencia con señal)
  - Misma zona geográfica (distrito Salamanca)
  - Características visuales similares (rama tapando STOP)
  - Ordenadas por relevancia (similarity score)
- Recupera top 3-5 ejemplos más relevantes

### 3. ENRIQUECIMIENTO DEL PROMPT

- Prompt base del agente + Ejemplos validados inyectados
- "Analiza esta imagen buscando interferencias con señales..."
- \*\*Contexto - Casos similares validados previamente:\*\*
- Ejemplo 1 (validado el 10-oct-2024):
  - Ubicación: Calle Serrano, 89
  - Situación: Rama de plátano tapando 60% de señal STOP
  - Decisión técnico: CONFIRMADA como crítica
  - Comentario: 'Requiere poda urgente, cruce escolar'
- Ejemplo 2 (validado el 08-oct-2024):
  - Ubicación: Calle Goya, 45
  - Situación: Sombra proyectada sobre señal
  - Decisión técnico: DESCARTADA
  - Comentario: 'No es vegetación, es sombra temporal'
- Basándote también en estos casos previos, analiza..."
- El agente tiene ahora MUCHO más contexto para su análisis

### 4. ANÁLISIS MEJORADO

- La IA analiza con ejemplos concretos de qué es incidencia real
- Aprende de matices: cuándo es crítico, cuándo es falso positivo
- Resultado: Detección más precisa desde el primer análisis

### 5. VALIDACIÓN Y RETROALIMENTACIÓN

- Técnico valida la nueva detección
- Esta validación SE AÑADE inmediatamente a la base de conocimiento
  - Disponible para próximos análisis en segundos

## **Beneficios del sistema RAG**

La principal ventaja del RAG frente al reentrenamiento tradicional de modelos es la velocidad de mejora. Mientras que reentrenar un modelo puede llevar entre uno y tres meses (requiere recopilar datos, procesarlos, entrenar con GPUs y validar), el RAG incorpora nuevos conocimientos en segundos. Cuando un técnico valida una incidencia, esa información queda

disponible inmediatamente para los siguientes análisis. Además, la complejidad técnica es considerablemente menor: no se necesita infraestructura MLOps ni equipos especializados en deep learning, solo una base de datos vectorial.

Otra característica importante es el aprendizaje acumulativo. Con cada validación que realizan los técnicos, el sistema construye una base de conocimiento más rica y matizada. No se trata de un modelo estático que se degrada con el tiempo, sino de un sistema que mejora continuamente con el uso.

Este enfoque también permite una especialización natural en el contexto de Madrid. El sistema va aprendiendo las particularidades del Platanus × hispanica (la especie dominante en la ciudad), entiende que una interferencia en Gran Vía tiene mayor criticidad que en una calle residencial, acumula patrones estacionales como la procesionaria en primavera, y asimila los criterios específicos que aplican los técnicos municipales.

Finalmente, el sistema ofrece transparencia completa. A diferencia de un modelo reentrenado donde no se puede explicar por qué tomó cierta decisión, aquí cada análisis puede justificarse mostrando exactamente qué casos históricos se utilizaron como referencia. Esto facilita la auditoría y permite depurar el sistema inspeccionando directamente la base de conocimiento.

### **Implementación técnica del RAG**

#### **Base de datos de conocimiento:**

Cada validación genera un registro que se almacena en formato estructurado:

```
{
  "id": "VAL-2024-10-15-00789",
  "caso_uso": "CU-01",
  "embeddings_visuales": [0.234, -0.891, 0.456, ...], // 1536 dimensiones
  "embedding_textual": [0.123, 0.567, -0.234, ...],
  "metadatos": {
    "ubicacion": {"lat": 40.416775, "lon": -3.703790, "distrito": "Salamanca"},
    "caracteristicas": ["arbol", "señal_trafico", "STOP", "obstruccion_parcial"],
    "contexto": "zona_escolar",
    "especie_arbol": "Platanus × hispanica"
  },
  "decision_humana": {
    "accion": "CONFIRMADA",
    "severidad": "CRITICA",
    "comentario": "Rama obstruye 70% de señal en cruce escolar",
    "tecnico": "usuario_042"
  },
  "imagenes": ["frame_12345_N.jpg", "frame_12346_E.jpg"]
}
```

#### **Proceso de recuperación:**

Cuando un agente va a analizar una nueva incidencia, el proceso de recuperación funciona en cinco pasos. Primero, se genera un embedding de la nueva detección combinando imagen y metadatos. Segundo, se realiza una búsqueda por similitud en la base de conocimiento usando similitud coseno. Tercero, se aplica un filtrado contextual para quedarse con casos del mismo tipo, de la misma zona geográfica (opcional) y con características similares. Cuarto, se

seleccionan los 3-5 ejemplos más relevantes. Y quinto, estos ejemplos se inyectan en el prompt como contexto adicional para el agente.

### **Ventaja diferencial del RAG**

El sistema RAG ofrece mejora continua sin necesidad de ciclos de reentrenamiento largos y costosos:

#### **Comparativa de enfoques:**

Característica	Reentrenamiento tradicional	Sistema RAG (GreenView®)
<b>Tiempo de mejora</b>	1-3 meses (requiere acumular dataset, entrenar y desplegar)	Segundos (disponible inmediatamente tras validación)
<b>Coste computacional</b>	Alto (requiere GPU, tiempo de entrenamiento)	Bajo (solo búsqueda vectorial)
<b>Transparencia</b>	Baja (modelo = caja negra)	Alta (se pueden revisar ejemplos utilizados)
<b>Reversibilidad</b>	Difícil (requiere nueva versión de modelo)	Inmediata (quitar o corregir ejemplos)
<b>Especialización local</b>	Requiere volumen masivo de datos	Aprende desde el primer ejemplo validado

El sistema RAG convierte cada validación en una mejora instantánea. Se hace realidad el principio de **“cuanto más se usa, mejor funciona”** sin necesidad de ciclos de reentrenamiento complejos.

#### **2.5.4. Potencial de reentrenamiento futuro**

Aunque la mejora inmediata se logra mediante RAG, el sistema también registra todas las validaciones humanas de forma estructurada. Esto genera un dataset de alta calidad que podría utilizarse en el futuro para reentrenar los modelos base de forma periódica, optimizando su rendimiento a largo plazo. Cada validación incluye las imágenes originales, la detección de la IA, la decisión del técnico (confirmar, corregir o descartar) y el contexto completo de la incidencia.

Este dataset acumulativo representa un activo valioso que permite tanto auditoría del sistema como posibles mejoras mediante técnicas de machine learning tradicional, complementando así la mejora inmediata que ya proporciona el RAG.

### **3. Resultados del piloto y validación de GreenView®**

Este capítulo presenta los resultados de la validación del sistema con datos reales. Se explican los dos enfoques complementarios de validación: comprobación práctica con avisos ciudadanos reales del Ayuntamiento (donde GreenView® detectó correctamente incidencias reportadas en 2025) y medición técnica rigurosa con métricas estándar sobre más de 1.000 imágenes validadas. Se presentan resultados consolidados comparando tres modelos de IA (GPT-4o, Gemini, Llama-4) y se desglosan resultados específicos para cada caso de uso con ejemplos documentados, justificaciones técnicas de la IA y métricas de

rendimiento. También se analiza la comparativa entre análisis de vistas vs. activos, el sistema de scoring y priorización, y se reconocen transparentemente las limitaciones y aspectos que requieren validación en fases futuras. La conclusión clave es que el éxito se mide en la capacidad de detectar suficientes incidencias críticas para justificar el valor preventivo de la inversión.

Durante el piloto hemos trabajado para medir y validar cómo funciona GreenView®. Los resultados nos muestran que el sistema puede detectar incidencias con buen nivel de fiabilidad y ayudar a priorizar las actuaciones de forma más eficiente.

### **3.1. ¿Cómo hemos validado el sistema?**

Para validar GreenView® hemos usado dos enfoques complementarios, uno práctico (comprobar si encuentra incidencias reales que ya conocíamos) y otro técnico (medir el rendimiento con métricas estándar). Esta combinación nos ha dado una visión completa de cómo funciona el sistema en condiciones reales.

#### **3.1.1. Comprobación con incidencias reales del Ayuntamiento**

El primer paso fue validar que el sistema detecta problemas reales. Para ello, partimos de **avisos ciudadanos que el Ayuntamiento nos facilitó**, que incluyen incidencias reportadas en 2025 por vecinos de Madrid sobre árboles que tapaban señales, farolas, o tenían mal estado.

**El proceso fue sencillo:**

1. Cogimos el aviso ciudadano con su ubicación
2. Buscamos imágenes de Street View de esa misma ubicación y fecha cercana al aviso
3. Pasamos esas imágenes por GreenView® sin decirle qué esperábamos encontrar
4. Comprobamos si el sistema detectaba la incidencia reportada

Esta validación fue importante porque nos permitió **comprobar con casos reales** que el sistema encuentra lo que sabemos que existe, no solo casos preparados para pruebas.

#### **Ejemplo 1: Señal de paso de peatones oculta por árbol**

**Aviso ciudadano:** 29/01/2025 - “No se ve la señal de paso de peatones, tapada por un árbol”



### Resultado del análisis de GreenView®:

```
{  
  "id_caso_uso": "CU-01",  
  "descripcion": "Interferencia con señal de tráfico",  
  "justificacion": "Las ramas del árbol en la parte central del encuadre ocultan  
aproximadamente un 80% de la señal de tráfico, lo que puede afectar la visibilidad  
para los conductores. La evidencia es clara y nítida en la imagen proporcionada.",  
  "scoring": "CU-01 score_base=15, Obstrucción alta (75-99%): +6 puntos,  
score_final=21.0",  
  "clasificacion": "TP"  
}
```

**GreenView® detectó correctamente la incidencia** que el ciudadano había reportado.

### **Ejemplo 2: Farola oculta por copa de árbol**

**Aviso ciudadano:** 07/01/2025 - “La calle está muy oscura por falta de iluminación, ya que las farolas se ven tapadas por las ramas de los árboles”



### Resultado del análisis de GreenView®:

```
{
  "id_caso_uso": "CU-01",
  "descripcion": "Ocultamiento de farola",
  "justificacion": "Las ramas del árbol en la parte superior derecha de la imagen ocultan aproximadamente un 80% de la farola. La evidencia es clara y se observa en el centro de la imagen.",
  "scoring": "CU-01 score_base=15, Obstrucción alta (75-99%): +6 puntos, score_final=21.0",
  "clasificacion": "TP"
}
```

**GreenView® identificó el problema de iluminación** reportado por el vecino.

### **Ejemplo 3: Interferencia con fachada**

**Aviso ciudadano:** 07/01/2025 - “Poda de árbol solicitada por ciudadano”



### Resultado del análisis de GreenView®:

```
{  
  "id_caso_uso": "CU-02",  
  "descripcion": "Interferencia de arbolado con fachada de edificio",  
  "justificacion": "Las ramas del árbol, ubicadas en la copa y lateral derecho,  
  están en contacto directo con la fachada del edificio, invadiendo balcones y  
  ventanas. Esto es visible en el centro de la imagen. La evidencia es nítida y  
  muestra una invasión clara de la vegetación sobre los elementos arquitectónicos.",  
  "scoring": "CU-02 score_base=9, invaden balcones: +4 puntos, score_final=13.0",  
  "clasificacion": "TP"  
}
```

**El sistema detectó la interferencia** que motivó la solicitud de poda.

De los avisos ciudadanos que comprobamos, GreenView® detectó correctamente la mayoría de las incidencias reportadas. Esto nos dio **confianza operativa**, el sistema encuentra problemas reales, no solo artefactos de laboratorio.

Pero también encontramos casos donde el sistema detectaba incidencias que, al revisar la imagen, considerábamos falsos positivos. Por ejemplo, ramas bajas que técnicamente no cumplen el galibo de 2,20m pero que en la práctica no molestan al paso. Estos casos nos ayudaron a **calibrar los umbrales** y ajustar los agentes.

### 3.1.2. Medición con métricas técnicas

La validación práctica nos dio confianza, pero necesitábamos **medir de forma rigurosa** el rendimiento del sistema. Para ello creamos datasets de validación con un proceso sistemático:

#### Proceso de creación de datasets:

1. **Selección de imágenes:** Elegimos cientos de imágenes de calles de Madrid (tanto de Street View como de capturas propias)
2. **Etiquetado manual:** Para cada imagen, marcamos qué casos de uso esperábamos encontrar basándonos en avisos ciudadanos históricos del Ayuntamiento, y revisión visual de las imágenes
3. **Marcado de casos esperados:** Le dijimos al sistema “en esta imagen esperamos detectar CU-01 (señal tapada)” o “en esta imagen no debería haber incidencias”
4. **Análisis con los agentes:** Pasamos todas las imágenes por GreenView® y comparamos lo que detectó con lo que esperábamos

Usamos las métricas estándar de la industria para medir sistemas de IA:

- **Precisión (Precisión):** De todo lo que el sistema dice que es una incidencia, ¿cuánto lo es realmente?
  - Precisión baja = muchos falsos positivos (detecta cosas que no son)
  - Fórmula:  $Precision = TP / (TP + FP)$
  - Interpretación: Mide la exactitud de las detecciones. Un valor alto implica pocos falsos positivos. Es decir, nos va indicar cuantas veces el sistema nos está indicando incidencias cuando no las hay.
- **Recall (Exhaustividad):** De todas las incidencias reales que existen, ¿cuántas encuentra el sistema?
  - Fórmula:  $Recall = TP / (TP + FN)$
  - Interpretación: Mide la capacidad de no dejar pasar incidencias reales. Un valor alto implica pocos falsos negativos. Todas las que se detecten son incidencias que no comunicará un ciudadano o no generará un problema mayor.
- **F1-Score:** Media armónica de las dos anteriores, indica el equilibrio global
  - Fórmula:  $F1 = 2 \times (Precision \times Recall) / (Precision + Recall)$
  - Interpretación: Indicador único que resume la calidad del agente. Se sitúa entre 0 y 1 (o 0–100 %), donde 1 indica desempeño perfecto.

En el desarrollo del sistema, hemos optado por priorizar que no se pierdan incidencias críticas, incluso si esto genera más falsos positivos. Es preferible revisar 100 alertas de las cuales 50 sean falsas antes que dejar escapar una incidencia crítica de seguridad.

Además, los falsos positivos se gestionan rápidamente, con el proceso Human-in-the-Loop un técnico puede revisar y descartar un falso positivo en poco tiempo. Es operativamente manejable.

### 3.1.3. Protocolo técnico de validación

Para garantizar que las métricas sean fiables y reproducibles, hemos seguido un protocolo sistemático:

#### Creación del dataset de validación:

1. Hemos seleccionado una muestra representativa de imágenes georeferenciadas
2. Hemos categorizado las incidencias según los casos de uso definidos
3. Hemos documentado los casos ambiguos y los criterios de decisión utilizados

#### Proceso de evaluación:

1. Hemos ejecutado los agentes de IA sobre el dataset de validación
2. Hemos comparado automáticamente las detecciones de IA con el etiquetado manual
3. Hemos clasificado los resultados en: Verdaderos Positivos (TP), Falsos Positivos (FP), Verdaderos Negativos (TN), Falsos Negativos (FN)
4. Hemos calculado las métricas usando las fórmulas estándar:
  - $\text{Precisión} = \text{TP}/(\text{TP}+\text{FP})$
  - $\text{Recall} = \text{TP}/(\text{TP}+\text{FN})$
  - $\text{F1} = 2 \times (\text{Precisión} \times \text{Recall}) / (\text{Precisión} + \text{Recall})$

## 3.2. Métricas de rendimiento

### 3.2.1. Importancia del volumen de incidencias detectadas

Para medir el rendimiento del sistema hemos utilizado métricas estándar de clasificación (precisión, recall y F1-score) sobre datasets de imágenes validadas por el Ayuntamiento. Sin embargo, lo más relevante no son los porcentajes absolutos, sino el **volumen de incidencias críticas que el sistema es capaz de detectar**.

El objetivo es identificar un número razonable de incidencias reales que permitan una gestión proactiva efectiva, sin saturar al equipo técnico con un volumen inmanejable de alertas. Un sistema que detecta 100 incidencias críticas al mes con un 60% de precisión es más útil que uno que detecta 10 con un 90%, siempre que el volumen total sea operativamente gestionable.

### 3.2.2. Evolución de las métricas con el entrenamiento de los agentes

Los datos muestran cómo el sistema ha ido mejorando conforme hemos cambiado y entrenado los agentes. Muestra de ello es el caso de uso CU-03 (Arbolado en mal estado), a continuación, se muestra la evolución de las métricas iniciales que obtuvimos con datasets que exclusivamente contenían imágenes con incidencias de arbolado en mal estado que provenían de datos de inspecciones facilitadas por el Ayuntamiento:

Fecha	Agente	Dataset	Recall	Comentario
05-ago-2025	v1.0	22 árboles	27%	Primera versión, detectaba solo 6 de 22
15-ago-2025	v2.0	217 árboles	60%	Mejoras internas, ya detecta 130 de 217
25-sep-2025	v3.5	282 árboles	80%	Versión actual, detecta 228 de 282

Esta mejora de **27% a 80% en recall** demuestra que el sistema aprende y mejora con el entrenamiento continuo basado en las validaciones del Ayuntamiento.

Cabe señalar también que en septiembre trabajamos con un conjunto de imágenes donde sabíamos que existían incidencias, mientras que en la parte final hemos ampliado el análisis a un conjunto más realista de imágenes donde no sabíamos de antemano si contenían incidencias o no. Este enfoque más riguroso redujo naturalmente la precisión (porque hay más imágenes sin incidencias que pueden generar falsos positivos), pero lo importante es que el recall se mantuvo alto, lo que significa que seguimos detectando la mayoría de las incidencias reales que existen.

### 3.2.3. Resultados consolidados

Hemos validado los agentes de IA de GreenView® con imágenes reales capturadas en las calles de Madrid, principalmente partiendo de un listado de incidencias reportadas por ciudadanos que provenían de datos facilitados por el Ayuntamiento. En el siguiente apartado del documento se desglosan los resultados por caso de uso.

Como resultado final se muestra la tabla siguiente que resume los resultados globales de los agentes de IA:

Modelo	F1 Global	Precisión	Recall	TP Total	FP Total	FN Total	Esperadas
<b>GPT-4O</b>	39.26%	28.07%	65.30%	32	82	17	50
<b>GEMINI</b>	42.49%	28.67%	82.00%	41	102	9	50
<b>LLAMA-4</b>	42.31%	40.74%	44.00%	22	32	28	50

**Nota metodológica:** Las validaciones se han realizado sobre muestras representativas de imágenes georeferenciadas.

Es importante contextualizar estos números. En septiembre de 2025 realizamos un primer análisis con GPT-4O que arrojó métricas más altas (Precisión: 74,3%, Recall: 78,9%, F1: 75,4%). Los resultados actuales muestran una precisión más baja (28%) pero mantienen un recall alto (65%), lo que significa que seguimos detectando la mayoría de las incidencias reales que existen y el hecho de tener una precisión más baja no es problemático sabiendo que disponemos del Human-in-the-Loop para descartar los falsos positivos de manera rápida y eficaz, más aun teniendo en cuenta su combinación con el sistema de scoring que nos permite priorizar las alertas.

Respecto a la interpretación de resultados, los resultados nos muestran que los agentes de IA funcionan con un enfoque claro, priorizar la detección sobre la perfección. Los valores de Precisión y Recall nos indican que:

1. **El sistema prioriza no perder incidencias críticas**, aunque genere más alertas para validar
2. **La cobertura de detección es alta**, identificando la mayoría de las incidencias reales presentes en el viario
3. **El balance está optimizado para seguridad**, prefiriendo revisar más alertas antes que perder problemas importantes

#### 3.2.4. Importancia de la medición rigurosa

Creemos que evaluar de forma rigurosa el rendimiento de los agentes de IA es importante por varias razones:

- **Fiabilidad operativa:** Medir de forma sistemática con métricas estándar (Precisión, Recall, F1-Score) permite al Ayuntamiento saber con objetividad cómo funciona el sistema. No basta con decir que la IA “funciona”; necesitamos cuantificar su nivel de acierto para que los técnicos municipales puedan confiar en sus detecciones y ajustar sus procesos de validación.

- **Trazabilidad y transparencia:** Al establecer umbrales de rendimiento mínimos (Precisión >40%, Recall >60%) y medir contra ellos, creamos un marco transparente que puede ser verificado por terceros. Esto es importante para futuros pliegos de contratación, donde será necesario exigir niveles de calidad verificables.

- **Mejora continua:** Las métricas nos dan una línea base cuantitativa que nos ayuda a medir cómo mejora el sistema con el tiempo. El sistema incorpora **dos niveles de mejora continua:** inmediata mediante recuperación de ejemplos (RAG) que aprovecha cada validación en segundos, y periódica mediante reentrenamiento que optimiza los modelos. Esta arquitectura de aprendizaje continuo muestra el valor del proceso de validación humana (human-in-the-loop) para capitalizar el conocimiento (ver sección 2.5.3 para detalles técnicos del sistema RAG).

En cualquier caso, el éxito del sistema no se mide solo por alcanzar un porcentaje alto de precisión en todas las detecciones. Lo que realmente importa es:

5. **El volumen de incidencias críticas detectadas:** Aunque tengamos falsos positivos, si detectamos cientos de incidencias reales de alta prioridad que de otro modo pasarían desapercibidas, el sistema cumple su objetivo.
6. **La gestión eficiente de los falsos positivos:** Con el proceso Human-in-the-Loop, cada técnico puede revisar y descartar un falso positivo en segundos, y además cada validación mejora el sistema inmediatamente (mediante RAG) y alimenta el reentrenamiento futuro.
7. **El cambio de paradigma:** Pasar de un modelo puramente reactivo (actuar solo cuando hay denuncia ciudadana) a uno proactivo (identificar problemas antes de que escalen) aporta un valor estratégico que va más allá de las métricas clásicas de precisión.
8. **La tolerancia operativa a falsos negativos en incidencias menores:** Si el sistema, por ejemplo, no detecta algún alcorque vacío de baja prioridad, no supone un problema grave porque el sistema reactivo tradicional seguirá funcionando. Lo crítico es no perder incidencias de seguridad.

### 3.2.5. Modelos comerciales vs open-source

Durante el piloto se planteó la necesidad de incorporar modelos opensource por la potencial reducción de costes que podíamos obtener. La arquitectura agnóstica de GreenView® permite trabajar con diferentes proveedores de IA. En base a esa incorporación se plantea la necesidad de evaluar la bondad de los modelos opensource en comparación con los comerciales.

#### Modelos comerciales “Pro” (GPT-4o, Gemini 2.5 Pro):

- **Ventajas:** Mayor precisión inmediata, rendimiento más predecible en casos complejos
- **Desventajas:** Coste por inferencia elevado, dependencia de proveedores externos, menor control sobre el modelo
- **Aplicación en GreenView®:** Los utilizamos en la fase de precisión del modelo híbrido, donde buscamos la máxima fiabilidad

#### Modelos open-source (Llama 4, modelos propios):

- **Ventajas:** Coste marginal muy bajo, independencia de proveedores
- **Desventajas:** Requiere mayor esfuerzo de validación, su rendimiento puede ser inferior en casos más complejos
- **Aplicación en GreenView®:** Los utilizamos en la fase de cobertura del modelo híbrido, donde priorizamos el análisis exhaustivo

La arquitectura agnóstica permite una transición progresiva hacia modelos open-source conforme mejoran en precisión. El Ayuntamiento puede empezar trabajando con modelos comerciales para casos críticos y, a medida que desarrolla datasets de entrenamiento propios, ir migrando hacia modelos open-source alojados en infraestructura propia, manteniendo o mejorando la calidad del sistema.

A fecha de redacción de este documento, hemos realizado pruebas comparativas con diferentes modelos open-source (como Llama 4) y modelos comerciales Pro (GPT-4o, Gemini 2.5 Pro).

Respecto a la tendencia de mejora de modelos open-source, vemos que estos modelos están mejorando rápidamente. En los últimos 12-18 meses hemos visto cómo modelos que antes solo estaban disponibles comercialmente se han liberado públicamente (por ejemplo, Grok de X.AI). Es razonable esperar que esta tendencia continúe, y que en 1-2 años los modelos open-source alcancen niveles de rendimiento comparables a los actuales modelos Pro.

En esta tabla vemos una comparación de indicadores de dos modelos, uno comercial y otro open-source.

<i>Indicador (few-shot)</i>	<i>GPT-4o</i>	<i>Llama 3.1 405B (open)</i>
<b>MMLU (5-shot)</b>	88.7%	88.6%
<b>MATH</b>	76.6%	73.8%
<b>HumanEval (0-shot)</b>	90.2%	89.0%

- **MMLU (5-shot):** Prueba general de conocimiento en 57 áreas. “5-shot” significa que el modelo ve 5 ejemplos antes de responder.
- **MATH:** Problemas de matemáticas de varios niveles de dificultad. Mide razonamiento matemático.
- **HumanEval (0-shot):** Retos de programación en Python. “0-shot” significa sin ejemplos previos.

**Fuente de datos:** Vellum.ai (resumen comparado): GPT-4o y Llama 3.1 405B en MMLU, MATH y HumanEval. <https://www.vellum.ai/blog/llm-benchmarks-overview-limits-and-model-comparison>

Para entender hacia dónde vamos, es útil mirar hacia atrás. Hace solo un par de años, Llama 2 70B se acercaba al rendimiento de GPT-3.5 en pruebas estándar. Hoy en día, Llama 3.1 405B prácticamente iguala a GPT-4o: 88.6% frente a 88.7% en MMLU, por ejemplo. La diferencia es mínima.

Esta convergencia es rápida y constante. Si mantenemos este ritmo, es probable que en 1-2 años los modelos open-source de nueva generación alcancen o superen el rendimiento de los modelos comerciales actuales en muchas tareas. No es una certeza, pero la tendencia está ahí.

Para el Ayuntamiento, esto es una buena noticia, significa que la estrategia de usar una arquitectura flexible no solo da independencia tecnológica hoy, sino que también abre la puerta a optimizar costes en el futuro sin perder capacidad.

### 3.2.6. Modelos open-source en detección sobre activos

En el modo de detección sobre activos, los modelos open-source actuales generan más falsos positivos que los comerciales Pro. Sin embargo:

- Este gap de rendimiento es mitigable: el proceso Human-in-the-Loop permite validar rápidamente estos falsos positivos
- Los modelos mejoran constantemente: cada nueva versión reduce el gap
- El sistema aprende: cada validación humana mejora la precisión de forma inmediata (RAG) y alimenta futuros entrenamientos

### 3.2.7. Proyección estratégica de los resultados

El sistema GreenView es cada vez más preciso por dos vías:

- **Mejora de los modelos base:** Las nuevas versiones de modelos open-source se pueden incorporar fácilmente gracias a la arquitectura agnóstica.
- **Aprendizaje continuo del sistema:** El conocimiento acumulado mediante las validaciones humanas se capitaliza y reutiliza.

Esto significa que el Ayuntamiento puede empezar con un mix de modelos comerciales y open-source, e ir migrando progresivamente hacia open-source conforme estos mejoran, optimizando recursos sin sacrificar calidad.

**Nota:** Nos hubiera gustado realizar pruebas más extensas con un dataset más amplio para obtener métricas aún más robustas de esta comparativa. Los datos actuales son representativos pero se beneficiarían de una validación a mayor escala.

### **3.3. Resultados por caso de uso**

Ahora vamos a desglosar los resultados caso por caso, mostrando ejemplos reales de detecciones, métricas específicas y las lecciones aprendidas en cada tipo de incidencia.

#### **3.3.1. CU-01: Interferencias con señales de tráfico**

##### **Qué detecta:**

El sistema identifica árboles cuyas ramas o follaje obstruyen señales de tráfico, semáforos o farolas, comprometiendo la seguridad vial. Es un caso de uso de **alta prioridad** por su impacto directo en la seguridad.

##### **Resultados obtenidos:**

<i>Modelo</i>	<i>F1</i>	<i>Precisión</i>	<i>Recall</i>	<i>TP</i>
<b>GEMINI</b>	40.00%	26.67%	80.00%	12/15
<b>GPT-4O</b>	40.82%	28.57%	71.43%	10/14
<b>LLAMA-4</b>	38.46%	45.45%	33.33%	5/15

Gemini ofrece el mejor recall (80%), lo que significa que detecta 8 de cada 10 casos reales. La precisión es moderada (~27%), pero preferimos esta configuración porque los falsos positivos se validan rápidamente.

##### **Ejemplos documentados:**

##### **Ejemplo 1: Señal de paso de peatones obstruida (TP)**

**Aviso ciudadano:** 29/01/2025 - “No se ve la señal de paso de peatones, tapada por un árbol, aligustre”



```
{  
  "id_caso_uso": "CU-01",  
  "descripcion": "Interferencia con señal de tráfico",  
  "justificacion": "Las ramas del árbol en la parte central del encuadre ocultan aproximadamente un 80% de la señal de tráfico, lo que puede afectar la visibilidad para los conductores. La evidencia es clara y nítida en la imagen proporcionada.",  
  "scoring": "CU-01 score_base=15, Obstrucción alta (75-99%): +6 puntos, score_final=21.0",  
  "clasificacion": "TP"  
}
```

**Ejemplo 2: Farola oculta por copa de árbol (TP)**

**Aviso ciudadano:** 07/01/2025 - “La calle está muy oscura por falta de iluminación, ya que las farolas se ven tapadas por las ramas de los árboles”



```
{
  "id_caso_uso": "CU-01",
  "descripcion": "Ocultamiento de punto de luz (farola) por la copa del árbol",
  "justificacion": "Las ramas del árbol en la parte superior derecha de la imagen ocultan aproximadamente un 80% de la farola. La evidencia es clara y se observa en el centro de la imagen.",
  "scoring": "CU-01 score_base=15, Obstrucción alta (75-99%): +6 puntos, score_final=21.0",
  "clasificacion": "TP"
}
```

### **Ejemplo 3: Semáforo parcialmente oculto (TP)**

**Aviso ciudadano: 05/05/2025 - "Poda solicitada por ciudadano"**



```
{  
  "id_caso_uso": "CU-01",  
  "descripcion": "Vegetación parcialmente ocultando un semáforo en la alineación de la calle",  
  "justificacion": "Semáforo en el centro izquierdo del encuadre, a media distancia, sobre la calzada. Presenta ocultamiento aproximado del 80% por ramas y follaje del árbol adyacente. El foco luminoso es visible, pero la obstrucción reduce la claridad de la señal para conductores que se aproximan. Ubicación crítica en cruce con paso de peatones.",  
  "scoring": "CU-01 score_base=15, Obstrucción alta (75-99%): +6 puntos, score_final=21.0",  
  "clasificacion": "TP"  
}
```

#### **Ejemplo 4: Falso positivo por perspectiva 2D**



```
{
  "id_caso_uso": "CU-01",
  "descripcion": "Obstrucción de farola por vegetación",
  "justificacion": "Las ramas en la copa del árbol, ubicadas en el centro de la imagen, cubren aproximadamente un 50% de la farola, lo que puede reducir la iluminación nocturna. Comparado con otros árboles en la imagen, este presenta una interferencia más significativa.",
  "scoring": "CU-01 score_base=15, Obstrucción moderada (50-74%): +3 puntos, score_final=18.0",
  "clasificacion": "FP (posible perspectiva 2D)"
}
```

**Nota:** Este caso ilustra el riesgo de perspectiva 2D. En la imagen parece que el árbol tapa la farola, pero en el espacio 3D real podría estar a varios metros de distancia. La validación humana descarta estos casos rápidamente.

#### **Lo que funciona bien:**

El sistema detecta correctamente cuando una señal está obstruida en un 30% o más. El scoring prioriza automáticamente los casos más graves y funciona tanto con el análisis de activos como con las vistas panorámicas.

#### **Los retos:**

El principal problema es la perspectiva 2D. A veces, en la imagen parece que un árbol tapa una señal, pero en la realidad están separados varios metros. También hay cierta imprecisión al estimar porcentajes de obstrucción, y las señales pequeñas o con poco contraste son más difíciles de detectar.

### 3.3.2. CU-02: Interferencias con fachadas

#### Qué detecta:

El sistema identifica árboles cuyas ramas invaden el espacio de edificios, tocando o entrando en balcones, ventanas, toldos o cualquier elemento de la fachada. Es un caso de uso de **prioridad media**, importante para prevenir daños materiales y quejas vecinales.

#### Resultados obtenidos:

Modelo	F1	Precisión	Recall	TP
<b>LLAMA-4</b>	41.18%	33.33%	53.85%	7/13
<b>GEMINI</b>	35.62%	21.67%	100.00%	13/13
<b>GPT-4O</b>	27.91%	20.00%	46.15%	6/13

Gemini tiene recall perfecto (100%), detecta todos los casos reales, aunque genera más falsos positivos. Llama-4 ofrece mejor equilibrio con F1 del 41%.

#### Ejemplos documentados:

##### **Ejemplo 1: Invasión severa de balcones (TP)**

**Aviso ciudadano:** 07/01/2025 - “Poda de árbol solicitada por ciudadano”



```

{
  "id_caso_uso": "CU-02",
  "descripcion": "Interferencia de arbolado con fachada de edificio",
  "justificacion": "Las ramas del árbol, ubicadas en la copa y lateral derecho, están en contacto directo con la fachada del edificio, invadiendo balcones y ventanas. Esto es visible en el centro de la imagen. La evidencia es nítida y muestra una invasión clara de la vegetación sobre los elementos arquitectónicos.",
  "scoring": "CU-02 score_base=9, invaden balcones: +4 puntos, score_final=13.0",
  "clasificacion": "TP"
}

```

**Ejemplo 2: Contacto con fachada y balcones (TP)**

**Aviso ciudadano: 02/01/2025 - “Poda de árbol solicitada por ciudadano”**



```

{
  "id_caso_uso": "CU-02",
  "descripcion": "Interferencia con fachada por contacto de ramas con balcones",
  "justificacion": "Las ramas en la parte media de la copa entran en contacto con la fachada del edificio colindante, invadiendo el espacio de balcones en plantas intermedias. La interferencia es visible y directa.",
  "scoring": "CU-02 score_base=9, contacto con fachada: +4 puntos, score_final=13.0",
  "clasificacion": "TP"
}

```

### Lo que funciona bien:

Gemini detecta el 100% de los casos reales. Cuando hay contacto directo visible entre el árbol y la fachada, la detección es sólida. Además, el scoring diferencia bien entre ramas que están simplemente cerca y ramas que realmente tocan el edificio.

### Los retos:

Aquí también aparece el problema de la perspectiva 2D: a veces parece que hay contacto cuando en realidad hay separación. Además, el sistema puede marcar como incidencia ramas que están muy cerca pero sin tocar realmente la fachada. La pregunta es: ¿a partir de cuántos centímetros de proximidad consideramos que hay problema? Esto requiere afinarlo con el Ayuntamiento.

Para gestionar estos casos, el scoring no trata todos los contactos como críticos, hay gradación según la severidad. Y los falsos positivos por proximidad sin contacto real se pueden descartar rápidamente en la validación.

### 3.3.3. CU-03: Arbolado en mal estado

#### Qué detecta:

El sistema identifica árboles con signos de deterioro, copa seca, ramas quebradizas, inclinación peligrosa, síntomas de enfermedad o plagas. Es un caso de uso de **alta prioridad** por el riesgo de caída.

#### Resultados obtenidos:

Modelo	F1	Precisión	Recall	TP
<b>GPT-4O</b>	75.00%	64.29%	90.00%	9/10
<b>LLAMA-4</b>	61.54%	50.00%	80.00%	8/10
<b>GEMINI</b>	58.83%	41.67%	100.00%	10/10

**Este es el caso de uso con mejores resultados.** GPT-4O alcanza 75% de F1 y 90% de recall. Es el caso más maduro del sistema.

#### Ejemplos documentados:

**Ejemplo 1: Árbol con copa >70% seca (TP)**



```
{
  "id_caso_uso": "CU-03",
  "descripcion": "Árbol con más del 70% de la copa seca y ramas quebradizas",
  "justificacion": "El árbol ubicado en el centro de la imagen presenta una
desfoliación severa, aproximadamente un 80% de la copa carece de hojas. Las ramas
muestran signos de debilidad estructural con segmentos quebradizos visibles. La
condición del árbol indica un estado avanzado de deterioro que podría comprometer
su estabilidad.",
  "scoring": "CU-03 score_base=12, copa seca >70%: +14 puntos, score_final=26.0",
  "clasificacion": "TP"
}
```

**Ejemplo 2: Árbol completamente seco (TP)**



```
{  
  "id_caso_uso": "CU-03",  
  "descripcion": "Árbol en estado crítico, prácticamente sin follaje",  
  "justificacion": "El árbol carece de hojas, con ramas completamente desnudas y quebradizas. No se observan brotes ni signos de vida vegetativa. El estado es incompatible con la época del año (imagen de julio), indicando que el árbol está muerto o en estado crítico terminal.",  
  "scoring": "CU-03 score_base=12, sin follaje: +14 puntos, score_final=26.0",  
  "clasificacion": "TP"  
}
```

El sistema detecta árboles muertos correctamente

### **Ejemplo 3: Árbol con ramas rotas y defoliación (TP)**



```
{
  "id_caso_uso": "CU-03",
  "descripcion": "Árbol con debilidad estructural severa",
  "justificacion": "Se observan ramas quebradizas colgando sin brotes. La copa presenta defoliación aproximada del 60-70%. Evidencia de debilidad estructural que representa riesgo de desprendimiento.",
  "scoring": "CU-03 score_base=12, defoliación >60%: +10 puntos, ramas quebradizas: +4 puntos, score_final=26.0",
  "clasificacion": "TP"
}
```

#### **Lo que funciona bien:**

Este es el caso de uso con mejores resultados del sistema. GPT-4O alcanza un F1 score del 75% y detecta el 90% de los casos reales. El sistema distingue bien entre un árbol seco y uno de hoja caduca, y tiene en cuenta la época del año: si una imagen es de julio y el árbol no tiene hojas, eso es señal de problema.

#### **Los retos:**

El principal desafío es la estacionalidad. En otoño, hay que tener cuidado de no confundir árboles sanos con árboles secos. También puede haber falsos positivos con árboles recién plantados que parecen débiles pero están perfectamente sanos. Y estimar el porcentaje exacto de defoliación desde una foto 2D no siempre es preciso.

Para minimizar estos problemas, el sistema considera la fecha de captura, compara con imágenes anteriores del mismo árbol cuando las hay disponibles, y no se basa solo en el follaje: también analiza ramas rotas, inclinación, grietas y otros indicadores de debilidad estructural

### 3.3.4. CU-05: Alcorques vacíos

#### Qué detecta:

El sistema identifica alcorques sin árbol o con tocones, que representan oportunidades de plantación o necesidad de limpieza/destoconado.

#### Resultados obtenidos:

Este caso de uso aún está en **validación limitada**. Hemos probado detecciones en condiciones favorables (alcorques visibles) con buenos resultados preliminares.

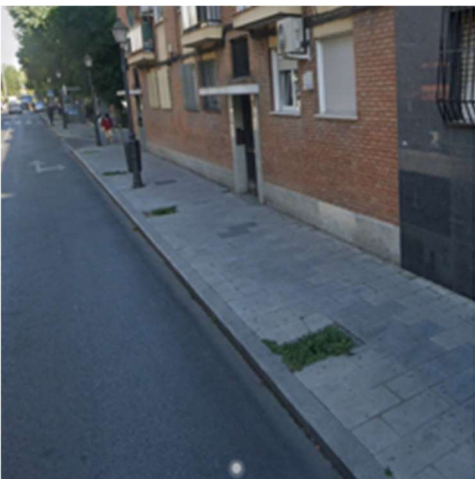
#### Ejemplos documentados:

##### Ejemplo 1: Alcorque vacío con riesgo de tropiezo (TP)



```
{  
  "id caso uso": "CU-05",  
  "descripcion": "Alcorque vacío sin árbol, con riesgo de tropiezo",  
  "justificacion": "Se observa un alcorque delimitado sin árbol vivo en su interior. El hueco presenta bordes elevados que podrían representar riesgo de tropiezo para peatones. El alcorque está limpio, listo para plantación.",  
  "scoring": "CU-05 score base=5, riesgo tropiezo: +3 puntos, score final=8.0",  
  "clasificacion": "TP"  
}
```

##### Ejemplo 2: Alcorque vacío en acera (TP)



```
{  
  "id caso uso": "CU-05",  
  "descripcion": "Alcorque sin árbol identificado en la acera",  
  "justificacion": "Alcorque visible en la acera sin presencia de árbol vivo. El espacio está disponible para nueva plantación. No se observan restos de tocón ni basura.",  
  "scoring": "CU-05 score base=5, score final=5.0",  
  "clasificacion": "TP"  
}
```

#### Los retos:

Este caso de uso tiene un problema crítico: los coches aparcados tapan los alcorques en la mayoría de las capturas. Además, los alcorques son objetos pequeños difíciles de detectar en imágenes panorámicas, y desde el vehículo de captura el ángulo no siempre permite ver el interior.

Para mejorar la cobertura, hemos identificado varias estrategias: capturar en diferentes horarios (madrugada, sábados, festivos) cuando hay menos coches aparcados, o planificar rutas en agosto cuando la ocupación de aparcamiento es menor. En zonas críticas, se podría usar captura peatonal con mochila 360°. También el modelo de vistas puede ayudar, ya que detecta alcorques incluso cuando están parcialmente visibles.

El siguiente paso es validar con un dataset más amplio y probar cuál funciona mejor: el modelo de vistas o el de activos.

### 3.3.5. CU-06: Obstrucción de vías peatonales

#### Qué detecta:

El sistema identifica ramas bajas que invaden el espacio peatonal, reduciendo el gálibo (altura libre) por debajo de umbrales de accesibilidad o el ancho útil de la acera.

#### Resultados obtenidos:

Modelo	F1	Precisión	Recall	TP
<b>GEMINI</b>	45.45%	45.45%	45.45%	5/11
<b>GPT-4O</b>	27.27%	18.18%	54.55%	6/11
<b>LLAMA-4</b>	25.00%	40.00%	18.18%	2/11

Este caso de uso está **en proceso de recalibración**. Genera un número elevado de falsos positivos porque los umbrales iniciales (2,20m) son muy exigentes.

#### Ejemplos documentados:

### **Ejemplo 1: Obstrucción severa de acera (TP)**



```
{
  "id_caso_uso": "CU-06",
  "descripcion": "Obstrucción severa de la acera por ramas bajas y vegetación densa",
  "justificacion": "Las ramas del árbol están a una altura estimada de 1.80 m sobre la acera, obstaculizando el paso de peatones. Además, la vegetación densa reduce el ancho útil de la acera a aproximadamente 1.0 m. La obstrucción es visible en el centro de la imagen.",
  "scoring": "CU-06 score_base=7, altura <1.8m: +3 puntos, ancho <1.0m: +4 puntos, score_final=14.0",
  "clasificacion": "TP"
}
```

## **Ejemplo 2: Falso positivo por umbral demasiado estricto (FP)**



```
{  
  "id caso uso": "CU-06",  
  "descripcion": "Obstrucción parcial de la acera por ramas bajas",  
  "justificacion": "Existen ramas a una altura estimada de 1.80 m sobre la acera,  
  lo que puede obstaculizar el paso de peatones. La evidencia es parcial pero sugiere  
  la necesidad de intervención.",  
  "scoring": "CU-06 score base=7, altura <2.2m: +3 puntos, score final=10.0",  
  "clasificacion": "FP"  
}
```

**Nota:** En esta imagen las ramas técnicamente están por debajo de 2,20m, pero en la práctica no molestan al paso. Este tipo de FP es común con el umbral actual.

### **Lo que funciona bien:**

Cuando la obstrucción es clara y severa (ramas por debajo de 1,80m o que reducen significativamente el ancho útil de la acera), el sistema la detecta correctamente. GPT-4O alcanza un 54% de recall, lo que significa que encuentra más de la mitad de los casos reales. El scoring considera múltiples factores, no solo la altura de las ramas: también tiene en cuenta si la vegetación reduce el ancho de paso, lo que da una valoración más completa del problema.

### **Los retos:**

El umbral de 2,20m resulta demasiado estricto y genera miles de detecciones de ramas que técnicamente no cumplen el gálibo pero que en la práctica no molestan al paso. Además, medir alturas con precisión desde una imagen 2D es complicado. Y no toda rama baja es un problema: depende del contexto, del ancho de la acera, del tránsito peatonal.

La idea es ajustar el umbral a alturas más realistas (2,0m o 1,9m), incorporar el contexto del ancho de acera (una rama baja en una acera de 3 metros es menos crítica que en una de 1,5m), y priorizar por ubicación: zonas escolares, hospitales y paradas de transporte deberían tener umbrales más exigentes.

### 3.3.6. CU-08: Nidos de gran tamaño (cotorras)

#### Qué detecta:

El sistema identifica nidos de gran tamaño en la copa de los árboles, típicamente construidos por cotorras argentinas, que pueden representar riesgo de desprendimiento o necesidad de gestión de fauna.

#### Resultados obtenidos:

Modelo	F1	Precisión	Recall	TP
<b>GPT-4O</b>	66.67%	50.00%	100.00%	1/1
<b>GEMINI</b>	50.00%	33.33%	100.00%	1/1
<b>LLAMA-4</b>	0.00%	0.00%	0.00%	0/1

**Dataset muy pequeño** (solo 1-2 casos validados), pero los resultados son prometedores. GPT-4O y Gemini detectan el 100% de los casos. Llama-4 falla completamente en este caso de uso.

#### Ejemplos documentados:

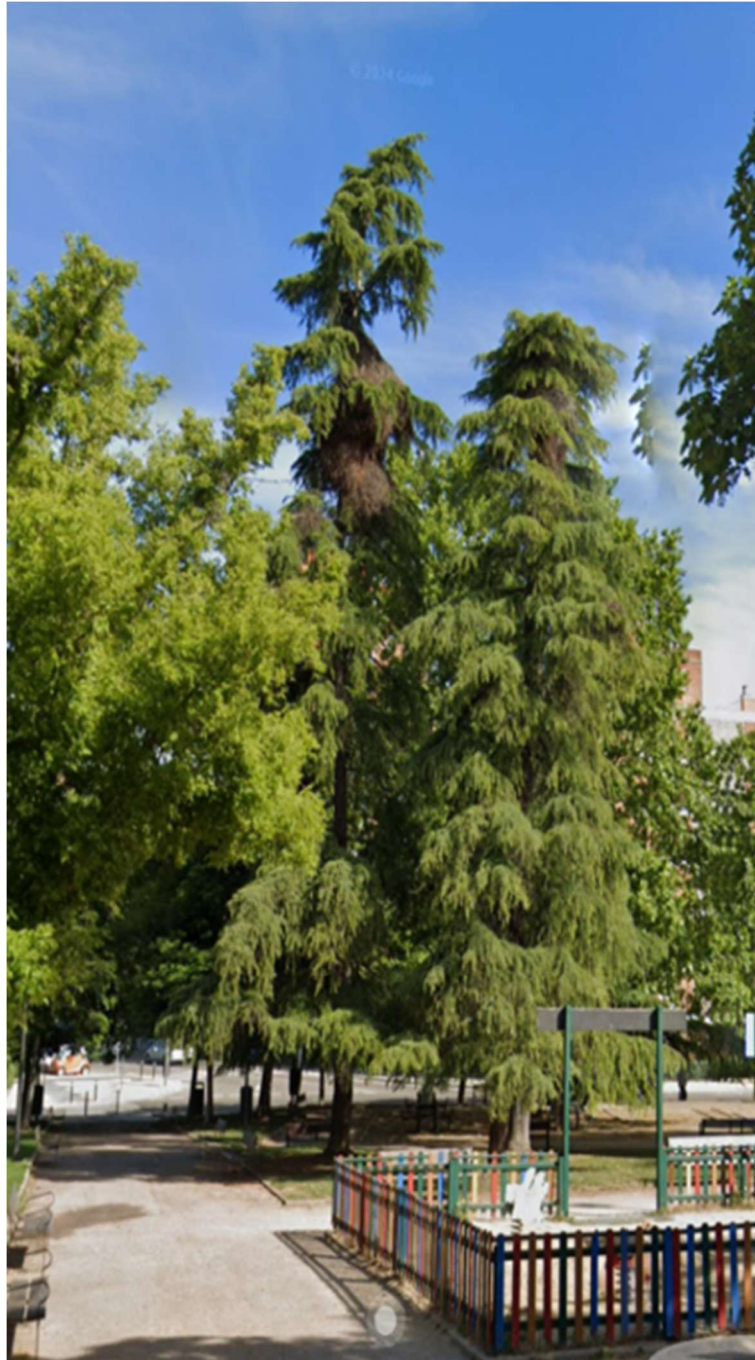
**Ejemplo 1: Nido grande en copa de árbol (TP)**

**Aviso ciudadano: 04/07/2025 - "2 nidos de cotorras en el árbol"**



```
{  
  "id caso uso": "CU-08",  
  "descripcion": "Nido de gran tamaño en la copa del árbol",  
  "justificacion": "Se observa un nido de gran tamaño en la copa superior del árbol, con una estructura esférica de aproximadamente 80 cm de diámetro. La ubicación del nido es en el centro del encuadre y es claramente visible.",  
  "scoring": "CU-08 score base=6, riesgo desprendimiento: +4 puntos, score final=10.0",  
  "clasificacion": "TP"  
}
```

**Ejemplo 2: Nido activo con ramas frescas (TP)**



```
{  
  "id_caso_uso": "CU-08",  
  "descripcion": "Nido de gran tamaño observado en la copa del árbol, posiblemente  
de cotorras",  
  "justificacion": "Se observa una estructura esférica de aproximadamente 80 cm de  
diámetro en la copa superior del árbol, construida con ramas. La estructura parece  
activa, ya que se observan ramas frescas. La ubicación del nido es en el centro del  
encuadre y es claramente visible.",  
  "scoring": "CU-08 score_base=6, riesgo desprendimiento: +4 puntos,  
score_final=10.0",  
  "clasificacion": "TP"  
}
```

**Lo que funciona bien:**

GPT-4O y Gemini detectan correctamente los nidos grandes y visibles. El sistema puede estimar el tamaño aproximado e incluso identificar si el nido parece activo observando si hay ramas frescas.

#### **Los retos:**

El dataset es muy limitado: solo hemos validado con 1-2 casos. Llama-4 falla completamente en este caso de uso y necesita más entrenamiento específico. Además, en verano con follaje denso, los nidos pueden quedar ocultos entre las hojas.

Para mejorar este caso de uso, el siguiente paso es ampliar el dataset de validación con más casos, entrenar específicamente con más ejemplos, y evaluar si la captura en invierno (sin hojas) mejora la detección.

#### **3.3.7. CU-09: Tocones (Exploratorio)**

**Estado:** Caso exploratorio, validado solo con modelo de vistas

#### **Qué detecta:**

Tocones de árboles talados que permanecen en alcorques y necesitan ser retirados.

#### **El descubrimiento:**

Este caso de uso funciona mejor con el modelo de vistas panorámicas que con el de activos individuales. Los tocones son elementos pequeños que necesitan contexto visual amplio para ser identificados correctamente: ver el alcorque completo y su entorno ayuda más que enfocarse solo en el árbol.

El siguiente paso es validar con un dataset más amplio, determinar si vale la pena desarrollar un agente específico o si el modelo de vistas es suficiente, y evaluar la prioridad operativa: ¿es crítico detectar tocones o se puede gestionar de forma reactiva?

#### **3.3.8. CU-10: Desbroce/escarda (Exploratorio)**

**Estado:** Caso exploratorio, validado solo con modelo de vistas

#### **Qué detecta:**

Vegetación espontánea excesiva en alcorques (más de 20cm de altura o más del 30% de cobertura) que requiere limpieza.

#### **El descubrimiento:**

Como en el caso anterior, funciona mejor con el modelo de vistas porque necesita contexto visual amplio para evaluar la densidad y altura de la vegetación en el alcorque.

Los próximos pasos son validar con más casos, calibrar umbrales con el Ayuntamiento (¿cuándo consideramos que la vegetación es “excesiva?”) y evaluar si tiene sentido combinar este caso de uso con el CU-05 (alcorques vacíos) en un único agente de gestión de alcorques.

#### **3.3.9. CU-12: Papeleras llenas (Exploratorio)**

**Estado:** Caso exploratorio, no validado extensivamente

**Qué detecta:**

Papeleras con llenado >70%, residuos sobresaliendo o residuos dispersos alrededor, indicando necesidad de vaciado.

**Lo que hemos probado:**

Hemos realizado pruebas preliminares con imágenes cercanas de papeleras. El sistema puede identificar el nivel de llenado estimado, detectar si hay residuos sobresaliendo de la boca, y ver si hay residuos dispersos en el suelo alrededor.

**Ejemplo:**



```
{  
  "id caso uso": "CU-12",  
  "descripcion": "Papeleras con llenado superior al 70%, residuos visibles sobresaliendo",  
  "justificacion": "La papeleras muestra un llenado estimado del 75%, con residuos sobresaliendo de la boca. No se observan residuos dispersos alrededor. La imagen fue capturada en un área pública.",  
  "scoring": "CU-12 score base=5, llenado >70%: +1 punto, score final=6.0",  
  "clasificacion": "TP"  
}
```

**Los retos:**

El problema principal es que las pruebas se hicieron con fotos cercanas que no se parecen a las capturas desde el vehículo. Desde el coche en movimiento, las papeleras quedan muy lejos y pequeñas en la imagen. Además, coches, peatones y mobiliario urbano las tapan frecuentemente.

Este caso de uso necesita un modelo de vistas con capacidad para detectar objetos pequeños, o bien captura específica a menor velocidad enfocada en mobiliario urbano. Los próximos pasos serían hacer una prueba piloto con captura a baja velocidad (10-15 km/h) o con paradas en zonas con muchas papeleras, conseguir ejemplos realistas capturados desde el vehículo, evaluar si el coste de captura específica justifica el beneficio, y coordinarse con el servicio de limpieza para definir qué umbrales de llenado son realmente relevantes.

### **3.4. Comparativa: Análisis de vistas vs análisis de activos**

GreenView® utiliza dos estrategias complementarias de análisis: **detección sobre vistas panorámicas** (análisis contextual de toda la imagen) y **detección sobre activos individuales** (análisis centrado en cada elemento georeferenciado). Cada enfoque tiene fortalezas y limitaciones diferentes.

#### **Las vistas tienen falsos negativos mayoritariamente de bajo riesgo (75%)**

Lo que se escapa en el análisis de vistas son principalmente elementos pequeños o poco visibles en la imagen panorámica: pequeñas papeleras, tocones de baja altura, alcorques vacíos difíciles de distinguir por el ángulo de captura. Tiene un **8% de falsos negativos de alto riesgo** que corresponde principalmente a elementos que directamente no aparecen en la imagen, bien sea por el ángulo de visión o por condiciones de iluminación desfavorables.

#### **Los activos tienen un porcentaje mucho mayor de falsos negativos de alto riesgo (40%)**

Esto ocurre porque el análisis sobre activos **no puede detectar incidencias de contexto urbano**. Si un árbol interfiere con una señal de STOP pero el análisis se centra solo en el árbol sin ver la señal, la incidencia crítica se pierde. También se pierden interferencias con fachadas, obstrucciones de vías, etc.

No obstante, esto se compensa con el análisis de primer nivel del agente de vistas, ya que tiene un mayor contexto de la imagen panorámica y puede identificar estas relaciones espaciales entre elementos.

#### **Ejemplo comparativo: Mismo punto, diferentes enfoques**

Para ilustrar las diferencias, veamos un mismo punto analizado con ambos enfoques:

##### **Análisis de vista panorámica:**



*Vista panorámica*

```

{
  "casos_uso_detectados": ["CU-02", "CU-09", "CU-10", "CU-12"],
  "total_incidencias": 4,
  "detecciones": [
    {
      "id_caso_uso": "CU-02",
      "descripcion": "Interferencia de arbolado con fachada de edificio",
      "justificacion": "Las ramas del árbol interfieren con el edificio contiguo. Se observa contacto directo con la fachada de ladrillo y las ventanas de las plantas superiores."
    },
    {
      "id_caso_uso": "CU-09",
      "descripcion": "Presencia de tocón en alcorque",
      "justificacion": "Se localiza un tocón de un árbol talado en un alcorque en la acera opuesta. El diámetro estimado es de 30-40 cm."
    },
    {
      "id_caso_uso": "CU-10",
      "descripcion": "Necesidad de desbroce/escarda en alcorque con tocón",
      "justificacion": "En el mismo alcorque ha crecido vegetación espontánea que supera los 20 cm de altura, cubriendo aproximadamente el 30% del alcorque."
    },
    {
      "id_caso_uso": "CU-12",
      "descripcion": "Residuos abandonados en vía pública",
      "justificacion": "Se observan dos bolsas de basura de gran tamaño depositadas en el suelo junto al bordillo."
    }
  ]
}

```

**El modelo de vistas detecta 4 incidencias** gracias al contexto amplio de la escena

**Análisis del activo individual (mismo punto):**

```

{
  "casos_uso_detectados": ["CU-01", "CU-02"],
  "total_incidencias": 2,
  "detecciones": [
    {
      "id_caso_uso": "CU-01",
      "descripcion": "Ocultamiento de punto de luz (farola) por la copa del árbol",
      "justificacion": "El denso follaje cubre por completo la luminaria. Se estima una reducción de la iluminación efectiva superior al 80%."
    },
    {
      "id_caso_uso": "CU-02",
      "descripcion": "Invasión severa de fachada y balcones",
      "justificacion": "El follaje invade el espacio aéreo de al menos dos balcones en los pisos superiores y está en contacto directo con la pared."
    }
  ]
}

```

**El modelo de activos detecta 2 incidencias** enfocándose en el árbol específico

**Conclusión: Son complementarios**

- **Vistas:** Detectan tocones, alcorques, papeleras, residuos (contexto urbano amplio). Pierden detalles específicos de activos individuales.
- **Activos:** Detectan interferencias específicas del árbol con precisión. Pierden elementos pequeños que no son el activo analizado.

El sistema usa **ambos enfoques en paralelo** para maximizar la cobertura. Lo que uno pierde, el otro probablemente lo detecta.

### **3.5. Sistema de scoring y priorización**

Detectar incidencias está bien, pero lo realmente útil es saber cuáles atender primero. GreenView® incorpora un sistema de priorización automática que asigna una puntuación a cada incidencia detectada. Esto permite ordenar las actuaciones por criticidad y convertir miles de detecciones en una lista manejable y priorizada.

**Sobre el estado actual del sistema:** El esquema de puntuación que presentamos aquí es un modelo preliminar configurado durante el piloto. Las ponderaciones, umbrales y categorías de criticidad todavía están pendientes de calibración con el Ayuntamiento. Antes del despliegue operativo completo, el Área de Zonas Verdes aprobará formalmente las categorías finales y sus rangos. El sistema está diseñado para ser completamente configurable y poder ajustarse según la experiencia real.

#### **Sobre los falsos positivos:**

No todos los casos de uso generan el mismo nivel de falsos positivos. Algunos tienen criterios más amplios y, por tanto, detectan más cosas:

**CU-02 (Interferencias con fachadas)** y **CU-06 (Obstrucción de vías)** generan más detecciones porque el criterio de proximidad es más amplio. Esto produce más falsos positivos, pero son casos que se validan rápidamente (un técnico descarta en segundos una rama que no toca realmente una fachada), no son críticos para la seguridad inmediata (por eso el scoring les asigna prioridad media-baja), y se pueden mitigar ajustando umbrales o mediante el aprendizaje continuo del sistema.

**CU-01 (Interferencias con señales)** y **CU-03 (Arbolado en mal estado)** tienen criterios más restrictivos y generan menos falsos positivos, pero detectan situaciones de mayor criticidad.

Esta diferencia es normal y manejable. El scoring automático filtra por criticidad real, permitiendo que los técnicos se concentren primero en validar las incidencias más importantes.

#### **3.5.1. Configurabilidad según prioridades municipales**

**Nota sobre implementación:** El sistema de configuración descrito en esta sección forma parte de la hoja de ruta de desarrollo del producto. Actualmente, las configuraciones de scoring se realizan a nivel técnico con soporte del equipo de SIGO. El objetivo es que el Ayuntamiento disponga de estas herramientas de configuración autónoma en fases posteriores de desarrollo.

El sistema de scoring de GreenView® es completamente configurable, permitiendo al Ayuntamiento adaptar los criterios de priorización a sus políticas y prioridades específicas:

#### **Capacidades de configuración:**

- **Ajuste de ponderaciones:** Los técnicos municipales podrán modificar los puntos asignados a cada factor según la importancia relativa que deseen dar a cada criterio.

- **Definición de umbrales de criticidad:** Los rangos de puntuación que definen las categorías (Crítica, Alta, Media, Baja) podrán ser personalizados.

### 3.5.2. Proceso de co-diseño del scoring

El sistema de scoring de GreenView® se ha diseñado desde el principio como una herramienta colaborativa que se configura y mantiene conjuntamente entre el proveedor tecnológico y el Ayuntamiento. Esto garantiza que las prioridades municipales se reflejen fielmente en el sistema y que el Ayuntamiento mantiene control total sobre la evolución de los criterios.

#### Cómo se ha diseñado el scoring

##### Definición de criterios iniciales

El esquema preliminar se estableció mediante reuniones con técnicos de Zonas Verdes para identificar qué factores determinan la urgencia real de una incidencia. Comparamos qué tipos de incidencias son más críticas (por ejemplo, ¿es más urgente una obstrucción de señal de STOP o un árbol con síntomas de enfermedad?). A partir de ahí, determinamos rangos de puntuación para las categorías de criticidad (Crítica, Alta, Media, Baja) y revisamos incidencias detectadas para ajustar ponderaciones según el criterio experto municipal.

#### Herramientas de configuración del scoring

El Ayuntamiento dispondrá de herramientas de configuración directa que le permiten ajustar el scoring sin dependencia del proveedor tecnológico.

##### Panel de administración de reglas de scoring

Una interfaz web intuitiva permitirá visualizar en una tabla completa todos los factores, ponderaciones y umbrales configurados. Desde ahí se podrán editar las ponderaciones modificando los puntos asignados a cada factor y ajustar los rangos de puntuación que determinan cada categoría de criticidad.

##### Ejemplo de interfaz (pseudocódigo visual):

```

CU-01: Interferencia con señales de tráfico
├─ Tipo de señal obstruida:
│   ├── STOP / Ceda el Paso: [6] puntos [editar]
│   ├── Señal reguladora: [4] puntos [editar]
│   └── Señal informativa: [2] puntos [editar]
├─ Nivel de obstrucción:
│   ├── Total (100%): [9] puntos [editar]
│   ├── Alta (75-99%): [6] puntos [editar]
│   └── Moderada (50-74%): [3] puntos [editar]
├─ Contexto adicional:
│   ├── Zona escolar: [+3] puntos [✓ activo]
│   └── Paso de peatones: [+3] puntos [✓ activo]
└─ Umbrales de criticidad:
    ├── - Crítica: [15-20] puntos [editar]
    ├── - Alta: [10-14] puntos [editar]
    ├── - Media: [5-9] puntos [editar]
    └── - Baja: [1-4] puntos [editar]

[Previsualizar cambios] [Guardar nueva versión]

```

### **Garantías de independencia municipal**

El diseño del sistema de scoring asegura que el Ayuntamiento puede operar de forma autónoma.

**Independencia operativa:** El Ayuntamiento podrá ajustar ponderaciones y umbrales directamente desde la interfaz web, sin intervención del proveedor. Las modificaciones se aplicarán inmediatamente sin necesidad de redespliegue técnico, y se podrán realizar tantos ajustes como se considere necesario sin limitaciones de frecuencia.

**Independencia técnica:** Se proporcionará un manual de usuario detallado del panel de configuración y sesiones de capacitación para responsables municipales. El soporte de SIGO es consultivo y no es requerido para cambios rutinarios.

**Independencia estratégica:** El Ayuntamiento define qué es crítico y qué no. Puede ajustar el scoring ante cambios en normativa o políticas municipales, y refinarlo continuamente basándose en su propia experiencia operativa.

### **3.5.3. Valor estratégico del enfoque preventivo**

El sistema de scoring representa un cambio fundamental en el criterio de éxito de la gestión del arbolado urbano.

#### **El criterio de viabilidad se basa en el impacto:**

El sistema se justifica económicamente si detecta anualmente un volumen relevante de incidencias de alta criticidad que permitan actuar preventivamente, independientemente de las métricas clásicas de Precision/Recall globales.

#### **El valor de prevenir una sola incidencia crítica:**

Una sola incidencia grave prevenida puede generar un ahorro superior al coste anual del sistema completo. Desde el punto de vista económico directo, evita responsabilidades patrimoniales, costes de intervención urgente no planificada, y reclamaciones ciudadanas. Estratégicamente, mejora la seguridad pública, permite optimizar recursos mediante planificación anticipada, y mejora la percepción ciudadana del servicio.

#### **Conclusión:**

GreenView® redefine el criterio de éxito: **el valor no está solo en detectar el máximo número de incidencias, sino en asegurar que las incidencias críticas sean identificadas a tiempo para actuar preventivamente.**

### **3.6. Limitaciones y próximos pasos**

Es importante ser transparentes sobre lo que hemos demostrado con datos reales y lo que aún son hipótesis pendientes de validación en fases futuras.

#### **3.6.1. Demostrado con datos**

Los siguientes aspectos han sido probados y medidos con evidencia objetiva:

- **Rendimiento de los modelos en dataset de validación (~1,000 imágenes):** Hemos cuantificado Precisión, Recall y F1-Score para los casos de uso principales (CU-01, CU-02, CU-03, CU-06, CU-08) con datos reales de Madrid
- **Funcionamiento de la arquitectura agnóstica:** Validado que el sistema puede trabajar indistintamente con GPT-4o, Gemini 2.5 Pro y Llama 3.1 405B sin cambios arquitectónicos
- **Capacidad de detección de los casos de uso específicos:** Confirmado que los agentes pueden identificar interferencias con señales, fachadas, arbolado en mal estado, obstrucciones de vías y nidos de gran tamaño
- **Proceso Human-in-the-Loop y mejora mediante RAG:** Demostrado que las validaciones humanas se integran inmediatamente en el sistema RAG, mejorando las respuestas en tiempo real
- **Detección sobre vistas vs. detección sobre activos:** Validado que ambos enfoques son complementarios, con perfiles de falsos negativos diferentes

#### **3.6.2. Hipótesis a validar en siguiente fase**

Los siguientes aspectos requieren validación en un piloto operativo más amplio:

- **Escalabilidad a decenas de miles de imágenes:** Aunque hemos procesado ~1,000 imágenes, es necesario confirmar que el sistema mantiene rendimiento y tiempos de respuesta aceptables con volúmenes 10x o 100x mayores
- **Comportamiento del scoring en operativa real continua:** El sistema de scoring ha sido calibrado con casos reales, pero falta validación con flujos de trabajo operativos continuos del Ayuntamiento durante varios meses
- **Evolución de modelos open-source en 1-2 años:** La proyección de convergencia de modelos open-source hacia niveles Pro se basa en tendencias históricas, pero requiere seguimiento temporal para confirmarse
- **Impacto real en prevención de incidencias críticas:** Es necesario medir cuántas incidencias detectadas proactivamente habrían escalado a problemas graves sin intervención, lo que requiere seguimiento longitudinal

## 4. Viabilidad y valor

Este capítulo demuestra que GreenView® no solo funciona técnicamente, sino que es viable económicamente y sostenible a largo plazo. Se presenta un modelo de costes transparente desglosado en cuatro componentes: plataforma tecnológica (~60-70k/año), captación de imágenes (~54k/año por equipo), procesamiento de IA (variable, con diferencia 20x entre modelos Pro y open-source), y asistencia técnica (~15-20k/año). Se explica cómo la arquitectura híbrida optimiza coste-precisión reduciendo el coste de IA en ~85%. Se presenta un escenario de referencia ilustrativo (1.500 km de viario, coste total anual 137-150k€) y se analizan las economías de escala (coste por km baja de 17€ a 10,4€ al triplicar cobertura). El capítulo detalla las estrategias de independencia tecnológica, la gestión de riesgos con sus mitigaciones, y las capacidades de escalabilidad en múltiples dimensiones. Finalmente, incluye un anexo técnico completo con especificaciones recomendadas para futuros pliegos que garanticen calidad, independencia y evolución continua.

Los capítulos anteriores han demostrado que GreenView® funciona desde el punto de vista técnico: la arquitectura de análisis híbrido detecta incidencias con precisión validada, el sistema de aprendizaje continuo mejora progresivamente, y la independencia tecnológica está garantizada mediante diseño agnóstico.

La viabilidad de una solución tecnológica se mide también por su sostenibilidad económica, su escalabilidad operativa y su alineamiento con los intereses de largo plazo de la administración. Este capítulo aborda el modelo de costes, las estrategias de mitigación de riesgos y las especificaciones técnicas para futuros pliegos.

### 4.1. Análisis de costes de despliegue

La viabilidad económica de GreenView® se fundamenta en un modelo de costes transparente, escalable y optimizado. Durante el piloto se ha realizado una evolución del modelo de costes que permite al Ayuntamiento comprender con precisión los factores que determinan el coste total de la solución y planificar su despliegue de forma estratégica.

#### **4.1.1. Evolución del modelo de costes: de unitario a desglosado**

En el desarrollo del modelo de costes, inicialmente se presentaba de forma unitaria (coste por árbol o por imagen), donde todos los componentes (fijos y variables) se diluían en función del volumen total de captura. Si bien este enfoque es útil para comparaciones rápidas, resultaba insuficiente para una planificación operativa rigurosa.

El modelo actual separa claramente los **costes fijos** (independientes del volumen) de los **costes variables** (proporcionales a la actividad), lo que aporta **transparencia** para identificar qué componentes son estructurales y cuáles dependen de la intensidad de uso, permite **optimización** al diseñar estrategias que maximicen el aprovechamiento de los recursos fijos, garantiza **escalabilidad** al proyectar con precisión el impacto económico de diferentes escenarios de despliegue, y ofrece **flexibilidad** para adaptar el modelo a las necesidades reales del Ayuntamiento sin sorpresas en los costes.

## 4.1.2. Componentes del modelo de costes

El modelo de costes de GreenView® se estructura en cuatro bloques principales:

### 4.1.2.1. Plataforma tecnológica SiGO (coste fijo anual)

**Coste anual estimado: ~60.000 - 70.000 €**

Este componente incluye todos los servicios y sistemas necesarios para el funcionamiento de la plataforma de análisis y gestión de incidencias, incluyendo la infraestructura cloud (Azure) y todos los servicios asociados:

- **Infraestructura cloud y almacenamiento:** Servidor de procesamiento y orquestación de agentes de IA, almacenamiento seguro de vídeos brutos, imágenes panorámicas, vistas procesadas y resultados, base de datos para incidencias, activos y metadatos, y sistema de copias de seguridad y políticas de recuperación.
- **Procesamiento de vídeo e imágenes:** Procesado de los vídeos brutos capturados por los vehículos, generación de imágenes panorámicas 360°, extracción de vistas direccionales (cardinales) y detección y recorte de activos mediante modelos YOLO.
- **Motor de inteligencia artificial:** Orquestador de agentes (sistema central que lanza y gestiona los análisis), motor de scoring para la priorización automática de incidencias y sistema de gestión de múltiples proveedores de IA (arquitectura agnóstica).
- **Plataforma de validación y gestión:** Interfaz web de Human-in-the-Loop para validación por expertos, generación automática de tablas de incidencias y cuadros de mando, panel de administración y configuración de agentes, y sistema de trazabilidad y auditoría.
- **Supervisión técnica especializada:** Dedicación de un técnico especializado (mínimo 4 horas diarias), soporte técnico y resolución de incidencias, y supervisión del correcto funcionamiento de los agentes.
- **Mantenimiento evolutivo:** Mejora continua de los agentes de IA existentes, capacidad para incorporar nuevos agentes, modelos o proveedores, actualizaciones de seguridad y funcionalidad, y adaptación a normativas (GDPR, protección de datos).

*Nota: Este mantenimiento evolutivo garantiza la operativa continua de la plataforma. Dado que se prevé la necesidad de evolucionar los agentes con el feedback operativo, incorporar nuevos casos de uso según prioridades municipales emergentes, integrar nuevos modelos de IA (especialmente open-source), y desarrollar integraciones específicas (ej. gemelo digital), se recomienda complementar este componente con el componente D “Asistencia técnica y evolución de plataforma” (15-20k/año), que proporciona capacidad de desarrollo estratégico mediante una bolsa de horas técnicas.*

- **Gestión del proyecto:** Coordinación técnica y estratégica, seguimiento de objetivos y KPIs, y elaboración de informes periódicos.

Este coste es **independiente del volumen de captura** y representa la inversión necesaria para mantener la plataforma operativa y en continua evolución.

#### **4.1.2.2. Captación de imágenes (coste variable por equipo)**

**Coste anual por equipo de captura: ~53.900 €**  
**Coste unitario por kilómetro capturado: ~5,5 €/km**

La captación de imágenes es el componente variable más relevante del modelo. Cada equipo de captura está compuesto por:

- **Recursos humanos (1 operador):** Salario y costes de empresa de un técnico cualificado. **Coste anual estimado: ~42.000 €.**
- **Vehículo (modelo renting/leasing):** Cuota anual que incluye mantenimiento, seguro, impuestos y reparaciones. **Coste anual estimado: ~8.000 €.**
- **Equipamiento tecnológico (hardware):** Kit de captura (cámara 360°, GPS de alta precisión, soportes) y estación de volcado de datos (que puede ser compartida por varios vehículos), con renovación cada dos años. **Coste anual amortizado: ~2.000 €.**
- **Costes operativos variables:** Combustible, consumibles y otros gastos menores. **Coste anual estimado: ~2.000 €.**

#### **IMPORTANTE - ESTIMACIONES OPERATIVAS:**

Las siguientes cifras son **estimaciones** basadas en análisis de referencias operativas de captura urbana con vehículos, no en capturas reales realizadas durante el piloto:

- **Jornada de conducción efectiva:** 5,5 horas (estimado)
- **Velocidad media efectiva:** ~8,2 km/h (estimado, considerando tráfico, paradas técnicas y condiciones urbanas)
- **Rendimiento por jornada:** ~45 km/día (estimado)
- **Datos generados por jornada:** ~300 GB (estimado)
- **Días operativos al año:** ~220 días laborables (estimado, incluyen descuentos por vacaciones, formación y mantenimiento)

Factores que requerirán validación operativa en fase de despliegue:

- **Variabilidad estacional:** Diferencias entre verano/invierno, horarios pico/valle
- **Densidad de tráfico variable:** Zonas con mayor congestión reducirán la velocidad media
- **Complejidad del viario:** Calles estrechas, zonas peatonales, obras pueden requerir captura a pie

Se recomienda realizar una **fase de calibración operativa** en diferentes distritos y estaciones para validar estas estimaciones antes de establecer los rendimientos definitivos para la planificación a gran escala.

#### **Capacidad anual de un equipo de captura:**

- **Kilómetros capturados al año:** ~9.900 km

- **Coste total operativo anual:** ~53.900 €
- **Coste por kilómetro:** ~5,5 €

#### **Inversión inicial (CAPEX):**

El hardware necesario para la captura requiere una inversión inicial moderada:

- **Kit de vehículo: 900 - 1.550 € (cámara, GPS, soportes)**
- **Estación de volcado: 2.300 - 3.700 €**
- **Total por equipo: 3.200 - 5.250 €**

Esta inversión se amortiza en 2 años y está incluida en el coste anual del equipamiento tecnológico.

**NOTA SOBRE PRECIOS:** Los precios del kit de captura deben ser validados con proveedores actuales antes del despliegue definitivo. Los costes indicados se basan en referencias de mercado del momento del piloto y pueden requerir actualización.

#### **4.1.2.3. Procesamiento de inteligencia artificial (coste variable por imagen)**

Este es el componente más estratégico del modelo, ya que su coste varía significativamente según la tecnología utilizada y la estrategia de análisis implementada.

#### **Costes por modelo de IA (€ por imagen/agente):**

Proveedor	Modelo	Open-Source	Coste por imagen
OpenAI	GPT-5	No	0,03 - 0,06 €
OpenAI	GPT-4o	No	0,02 - 0,03 €
Google	Gemini 2.5 Pro	No	0,02 - 0,05 €
Google	Gemini 2.5 Flash	No	0,003 - 0,008 €
Meta	Llama 4 Maverick*	Sí	0,0005 - 0,002 €

**Diferencia de coste:** Los modelos open-source (Llama 4, Qwen, Mistral, entre otros) representan aproximadamente el **5% del coste** de los modelos comerciales premium (GPT-5, Gemini Pro), o incluso menos.

#### **Implicaciones de los modelos Open-Source:**

Durante el piloto se han realizado pruebas comparativas con modelos open-source, obteniendo los siguientes hallazgos preliminares:

1. **Desempeño en análisis de activos:** En datasets de inspecciones (CU-03), el rendimiento es prácticamente igual, con una precisión ligeramente inferior (95% vs 100%) debido a más falsos positivos, pero con un recall adecuado. En datasets mixtos más complejos, el rendimiento baja, aunque la hipótesis es que las incidencias no detectadas corresponden a scores de prioridad más bajos.
2. **Ventaja inesperada en análisis de vistas:** Los modelos open-source funcionan especialmente bien en el análisis de vistas de entorno urbano. En algunos casos,

detectan **más incidencias con análisis de vistas que con análisis de activos**, lo cual es una ventaja significativa considerando el menor coste de procesamiento.

3. **Importancia del scoring:** El factor crítico no es detectar el 100% de incidencias, sino **no perder las incidencias críticas**. El sistema de scoring permite cuantificar qué se está omitiendo y confirmar que el trade-off coste-cobertura es aceptable.
4. **Validación Humana:** El sistema Human-in-the-Loop actúa como filtro de calidad. Los falsos positivos (que también aparecen con modelos “pro” en datasets grandes) se descartan rápidamente en la validación.
5. **Evolución continua:** Cada año surgen nuevos modelos open-source más precisos, y además ofrecen mayores facilidades para fine-tuning y reentrenamiento, lo que aumenta su potencial de mejora.

### **Casos de uso detectados por cada agente:**

En las pruebas realizadas, cada árbol es analizado por agentes especializados que cubren los siguientes casos de uso:

- CU-01: Interferencias con seguridad vial
- CU-02: Interferencias con fachadas
- CU-03: Mal estado del arbolado
- CU-05: Alcorques vacíos
- CU-06: Obstrucción de vías
- CU-08: Nidos de cotorras
- CU-10: Desbroce
- CU-12: Papeleras

*Nota: Los casos de uso CU-10 (desbroce) y CU-12 (papeleras) no han sido analizados con el mismo nivel de detalle durante el piloto, pero el proceso técnico para su detección mediante el agente de vistas está resuelto y puede implementarse según las necesidades del Ayuntamiento.*

#### **4.1.2.4. Asistencia técnica y evolución de plataforma (coste variable/semi-fijo)**

**Coste anual estimado: 15.000 - 20.000 €**

Este componente garantiza la **capacidad de evolución continua** de la plataforma según las necesidades operativas del servicio y los objetivos estratégicos del Ayuntamiento. Se estructura como una **bolsa anual de horas técnicas** que permite implementar mejoras, desarrollar nuevas funcionalidades y adaptar el sistema a requisitos emergentes.

#### **¿Qué incluye este servicio?**

- **Desarrollo de nuevos casos de uso:** Implementación de agentes de IA para detectar nuevas incidencias identificadas como prioritarias, adaptación de modelos existentes a nuevas especies de árboles o elementos urbanos, e integración de capacidades detectivas adicionales según necesidades operativas.

- **Mejora continua de modelos:** Refinamiento de agentes existentes basándose en feedback operativo, optimización de parámetros de detección (umbrales, sensibilidad, criterios de scoring), reentrenamiento periódico de modelos con datasets curados del Ayuntamiento y ajustes estacionales necesarios.
- **Integraciones y extensiones:** Desarrollo de conectores con nuevos sistemas municipales, **integración con plataforma de gemelo digital urbano** para sincronización bidireccional de datos, implementación de nuevas exportaciones de datos o formatos requeridos, y conexión con sistemas de las empresas contratistas.
- **Soporte técnico avanzado:** Resolución de incidencias complejas que requieren desarrollo, asesoramiento técnico en optimización de flujos de trabajo, formación especializada para nuevos usuarios o funcionalidades, y análisis de rendimiento con propuestas de mejora.
- **Personalización operativa:** Ajustes en la interfaz de usuario según feedback de validadores, configuración avanzada de reglas de negocio y automatizaciones, y desarrollo de informes personalizados o cuadros de mando específicos.

#### Estructura de la bolsa de horas:

Concepto	Tarifa	Volumen anual	Coste anual
Horas de ingeniería técnica	40 €/h	375 - 500 h	15.000 - 20.000 €

#### Equivalencia en dedicación:

- **375 horas/año**  $\approx$  1,7 horas/día laboral  $\approx$  8,5 horas/semana
- **500 horas/año**  $\approx$  2,3 horas/día laboral  $\approx$  11,5 horas/semana

#### Modelo de consumo:

La bolsa de horas puede consumirse de forma flexible según las necesidades:

1. **Modelo distribuido:** Pequeñas mejoras incrementales durante todo el año
  - Ejemplo: Ajustes semanales de parámetros, refinamiento continuo
2. **Modelo por proyectos:** Concentración de horas en desarrollos específicos
  - Ejemplo: Desarrollo de integración con gemelo digital (80-120h de proyecto)
3. **Modelo híbrido:** Combinación de mantenimiento continuo + proyectos puntuales
  - Ejemplo: 200h de mejoras continuas + 300h para 2-3 proyectos específicos

#### Gestión y transparencia:

- **Registro detallado:** Todas las horas consumidas quedan registradas con descripción de la tarea realizada
- **Informes mensuales:** Reporte de horas consumidas, actividades realizadas y saldo disponible
- **Priorización conjunta:** El Ayuntamiento decide qué desarrollos priorizar según sus necesidades

- **Sin penalización por no uso:** Horas no consumidas no generan coste adicional (modelo pay-per-use dentro del límite anual)

### Valor estratégico:

Este componente es fundamental para asegurar que GreenView® **no sea una tecnología estática**, sino una **plataforma viva** que:

1. **Se adapta:** A cambios normativos, nuevas prioridades municipales, feedback operativo
2. **Evoluciona:** Incorpora mejoras tecnológicas, nuevos modelos de IA, funcionalidades emergentes
3. **Capitaliza aprendizaje:** Convierte la experiencia operativa en mejoras concretas del sistema
4. **Mantiene independencia:** Permite al Ayuntamiento dirigir la evolución tecnológica según sus intereses

### Diferenciación con el coste de plataforma:

- **Plataforma SiGO (60-70k/año):** Mantiene la infraestructura operativa, sin cambios funcionales
- **Asistencia técnica (15-20k/año):** Desarrolla nuevas capacidades y mejoras específicas

Ambos componentes son complementarios: la plataforma garantiza la operativa diaria, la asistencia técnica garantiza la evolución estratégica.

### 4.1.3. Estrategia híbrida: optimización coste-precisión

La innovación clave del modelo económico de GreenView® es el **análisis híbrido**, que combina lo mejor de ambos mundos:

#### 4.1.3.1. Arquitectura del modelo híbrido

1. **Primera fase - Triaje con Open-Source (nivel de vista):**
  - Análisis inicial de todas las vistas direccionales con modelos de bajo coste
  - Detección de posibles incidencias en la escena completa
  - Umbral de sensibilidad alto para no perder incidencias relevantes
  - Generación de un conjunto reducido de vistas sospechosas
2. **Segunda fase - Confirmación con modelos Pro (nivel de activo):**
  - Cuando una vista presenta indicios de incidencia, se activa el análisis de activos
  - **Consideración importante:** Como no sabemos qué activo específico causa la incidencia detectada en la vista, **se analizan todos los activos presentes en esa vista**
  - Estimación: ~4 activos (árboles) por vista de media
  - Uso de modelos de alta precisión (GPT-5, Gemini Pro) para confirmar y clasificar cada activo
  - Identificación precisa del activo responsable y caracterización detallada de la incidencia

#### 4.1.3.2. Implicación en costes

Si se detectan 20.000 incidencias en el nivel de vistas:

- **NO** se realizan 20.000 análisis de activos
- **SÍ** se realizan ~80.000 análisis de activos (20.000 vistas × 4 activos/vista)

Esta multiplicación es necesaria porque:

- El modelo de vistas identifica que hay una incidencia en la escena
- Pero NO identifica QUÉ árbol específico la causa
- Por lo tanto, hay que analizar todos los árboles visibles con el modelo de precisión

#### **4.1.3.3. Resultado**

Cobertura completa con coste optimizado. El modelo híbrido sigue siendo mucho más eficiente que analizar los ~6,5 millones de activos totales con modelos Pro (1.620.000 vistas × 4 activos/vista de media = 6.480.000 activos).

#### **4.1.4. Escenario de referencia: análisis de viabilidad**

Para ilustrar la viabilidad del modelo, se presenta un escenario de referencia basado en un análisis preliminar de las calles de Madrid con arbolado:

##### **Contexto:**

Según el [Plan de Infraestructura Verde y Biodiversidad del Ayuntamiento de Madrid](#), la ciudad cuenta con **más de 2.600 km de calles con arbolado**. Para efectos de este análisis de viabilidad económica, se presenta un escenario simplificado con una cifra conservadora de **~1.500 km** que permite ilustrar el modelo de costes sin pretender representar la cobertura total real.

**IMPORTANTE:** Las cifras de este escenario (1.500 km, 8.100 km/año, 405.000 puntos de muestreo) son **ejemplos ilustrativos** para análisis de viabilidad económica. **No representan** el alcance definitivo de un despliegue operativo, que deberá ser dimensionado con datos GIS oficiales actualizados del Ayuntamiento considerando: (a) la red real de calles con arbolado (~2.600 km), (b) la estrategia de frecuencia óptima por jerarquía viaria, y (c) los objetivos de cobertura y priorización que establezca el área de Zonas Verdes.

##### **Estrategia de frecuencia por ferarquía viaria (modelo intensivo):**

Nivel	Tipo de Vía	% de la red	Km	Frecuencia	Total km/año
1	Vías Principales	15%	225 km	12 veces/año	2.700 km
2	Vías Secundarias	35%	525 km	6 veces/año	3.150 km
3	Vías Residenciales	50%	750 km	3 veces/año	2.250 km
<b>TOTAL</b>		<b>100%</b>	<b>1.500 km</b>		<b>8.100 km/año</b>

##### **Parámetros de captura y procesamiento:**

- **Longitud anual recorrida:** 8.100 km
- **Puntos de muestreo:** 405.000 (cada 20 metros)

- **Vistas generadas:** 1.620.000 imágenes (4 vistas por punto)
- **Activos detectados:**  $\sim 4$  activos por vista  $\times$  2 agentes = 3.240.000 inferencias

## Desglose de costes del escenario (modelo híbrido)

### Costes fijos:

- Plataforma SiGO: **60.000 €/año**
- Asistencia técnica y evolución: **17.500 €/año** (punto medio de 15-20k)
- Captación anual: **54.000 €/año** (1 equipo para 8.100 km)

### Costes variables de IA (modelo híbrido):

- Fase 1 - Open-Source (todas las vistas):  $1.620.000 \times 0,002 \text{ €} = \mathbf{3.240 \text{ €}}$
- Fase 2 - Modelos Pro (análisis de activos en vistas con incidencias):
  - Con 20.000 incidencias detectadas:  $20.000 \text{ vistas} \times 4 \text{ activos/vista} \times 0,04 \text{ €} = \mathbf{3.200 \text{ €}} \rightarrow \mathbf{\text{Total IA: 6.440 €}}$
  - Con 50.000 incidencias detectadas:  $50.000 \text{ vistas} \times 4 \text{ activos/vista} \times 0,04 \text{ €} = \mathbf{8.000 \text{ €}} \rightarrow \mathbf{\text{Total IA: 11.240 €}}$
  - Con 100.000 incidencias detectadas:  $100.000 \text{ vistas} \times 4 \text{ activos/vista} \times 0,04 \text{ €} = \mathbf{16.000 \text{ €}} \rightarrow \mathbf{\text{Total IA: 19.240 €}}$

*Nota: Se analizan todos los activos de cada vista donde se detectó incidencia porque no sabemos qué activo específico la causa hasta realizar el análisis de precisión.*

### Coste total anual del escenario:

Incidencias detectadas	Coste modelo PRO	Subtotal (OS + PRO)	TOTAL
20.000	3.200 €	6.440 €	<b>137.940 €</b>
50.000	8.000 €	11.240 €	<b>142.740 €</b>
100.000	16.000 €	19.240 €	<b>150.740 €</b>

### Interpretación:

1. **El coste de IA crece con las incidencias pero sigue siendo manejable:**
  - Con 20.000 incidencias: ~5% del coste total (6.440 €)
  - Con 100.000 incidencias: ~13% del coste total (19.240 €)
2. **Los costes fijos son dominantes:** El 87-95% del coste está en la plataforma, asistencia técnica y la captación
3. **El modelo híbrido es altamente eficiente:** Sin modelo híbrido, analizar todos los activos (~8,1M) con modelos Pro costaría ~324.000 €, más del doble del coste total del sistema completo. El híbrido reduce el coste de IA en un 94% incluso con 100.000 incidencias detectadas
4. **La asistencia técnica representa el 12-13% del coste total:** Esta inversión garantiza la capacidad de evolución y adaptación continua del sistema

**NOTA:** El rango de 20.000-100.000 incidencias es una **proyección estimativa** para este escenario. El número real de incidencias dependerá del estado del arbolado, las condiciones climáticas, la configuración de los umbrales de detección y la frecuencia de captura. Durante la fase operativa se ajustarán estos parámetros basándose en datos reales.

#### 4.1.5. Factores de optimización dinámica

El modelo de costes no es estático. En fases futuras, el sistema podría permitir ajustar la frecuencia de captura en base a un **índice de riesgo** que considere:

- **Características del árbol:** Especie, edad, estado de salud aparente
- **Histórico de incidencias:** Frecuencia y gravedad de incidencias previas detectadas
- **Contexto urbano:** Aforo de la vía, proximidad a colegios/hospitales, categoría de la calle
- **Eventos climáticos:** Intensificación de la vigilancia tras episodios de viento o nieve

Esta capacidad de ajuste dinámico permitiría concentrar recursos donde son más necesarios, maximizando el impacto preventivo sin incrementar costes.

#### 4.1.6. Proyección de costes y economía de escala

El modelo de costes demuestra una **economía de escala favorable**:

**Coste unitario por kilómetro inspeccionado:**

- 8.100 km/año → 137.940 € (20k incidencias) → **17,0 €/km**
- Con 2 equipos y 16.200 km/año → 195.140 € (40k incidencias) → **12,0 €/km** (-29%)
- Con 3 equipos y 24.300 km/año → 252.340 € (60k incidencias) → **10,4 €/km** (-39%)

#### 4.1.7. Capacidades de evolución y escalabilidad del sistema

GreenView® ha sido diseñado como una **plataforma evolutiva** que puede crecer y adaptarse a las necesidades futuras del Ayuntamiento sin requerir rediseños estructurales. La escalabilidad no se limita a aumentar el volumen de captura, sino que incluye múltiples dimensiones de crecimiento y mejora continua.

##### 4.1.7.1. Evolución funcional: Mejora continua de capacidades

- **Refinamiento de casos de uso existentes:** El sistema permite la mejora de precisión mediante reentrenamiento con dataset municipal curado, la optimización de umbrales de detección basados en feedback operativo, el ajuste de reglas de scoring según experiencia real de priorización, y la especialización en especies arbóreas específicas de Madrid.
- **Incorporación de nuevos casos de uso:** La plataforma contempla el desarrollo de agentes adicionales según prioridades municipales emergentes, la detección de nuevas tipologías de incidencias identificadas durante la operativa, la extensión a otros elementos urbanos (pavimentos, señalización, mobiliario), y casos de uso estacionales (riego, eventos climáticos).
- **Mejora de modelos de IA:** La estrategia de mejora incluye la integración progresiva de modelos open-source propios alojados en infraestructura municipal, la adopción de nuevos modelos comerciales más precisos o económicos conforme emergen, el reentrenamiento periódico que especializa los modelos en la realidad de Madrid, y la optimización del modelo híbrido para maximizar el ratio coste-precisión.

#### **4.1.7.2. Escalabilidad operativa: Intensificación de la cobertura**

- **Aumento de frecuencia de captura:** El sistema permite el incremento de pasadas anuales en zonas críticas (de 3-6x a 12x por año), la captura estacional intensiva post-temporales, recapturas dirigidas en áreas con incidencias recurrentes detectadas, y el ajuste dinámico de frecuencia basado en índice de riesgo por zona.
- **Ampliación de equipos de captura vehicular:** La escalabilidad contempla la incorporación de equipos de captura adicionales según crecimiento de cobertura, la optimización de rutas para maximizar kilómetros capturados por equipo, y la coordinación con vehículos municipales existentes (limpieza, mantenimiento), con un **coste incremental de ~54.000 €/año por equipo adicional.**
- **Integración de fuentes de captura complementarias:**

**Captura mediante mochilas portátiles (incluido en el reto original):** Zonas peatonales, parques y calles estrechas inaccesibles en vehículo. Kit de mochila 360° (~1.200-1.800 €) + operador a pie con rendimiento de ~2-3 km/día en zonas de alta densidad peatonal. Caso de uso: complemento para alcorques ocultos por vehículos aparcados.

**Aprovechamiento de imágenes de otras fuentes municipales:** Google Street View y plataformas similares para validación de contexto histórico, integración de captura en vehículos de empresas contratistas y servicios municipales (limpieza, recogida de residuos, Policía Municipal), y captura aérea mediante drones para zonas de difícil acceso o inspección de copas de árboles.

#### **4.1.7.3. Escalabilidad técnica: Capacidad de procesamiento**

La arquitectura se basa en diseño cloud nativo (Azure) con escalado automático según carga, procesamiento paralelo de imágenes sin cuellos de botella estructurales, y proyección técnica sin límites conocidos para volúmenes escalables según las necesidades del Ayuntamiento.

La optimización incluye mejoras en eficiencia de algoritmos de detección (YOLO, extracción de vistas), reducción de tiempos de procesamiento mediante paralelización mejorada, gestión inteligente de caché y almacenamiento para minimizar costes, e integración de nuevas GPUs o aceleradores según disponibilidad.

#### **4.1.7.4. Escalabilidad económica: Economía de escala**

Equipos de captura	Km/año capturados	Coste plataforma SiGO	Coste por km (solo plataforma)
1 equipo	9.900 km	60.000 €	6,1 €/km
2 equipos	19.800 km	60.000 €	3,0 €/km (-49%)
3 equipos	29.700 km	60.000 €	2,0 €/km (-67%)

La plataforma SiGO (60-70k/año) es un coste fijo que se diluye al aumentar la cobertura, generando economías de escala favorables. La transición a modelos open-source reduce el coste de inferencia en más del 90%, la mejora de precisión reduce la carga de validación humana, y la optimización del modelo híbrido minimiza el uso de modelos premium.

#### **4.1.7.5. Escalabilidad de integración: Conexión con ecosistema municipal**

La integración contempla la sincronización bidireccional de datos entre GreenView® y el gemelo digital, la visualización 3D de incidencias en contexto urbano completo, simulaciones predictivas de evolución de arbolado, y una base para planificación urbana basada en datos reales.

El sistema ofrece APIs abiertas para integración con sistemas de empresas contratistas, exportación automática a sistemas de gestión de mantenimiento (GMAO), integración con plataformas de denuncias ciudadanas (010, app móvil), y conexión con sistemas de emergencias y protección civil. La plataforma permite el intercambio de mejores prácticas y configuraciones de scoring, la compartición de datasets anonimizados para mejorar modelos, la colaboración en desarrollo de nuevos casos de uso, y el establecimiento de estándares abiertos para inspección de arbolado urbano.

#### **4.1.7.6. Conclusión sobre escalabilidad**

GreenView® no es una solución estática con límites de crecimiento, sino una **plataforma diseñada para evolucionar** en múltiples dimensiones:

1. **Funcionalmente:** Nuevos casos de uso y mejora continua de modelos
2. **Operativamente:** Aumento de frecuencia, equipos y fuentes de captura complementarias
3. **Técnicamente:** Arquitectura cloud sin límites conocidos de procesamiento
4. **Económicamente:** Economías de escala favorables al crecer la cobertura
5. **Estratégicamente:** Integración con ecosistema municipal y capacidad de interoperabilidad

La plataforma está **preparada para crecer al ritmo que el Ayuntamiento determine**, sin requerir inversiones estructurales adicionales ni rediseños arquitectónicos.

#### **4.1.8. Justificación de la inversión: valor estratégico vs. coste operativo**

Una vez presentado el modelo de costes y sus capacidades de escalabilidad, cabe plantear la pregunta clave: **¿cómo se justifica esta inversión?**

La respuesta no puede basarse en una comparación directa con el modelo actual de gestión (que involucra múltiples contratos y procesos no siempre cuantificables de forma unitaria), sino en el **valor potencial** que puede aportar la anticipación proactiva.

La justificación de GreenView® se plantea más desde el **valor de poder anticiparse a los problemas** que desde una comparación coste-beneficio unitaria con los métodos reactivos actuales.

En un contexto donde el Ayuntamiento gestiona decenas de miles de incidencias anuales de forma reactiva (denuncias ciudadanas, avisos de emergencia), la hipótesis es que el coste de respuesta podría ser mayor que el de la anticipación:

#### **Comparativa de costes: Gestión reactiva vs. Gestión proactiva**

- **Gestión reactiva tradicional:**

- Movilización de recursos sin planificación previa
- Intervención urgente con sobrecoste operativo
- Gestión de reclamaciones ya activadas
- Exposición a responsabilidad patrimonial
- Impacto reputacional de problemas visibles
- **Gestión proactiva con GreenView®:**
  - Identificación temprana y planificación optimizada de actuaciones
  - Intervención preventiva con menor coste unitario
  - Prevención de reclamaciones antes de que se produzcan
  - Reducción de exposición a responsabilidad patrimonial
  - Gestión anticipada que mejora la percepción ciudadana

### **Valor potencial de la prevención:**

La prevención de incidencias críticas podría generar valor de diversas formas, aunque su cuantificación precisa es compleja y depende de múltiples factores:

- **Reducción de exposición a riesgos:** El sistema puede ayudar a identificar situaciones que podrían derivar en accidentes, aunque no garantiza su prevención total
- **Optimización de intervenciones:** Menor necesidad de movilizaciones urgentes no planificadas, permitiendo una gestión más ordenada de los recursos
- **Mejora en la gestión de reclamaciones:** Mayor capacidad de respuesta basada en evidencia documental del estado de los activos
- **Planificación más informada:** Conocer con antelación la ubicación y criticidad de las intervenciones necesarias facilita la optimización de rutas y recursos
- **Reducción de costes asociados:** Potencial disminución de costes derivados de responsabilidad patrimonial, aunque su materialización es incierta

Es importante señalar que el valor real del sistema se fundamenta más en la mejora continua de la gestión operativa y la reducción gradual de riesgos que en un cálculo directo de ahorro por incidente específico prevenido.

### **Enfoque de viabilidad basado en mejora operativa:**

El sistema busca justificarse a través de la detección de un volumen relevante de incidencias que permitan mejorar la gestión preventiva, complementando (no sustituyendo) los mecanismos reactivos existentes. En este enfoque:

1. Los falsos positivos tienen un coste de gestión bajo (~10 seg/incidencia) mediante Human-in-the-Loop, permitiendo configurar umbrales de sensibilidad altos sin impacto operativo significativo.
2. El objetivo no es alcanzar una exhaustividad del 100%, sino identificar suficientes incidencias críticas para mejorar la gestión del arbolado urbano.
3. El valor se construye progresivamente con el uso continuado del sistema y el aprendizaje acumulado del equipo municipal.

Este enfoque de viabilidad no requiere que el sistema tenga exhaustividad absoluta, sino que sea capaz de identificar un número suficiente de incidencias críticas para justificar la inversión a través del valor preventivo generado.

### Cambio de paradigma: de reactivo a proactivo

Aspecto	Modelo Actual	GreenView®
<b>Cobertura</b>	Limitada a denuncias y rutas programadas	100% del viario con arbolado, múltiples veces al año
<b>Detección</b>	Reactiva (problema ya evidente)	Proactiva (identificación temprana)
<b>Evidencia</b>	Variable, sin registro fotográfico sistemático	Registro fotográfico completo y geolocalizado
<b>Priorización</b>	Criterio subjetivo, orden de llegada	Scoring objetivo basado en criticidad
<b>Trazabilidad</b>	Limitada	Histórico completo de evolución de cada activo
<b>Conocimiento</b>	Disperso entre contratistas	Capitalizado internamente en dataset propio

### Valor estratégico no cuantificable directamente:

1. **Reducción de riesgo de accidentes:** Detección temprana de situaciones peligrosas antes de que generen incidentes
2. **Optimización de contratistas:** Datos objetivos para planificar actuaciones y auditar trabajos realizados
3. **Mejora en percepción ciudadana:** Respuesta proactiva vs. reactiva a problemas del espacio público
4. **Soberanía del dato:** Generación de un activo informacional propio del Ayuntamiento
5. **Base para decisiones estratégicas:** Datos para optimizar inversiones en renovación de arbolado y diseño urbano

**Conclusión:** GreenView® no compite en coste directo con el modelo actual, sino que **añade una capa de inteligencia y anticipación** que transforma la gestión de zonas verdes de Madrid, convirtiendo un sistema reactivo en uno predictivo y basado en datos.

## **4.2. Independencia tecnológica y evolución continua**

El valor estratégico de GreenView® reside en su diseño, que busca facilitar la independencia tecnológica, la interoperabilidad y la capacidad de evolución del sistema. El enfoque se basa en varios principios que buscan reducir el riesgo de dependencia de proveedores externos (*vendor lock-in*).

### **4.2.1. Arquitectura agnóstica**

GreenView® ha sido diseñado para operar con diferentes proveedores de IA. Durante el piloto se ha validado su funcionamiento con distintos modelos (OpenAI, Google), lo que permite al Ayuntamiento seleccionar la tecnología más adecuada en cada momento según criterios de coste y rendimiento, evitando quedar vinculado a un único ecosistema.

### **4.2.2. Interoperabilidad con sistemas municipales**

La plataforma se ha desarrollado con un enfoque de interoperabilidad. Su diseño basado en APIs busca facilitar la integración con sistemas municipales existentes, como la Plataforma de Madrid Inteligente (MINT), y permitir la comunicación con futuras herramientas sin requerir desarrollos complejos.

### **4.2.3. Preparación para modelos open-source**

El diseño agnóstico abre la posibilidad de integrar modelos de IA open-source en el futuro. Esta capacidad permite al Ayuntamiento considerar una transición progresiva hacia infraestructura propia, lo que podría traducirse en una reducción de costes operativos y mayor control sobre la continuidad y evolución del servicio.

### **4.2.4. Capitalización del conocimiento interno**

El sistema de aprendizaje continuo, alimentado por las validaciones del personal municipal (human-in-the-loop), busca crear un activo de conocimiento especializado: modelos de IA que aprenden de la realidad específica de Madrid. Este conocimiento se acumula internamente, desarrollando una herramienta progresivamente más adaptada al contexto local, que permanece como propiedad del Ayuntamiento.

## **4.3. Gestión de riesgos y estrategias de mitigación**

Como cualquier innovación tecnológica, GreenView® presenta desafíos que han sido identificados durante el piloto y para los cuales se han diseñado estrategias de mitigación:

### **4.3.1. Riesgo: variabilidad en el rendimiento de los modelos de IA**

**Escenario:** Posibles variaciones en el rendimiento de los modelos de IA en diferentes condiciones operativas (luminosidad, estación del año, tipos de arbolado no vistos durante el entrenamiento).

**Contexto de evaluación:**

El criterio de éxito del sistema no se basa en alcanzar umbrales de Precision/Recall absolutos del 100%, sino en **detectar un volumen suficiente de incidencias críticas** que justifiquen la inversión mediante la anticipación proactiva.

Dado que el Ayuntamiento ya gestiona decenas de miles de incidencias anuales de forma reactiva, el valor diferencial de GreenView® reside en identificar problemas antes de que escalen, complementando (no sustituyendo) los mecanismos reactivos existentes.

#### **Mitigación implementada:**

1. **Sistema Human-in-the-Loop:** Los falsos positivos tienen un coste marginal bajo (~10 segundos/incidencia), lo que permite configurar umbrales de sensibilidad altos sin impacto operativo significativo. Es preferible detectar de más y filtrar rápidamente, que perder incidencias críticas.
2. **Enfoque en criticidad y valor preventivo:** Una sola incidencia grave prevenida (ej. caída de rama que afecta seguridad vial, interferencia que causa accidente) puede generar un ahorro superior al coste anual del sistema completo (responsabilidad patrimonial evitada, accidentes prevenidos, intervención urgente innecesaria).
3. **Complementariedad con gestión reactiva:** El sistema complementa los mecanismos reactivos existentes (denuncias, inspecciones tradicionales), añadiendo una capa de anticipación proactiva basada en evidencia visual y geolocalizada. Las incidencias menores no detectadas no comprometen la seguridad, ya que siguen siendo gestionadas por los canales tradicionales.
4. **Mejora continua dual:**
  - **Inmediata (RAG):** El sistema aprende de cada validación humana en segundos, mejorando progresivamente su rendimiento
  - **Periódica (reentrenamiento):** Los datos validados alimentan el reentrenamiento de modelos, especializándolos en la realidad específica de Madrid

**Resultado:** La posible variabilidad en rendimiento está controlada mediante el diseño operativo del sistema y el enfoque en impacto preventivo sobre incidencias críticas. El valor del sistema no depende de alcanzar exhaustividad absoluta, sino de su capacidad de identificar suficientes incidencias graves para justificar la inversión a través del ahorro preventivo generado.

#### **4.3.2. Riesgo: incremento de costes de proveedores de IA**

**Escenario:** Los precios de los modelos de IA comerciales (OpenAI, Google) suben significativamente, comprometiendo la viabilidad económica.

#### **Mitigación implementada:**

1. **Arquitectura agnóstica:** Capacidad de cambiar de proveedor sin afectar la operativa del sistema
2. **Modelos open-source disponibles:** Ya existen modelos open-source (Llama 4, Gemini Flash) con costes 10-30x inferiores a los modelos Pro

3. **Proyección de mejora continua:** Cada año surgen nuevos modelos open-source más precisos y económicos, mejorando la ecuación coste-rendimiento
4. **Infraestructura propia planificada:** Hoja de ruta para alojar modelos en infraestructura del Ayuntamiento, eliminando costes variables por inferencia

**Resultado:** La dependencia de precios de terceros está limitada temporalmente y existen alternativas viables inmediatas.

#### 4.3.3. Riesgo: Continuidad del Proveedor de Plataforma

**Escenario:** SIGO cesa su actividad o deja de dar soporte a la plataforma.

**Mitigación contractual recomendada:**

1. **Cláusula de acceso al código fuente:** En caso de cese de actividad, el Ayuntamiento debe tener derecho a acceder al código fuente de la plataforma o a un sistema de depósito en garantía (*escrow*)
2. **Propiedad total de los datos:** Todos los datos (imágenes, incidencias, validaciones, modelos entrenados) son propiedad exclusiva del Ayuntamiento con capacidad de exportación completa en cualquier momento
3. **Arquitectura basada en estándares:** Uso de APIs REST, formatos abiertos (GeoJSON, JSON) y bases de datos estándar que facilitan la migración a otra plataforma si fuera necesario
4. **Documentación técnica completa:** Acceso a documentación de arquitectura, APIs y procesos que permitan la continuidad operativa

**Resultado:** La dependencia del proveedor está mitigada mediante garantías contractuales y diseño técnico abierto.

#### 4.3.4. Riesgo: Escalabilidad técnica

**Escenario:** La plataforma no puede escalar al volumen total de arbolado de.

**Mitigación técnica:**

1. **Arquitectura cloud nativa:** Diseño basado en servicios escalables (Azure y procesamiento paralelo) sin límites técnicos conocidos
2. **Validación incremental:** El despliegue se realizará de forma gradual, validando la escalabilidad en cada fase antes de ampliar cobertura
3. **Procesamiento batch optimizado:** El análisis no es en tiempo real, permitiendo distribuir la carga computacional de forma eficiente

**Resultado:** No se identifican límites técnicos significativos para el escalado completo del sistema.

#### **4.4. Especificaciones técnicas recomendadas para futuros pliegos**

Se recomienda que los futuros pliegos de contratación para sistemas de inspección inteligente de arbolado urbano incluyan las siguientes especificaciones técnicas mínimas que garanticen la calidad, independencia tecnológica y evolución continua del sistema.

## Anexo Técnico: Especificaciones para Sistemas de Inspección Inteligente de Arbolado Urbano

### 1. Arquitectura del Sistema y Modelo de Análisis

#### 1.1. Arquitectura de Análisis Híbrido Multi-Nivel (REQUISITO OBLIGATORIO)

La solución deberá implementar una **arquitectura de análisis en dos niveles jerárquicos** que optimice la relación cobertura-precisión-coste:

- **Nivel 1 - Análisis de Cobertura (Nivel de Vista):**
  - Análisis inicial de vistas direccionales de la escena completa mediante modelos de IA optimizados para cobertura
  - Detección de indicios de incidencias sin requerir identificación previa del activo específico
  - Capacidad de procesamiento de vistas direccionales extraídas de panorámicas 360° (Norte, Sur, Este, Oeste)
  - Objetivo: Maximizar cobertura de detección minimizando falsos negativos
- **Nivel 2 - Análisis de Precisión (Nivel de Activo):**
  - Análisis detallado de activos específicos (árboles, alcorques) en zonas identificadas como sospechosas en Nivel 1
  - Utilización de modelos de IA de alta precisión para caracterización detallada de incidencias
  - Identificación del activo responsable y clasificación específica del tipo de incidencia
  - Objetivo: Precisión máxima en caracterización y atribución de incidencias

#### 1.2. Arquitectura Agnóstica Multi-Proveedor (REQUISITO OBLIGATORIO)

La plataforma deberá ser **completamente agnóstica** respecto al proveedor de modelos de IA:

- **Demostración obligatoria:** El licitador deberá realizar una **demostración en vivo** de cambio de proveedor de IA (ej. de OpenAI a Google Gemini a Meta Llama) procesando un mismo conjunto de imágenes, completando el cambio en **menos de 4 horas**
- **Capacidad multi-proveedor validada:**
  - Funcionamiento demostrado con al menos **3 proveedores de IA comerciales diferentes** (ej. OpenAI, Google, Anthropic, Meta)
  - Integración funcional con al menos **1 modelo open-source** ejecutado vía API de terceros o infraestructura propia
  - Sistema de gestión de costes que permita seleccionar proveedor óptimo según criterio coste-rendimiento por caso de uso
- **Independencia operativa:**
  - Cambio de proveedor sin afectar la operativa del sistema (sin redepliegue de código, solo configuración)
  - Sin dependencias hard-coded en modelos específicos

- Capacidad de operar con múltiples proveedores simultáneamente según tipo de análisis

### **1.3. Procesamiento de Imágenes Panorámicas 360° (REQUISITO OBLIGATORIO)**

El sistema deberá ser capaz de procesar imágenes panorámicas 360° capturadas desde vehículos o mochilas portátiles, extraer automáticamente vistas direccionales (mínimo 4 vistas: N, S, E, O), detectar y recortar activos mediante modelos de detección de objetos (tipo YOLO o similar), y geolocalizar cada vista/activo con precisión suficiente.

**Formatos soportados:** Vídeo en bruto, panorámicas equirectangulares, imágenes 360° estándar

## **2. Sistema de Aprendizaje Continuo Dual (REQUISITO OBLIGATORIO)**

La plataforma deberá implementar un **sistema de aprendizaje en dos niveles** que capitalice el conocimiento generado por el Ayuntamiento:

### **2.1. Aprendizaje Inmediato (Retrieval-Augmented Generation - RAG)**

- **Mejora en tiempo real:** Cada validación humana (Human-in-the-Loop) debe enriquecer inmediatamente la base de conocimiento del sistema
- **Disponibilidad instantánea:** El conocimiento validado debe estar disponible para análisis posteriores en **menos de 60 segundos**
- **Persistencia:** Las validaciones humanas deben almacenarse como ejemplos curados que el sistema recupera para mejorar análisis similares futuros

### **2.2. Reentrenamiento Periódico**

- **Mejora estructural:** Capacidad de reentrenar modelos de IA con el dataset municipal curado mediante validaciones humanas
- **Especialización:** Los modelos deben especializarse progresivamente en la realidad específica de Madrid (especies arbóreas locales, tipología de alcorques, contexto urbano)
- **Frecuencia:** Posibilidad de reentrenamiento trimestral o bajo demanda según volumen de validaciones acumuladas

## **3. Sistema de Priorización Configurable y Autónomo (REQUISITO OBLIGATORIO)**

### **3.1. Scoring Parametrizable sin Dependencia del Proveedor**

La plataforma deberá incluir un **panel web de configuración de reglas de scoring** que permita al personal municipal:

- **Configuración autónoma:** Modificar ponderaciones, umbrales y reglas de priorización **sin intervención del proveedor tecnológico**
- **Ajuste granular:** Definir factores de criticidad específicos para cada tipo de incidencia (ej. tipo de señal obstruida, severidad de daño, contexto urbano)
- **Reglas contextuales:** Activar/desactivar bonificaciones según eventos temporales (post-temporales, zonas escolares)

- **Versionado:** Sistema de versiones de configuración de scoring con capacidad de reversión y trazabilidad de cambios

### 3.2. Simulación de Cambios (Sandbox)

El sistema debe incorporar un modo de prueba que permita aplicar nuevas reglas de scoring sobre incidencias históricas antes de activarlas en producción, una comparativa visual de cambios en clasificación de criticidad antes/después, y validación de impacto operativo de cambios en reglas.

## 4. Casos de Uso y Agentes Especializados

### 4.1. Casos de Uso Mínimos Obligatorios

El sistema deberá incluir agentes de IA validados para detectar, como mínimo, los siguientes casos de uso:

- **CU-01:** Interferencias con elementos de seguridad vial (señales, semáforos)
- **CU-02:** Interferencias con fachadas y elementos urbanos
- **CU-03:** Detección de arbolado en mal estado (integridad estructural, salud foliar)
- **CU-05:** Alcorques vacíos, con tocones o elementos impropios
- **CU-06:** Obstrucción de vías (gálibo vehicular y peatonal)
- **CU-08:** Nidos de gran tamaño (cotorras)

### 4.2. Desarrollo de Nuevos Agentes

El adjudicatario deberá comprometerse a desarrollar **nuevos agentes para casos de uso prioritarios** definidos por el Ayuntamiento durante la vigencia del contrato, incluyendo el desarrollo de agentes estacionales (riego, eventos climáticos), la extensión a otros elementos urbanos según necesidades identificadas, y la integración de modelos específicos para especies arbóreas locales de Madrid.

### 4.3. IA Explicativa y Transparencia Algorítmica (REQUISITO OBLIGATORIO)

Todos los agentes de IA deberán generar **justificaciones técnicas comprensibles** para cada incidencia detectada que incluyan:

- **Descripción de observaciones:** Qué elementos o condiciones observó la IA en las imágenes analizadas
- **Factores de decisión:** Qué factores específicos llevaron a clasificar la situación como incidencia (ej. porcentaje de obstrucción, tipo de elemento afectado, severidad observada)
- **Nivel de confianza:** Indicación del grado de certeza en la detección
- **Referencias visuales:** Señalamiento en las imágenes de las áreas o elementos relevantes para la detección

#### Requisitos de formato:

- Las justificaciones deberán ser comprensibles para personal técnico municipal sin formación específica en IA

- Deberán almacenarse estructuradamente junto con cada detección para auditoría y trazabilidad
- Deberán ser visibles en la interfaz de validación Human-in-the-Loop
- Los factores identificados deberán ser utilizables para la parametrización del sistema de scoring

## **5. Interfaz Human-in-the-Loop y Operativa**

### **5.1. Módulo de Validación Ágil (REQUISITO OBLIGATORIO)**

La interfaz de validación humana deberá permitir:

- **Validación eficiente:** Diseño optimizado para validar incidencias en **menos de 15 segundos** por registro
- **Visualización contextual:** Evidencia visual geolocalizada, vista 360º navegable, contexto urbano en mapa
- **Acciones disponibles:** Confirmar, corregir clasificación, descartar, reasignar criticidad, añadir observaciones
- **Captura de conocimiento:** Cada validación alimenta automáticamente el sistema de aprendizaje dual (RAG + reentrenamiento)

### **5.2. Trazabilidad de Validaciones**

El sistema debe proporcionar registro detallado de todas las validaciones (usuario, timestamp, acción realizada, cambios aplicados), exportación de histórico de validaciones para auditoría interna, y capacidad de generar informes de actividad de validación por usuario y período.

## **6. Experiencia Demostrable y Validación Operativa (CRITERIO DE VALORACIÓN)**

### **6.1. Experiencia en Contexto Municipal de Madrid (Altamente Valorado)**

Se valorará especialmente la **experiencia demostrable** del licitador en:

- **Piloto validado con datos reales de Madrid:** Ejecución de piloto con captura y análisis de viario madrileño (kilómetros recorridos, incidencias detectadas, validaciones realizadas con técnicos municipales)
- **Conocimiento del contexto GIS municipal:** Experiencia en integración con sistemas GIS del Ayuntamiento, conocimiento de estructura de datos MINT, tablas de activos municipales
- **Validación de casos de uso específicos:** Demostración de detección validada de incidencias reales identificadas por técnicos del Ayuntamiento (no solo pruebas de concepto)

### **6.2. Capacidad en Gemelo Digital Urbano (Valorado)**

Se valorará experiencia demostrable en integración de sistemas de inspección con plataformas de gemelo digital urbano, sincronización bidireccional de datos entre sistemas de análisis y gemelos digitales, y visualización 3D de incidencias en contexto urbano completo.

### **6.3. Trayectoria en Mantenimiento Urbano y GIS (Valorado)**

Se valorará experiencia acreditada de **más de 15 años** en gestión de mantenimiento de infraestructuras urbanas, desarrollo de sistemas GIS para administraciones públicas, y conocimiento profundo de normativas municipales de arbolado y zonas verdes.

## **7. Propiedad del Dato y Portabilidad (REQUISITO OBLIGATORIO)**

### **7.1. Propiedad Integral de Datos**

- **Propiedad exclusiva del Ayuntamiento:** Todos los datos generados (vídeos, imágenes, detecciones, validaciones, modelos re-entrenados) son propiedad íntegra y perpetua del Ayuntamiento de Madrid
- **Sin restricciones de uso:** El Ayuntamiento puede usar, modificar, compartir o comercializar los datos sin limitación ni coste adicional

### **7.2. Portabilidad Total y API de Exportación**

- **API de exportación masiva:** Capacidad de exportar la totalidad de los datos en cualquier momento
- **Formatos estándar abiertos:** GeoJSON (datos geográficos), JSON (datos estructurados), JPEG/PNG (imágenes)
- **Acceso directo al repositorio:** El Ayuntamiento dispondrá de acceso programático directo a la base de datos
- **Sin penalización:** Exportación sin coste adicional ni restricciones temporales o de volumen

## **8. Interoperabilidad y Estándares (REQUISITO OBLIGATORIO)**

### **8.1. API REST Completa**

El sistema debe proporcionar documentación completa de API REST que permita, como mínimo, consulta de incidencias con filtros avanzados (tipo, fecha, criticidad, ubicación, estado), recuperación de detalles completos de incidencias incluyendo evidencias visuales, actualización de estado de incidencias desde sistemas externos, y webhooks configurables para notificaciones en tiempo real.

### **8.2. Integración con MiNT y Sistemas Municipales (REQUISITO OBLIGATORIO)**

La solución deberá garantizar integración completa con los sistemas municipales de gestión:

- **Integración obligatoria con MiNT:** Capacidad demostrable de exportación a las estructuras de datos estándar de la Plataforma de Madrid Inteligente (MiNT)
- **Sincronización automática con sistemas GIS:** Conectividad bidireccional con sistemas de información geográfica municipales
- **Compatibilidad con formatos estándar:** Soporte para formatos de datos y APIs utilizados por los sistemas corporativos del Ayuntamiento

## **9. Demostración Funcional Completa del Sistema (REQUISITO OBLIGATORIO)**

El licitador deberá realizar una **demostración funcional completa** del sistema operativo que incluya todos los componentes críticos:

### 9.1. Procesamiento Completo de Imágenes

- Carga de imágenes panorámicas 360°
- Extracción automática de vistas direccionales
- Detección y recorte de activos mediante modelos de detección de objetos
- Geolocalización precisa de detecciones

### 9.2. Análisis con Agentes de IA y Justificaciones

- Ejecución del análisis híbrido de dos niveles (vista + activo)
- Generación de detecciones con justificaciones técnicas comprensibles
- Visualización de los factores de decisión identificados por la IA
- Asignación automática de scoring de criticidad

### 9.3. Interfaz Human-in-the-Loop Operativa

- Validación de incidencias por personal técnico del Ayuntamiento
- Demostración de flujo completo: confirmar, corregir y descartar
- Tiempo de validación por incidencia < 15 segundos en casos evidentes
- Visualización de justificaciones técnicas en la interfaz

### 9.4. Sistema de Scoring Configurable

- Modificación en vivo de parámetros de scoring (ponderaciones, umbrales)
- Ejecución de simulación (sandbox) sobre incidencias históricas
- Visualización del impacto de cambios en la clasificación de criticidad
- Demostración de autonomía sin intervención del proveedor

### 9.5. Aprendizaje Inmediato (RAG)

- Demostración de cómo una validación humana enriquece la base de conocimiento
- Análisis de una incidencia similar posterior mostrando la mejora
- Disponibilidad del conocimiento validado en menos de 60 segundos

### 9.6. Cambio de Proveedor de IA

- Cambio de proveedor de IA (ej. OpenAI → Google Gemini) procesando el mismo conjunto de imágenes
- Tiempo de cambio: menos de 4 horas
- Sin redepliegue de código, solo configuración

### Requisitos de la demostración:

- Deberá realizarse con **imágenes reales del viario de Madrid** proporcionadas por el Ayuntamiento
- El sistema deberá estar completamente operativo, no simulaciones ni prototipos
- Personal técnico municipal deberá poder interactuar directamente con la interfaz

## **10. Seguridad y Cumplimiento Normativo (REQUISITO OBLIGATORIO)**

### **10.1. Cumplimiento del Esquema Nacional de Seguridad (ENS)**

El sistema deberá cumplir con los requisitos del Esquema Nacional de Seguridad (Real Decreto 311/2022, de 3 de mayo), aplicando las medidas de seguridad correspondientes a la categoría que se determine en el análisis de riesgos del Ayuntamiento de Madrid.

#### **Medidas de seguridad requeridas:**

- **Identificación y autenticación:** Control de acceso basado en roles con autenticación robusta para todos los usuarios del sistema
- **Protección de datos en reposo:** Cifrado de datos almacenados (mínimo AES-256) para imágenes, vídeos y bases de datos
- **Protección de datos en tránsito:** Comunicaciones cifradas mediante TLS 1.3 o superior entre todos los componentes del sistema
- **Auditoría y trazabilidad:** Registro completo de accesos, modificaciones y operaciones críticas con retención mínima de 2 años
- **Copias de seguridad:** Política de backup automatizada con almacenamiento redundante y geográficamente distribuido
- **Gestión de incidentes:** Procedimientos documentados de respuesta ante incidentes de seguridad

### **10.2. Cumplimiento del Reglamento General de Protección de Datos (RGPD)**

Además del ENS, el sistema deberá cumplir con el RGPD en el tratamiento de imágenes del espacio público:

- **Anonimización automática:** Desenfoque de rostros y matrículas de forma automática antes del almacenamiento definitivo
- **Minimización de datos:** Retención limitada de imágenes según finalidad, con eliminación automática tras el análisis
- **Derecho de información:** Señalización en vehículos de captura y publicación de zonas y períodos en web municipal
- **Evaluación de impacto (EIPD):** Documentación completa de la evaluación de impacto en protección de datos

### **10.3. Certificaciones y Auditorías**

El licitador deberá acreditar:

- **Certificación ISO 27001** en seguridad de la información (propia o del proveedor de infraestructura cloud)
- **Auditoría ENS:** Compromiso de realización de auditoría ENS en un plazo máximo de 12 meses desde el despliegue
- **Ubicación de datos:** Garantía de que todos los datos se almacenan en servidores ubicados en la Unión Europea

## **5. Conclusiones**

Este capítulo recoge los aprendizajes clave del piloto, explica cómo puede evolucionar el sistema y confirma que GreenView® está listo para su continuidad operativa. Se destacan nueve descubrimientos importantes surgidos del feedback del Ayuntamiento y se reconocen los aspectos que quedaron fuera del alcance del piloto. La conclusión confirma la viabilidad técnica y económica demostrada, el cambio de paradigma hacia gestión proactiva, y el valor del sistema en complementar (no sustituir) los mecanismos actuales manteniendo control humano sobre decisiones y evolución tecnológica.

El piloto de GreenView® ha demostrado que el sistema funciona. Los resultados son medibles. En casos prioritarios como el arbolado en mal estado alcanzamos un F1-Score superior al 75%, y en interferencias con seguridad vial detectamos 7 de cada 10 casos reales. Lo hemos validado con casos reales documentados por el Ayuntamiento y medido con métricas técnicas estándar. El proceso de validación humana es ágil, los técnicos pueden revisar incidencias en segundos. La arquitectura agnóstica funciona, hemos podido cambiar de proveedor de IA en cuestión de horas. Y lo que es fundamental, el sistema mejora con su uso. Cada validación que realiza un técnico alimenta el aprendizaje continuo mediante RAG y reentrenamiento. Cuanto más se utilice, mejor será su rendimiento.

Durante el proceso han surgido aprendizajes importantes gracias al feedback continuo del Ayuntamiento. Elementos que no estaban en el diseño inicial pero que han resultado fundamentales: el modelo híbrido de dos niveles, la arquitectura agnóstica que permite cambiar de proveedor, la incorporación de modelos open-source, el sistema de scoring configurable. Estos descubrimientos han mejorado el sistema.

Este capítulo recoge esos aprendizajes, explica cómo puede evolucionar el sistema y propone pasos concretos si el proyecto continúa.

### **5.1. Lo que hemos aprendido**

Durante el piloto hemos descubierto aspectos que no eran evidentes al principio. Algunos confirmaron nuestras hipótesis, mientras que otros nos llevaron a replantear el enfoque.

#### **5.1.1. La captura no es un problema técnico, es un problema operativo**

Las cámaras 360° y los GPS de precisión son tecnología madura y fiable. El reto no ha estado en el hardware, sino en pensar cómo se iba organizar la operativa de captura de forma que sea sostenible.

- **Planificar las rutas de captura no es trivial:** Hay que decidir qué calles recorrer primero, con qué frecuencia, evitando duplicidades innecesarias. Además, hay restricciones de tráfico, zonas de acceso limitado y la necesidad de coordinar con otras operaciones municipales para no interferir.
- **El volumen de datos es considerable:** Cada vehículo produciría unos 300 GB de imágenes al día. Esto obliga a tener protocolos claros de volcado, almacenamiento y gestión del ciclo de vida de los datos.

La captura física sigue siendo viable, pero el reto principal es organizativo, coordinar rutas, gestionar el volumen de datos y mantener procesos consistentes en una operativa municipal real.

Por eso, el piloto se ha centrado en lo realmente crítico, validar la arquitectura tecnológica, los algoritmos de IA y los flujos de trabajo. Utilizando imágenes de Google Street View hemos podido demostrar el funcionamiento y la viabilidad de los componentes fundamentales del sistema (detección de IA, scoring, validación humana) sin necesidad de montar una infraestructura completa de captura en campo. Esta aproximación ha resultado clave, nos ha permitido validar lo esencial primero.

### **5.1.2. El análisis en dos niveles hace viable el proyecto**

Uno de los descubrimientos más importantes del piloto fue el modelo híbrido de dos niveles. Este enfoque no solo optimiza costes, sino que ha resultado fundamental para hacer el proyecto operativamente viable, permite planificar mejor las rutas de captura, ampliar los casos de uso y tener una visión más completa.

- **El modelo híbrido que hemos desarrollado**, el sistema funciona en dos fases. Primero, un análisis rápido de las vistas panorámicas con modelos más sencillos y baratos busca indicios de incidencias en la escena completa. Esto nos da cobertura amplia con bajo coste. Cuando detectamos algo sospechoso, activamos la segunda fase: un análisis detallado de los activos concretos (árboles, alcorques) con modelos más precisos y caros.

Esta estrategia reduce el coste porque solo usamos los modelos caros cuando realmente hace falta. La mayoría de las imágenes se procesan con modelos baratos, y solo una fracción pequeña necesita el análisis de precisión.

Sin este modelo híbrido, el proyecto tal vez no sería viable económicamente. Analizar millones de imágenes con modelos caros tendría un coste muy elevado. Con el modelo híbrido, podemos escalar sin que los costes se disparen.

Además, este enfoque encaja bien con los modelos open-source, usamos modelos abiertos para el primer filtro (donde no necesitamos máxima precisión) y guardamos los modelos comerciales para cuando queremos estar seguros.

### **5.1.3. Detectar no es suficiente, hay que priorizar**

Otro de los aprendizajes importantes del piloto es que la detección masiva de incidencias cobra sentido cuando puedes priorizarlas. El sistema puede identificar cientos o miles de incidencias, pero el valor real está en saber cuáles atender primero. Sin priorización, se genera una lista muy extensa de alertas complicadas de gestionar.

Por ello hemos desarrollado un sistema de scoring que ordena las incidencias por criticidad, permitiendo saber qué es urgente y qué puede esperar.

Lo clave es que cada detección viene con su justificación, la IA explica qué vio y por qué lo considera una incidencia (tipo de señal tapada, porcentaje de obstrucción, si está en una zona escolar, etc.). Así, el Ayuntamiento puede usar esa justificación para parametrizar el scoring. Si

la IA dice “señal de STOP obstruida al 80%”, el Ayuntamiento decide cuántos puntos vale eso. Si dice “árbol cerca de un colegio”, el Ayuntamiento decidirá si eso suma puntos o no.

Esta parametrización no la hace el proveedor, la hace el Ayuntamiento. Y esto es importante: el Ayuntamiento mantiene el control sobre qué es prioritario y qué no, pudiendo recalibrar el sistema cuando sea necesario. Si hay un temporal, se ajustan los pesos para priorizar riesgo de caída. Si cambian las prioridades municipales, se modifican los criterios. El sistema se adapta a las necesidades reales.

#### **5.1.4. El valor está en prevenir, no en detectarlo todo**

Al principio del piloto teníamos la presión de alcanzar métricas muy altas: precisión del 95%, recall del 90%. Durante la validación descubrimos que las métricas son importantes, pero no son el único criterio de éxito.

El valor real está en detectar suficientes incidencias críticas para justificar el sistema. El Ayuntamiento ya gestiona decenas de miles de incidencias al año de forma reactiva (denuncias del 010, avisos de emergencia, inspecciones programadas). GreenView® no sustituye esos mecanismos, los complementa añadiendo una capa de anticipación.

El sistema se justifica si cada año detecta un volumen relevante de incidencias graves que permiten actuar antes de que se conviertan en problemas mayores. No hace falta detectar el 100% de los alcorques vacíos si detectamos un porcentaje de los árboles con riesgo de caída.

Los falsos positivos son manejables porque validarlos cuesta poco. Las incidencias menores que se nos escapan seguirán siendo gestionadas por los canales tradicionales. Lo importante es no perder incidencias graves, una rama peligrosa detectada a tiempo, una señal de STOP que se libera antes de que cause un accidente, un árbol en mal estado que se interviene preventivamente. Esas detecciones tienen un valor que supera el coste del sistema.

El sistema está pensado para este criterio. El modelo híbrido da cobertura amplia. El proceso Human-in-the-Loop gestiona los falsos positivos rápidamente. El scoring filtra y prioriza lo crítico. Y cada validación mejora el sistema inmediatamente. El éxito se mide en impacto preventivo real, no en cobertura absoluta.

#### **5.1.5. Cada validación humana mejora el sistema**

La independencia tecnológica no es solo usar software libre. Es tener control sobre cómo mejora el sistema. El proceso es simple. La IA detecta una posible incidencia. Un técnico municipal la revisa y la confirma, corrige o descarta. Esa validación no se pierde, se guarda y el sistema la usa de dos formas.

Primero, mejora inmediata. En cuestión de segundos, esa validación se añade a la base de conocimiento del sistema (vía RAG). Si mañana aparece una imagen similar, el sistema ya sabe cómo tratarla.

Segundo, mejora a largo plazo. Todas las validaciones se acumulan y pueden usarse para ajustar y mejorar el sistema. El objetivo es que con el tiempo el sistema se especialice en Madrid, sus especies de árboles, sus tipos de alcorques, sus características urbanas.

Los modelos comerciales están entrenados con datos genéricos. Con las validaciones del Ayuntamiento, el sistema va acumulando conocimiento específico. Cada técnico que valida una incidencia aporta su criterio, su experiencia de campo. Ese conocimiento queda registrado y puede aprovecharse para mejorar las detecciones futuras.

#### **5.1.6. La tecnología tiene que ser fácil de usar**

Hemos buscado que la tecnología se integre de forma natural en el trabajo diario, priorizando la usabilidad y la facilidad de uso.

- **Exportación a sistemas existentes:** El sistema permitirá exportar los resultados en formato tabular para su integración con MINT. Es compatible con estructuras de datos GIS y tendrá una API abierta para futuras conexiones.
- **Interfaz pensada para técnicos municipales:** La interfaz de validación permite revisar incidencias en poco tiempo. La visualización es intuitiva (mapas de calor, vistas 360° navegables) y no requiere formación técnica especializada.
- **Potenciar al técnico, no sustituirlo:** La IA propone, el humano decide. El sistema explica sus detecciones para que el técnico pueda validarlas con criterio. El objetivo no es automatizar completamente, sino liberar tiempo de tareas repetitivas para que los técnicos puedan dedicarse a lo que aporta más valor.

Hemos intentado que la interfaz sea lo suficientemente sencilla como para que los técnicos puedan integrarla en su trabajo sin complicaciones adicionales.

#### **5.1.7. Tres cambios importantes surgieron del feedback**

Al empezar el piloto teníamos un diseño inicial. Durante la validación con los técnicos del Ayuntamiento descubrimos que necesitábamos cambiar cosas importantes.

**Primero, añadir análisis de vistas panorámicas.** El diseño inicial solo analizaba árboles individuales. Si el sistema no detectaba un árbol, no lo analizaba. De ahí salió el modelo híbrido: primero analizamos las vistas panorámicas completas, luego refinamos con análisis de activos específicos. Resultando en una cobertura amplia + precisión.

**Segundo, incorporar modelos open-source.** Al principio solo usábamos modelos comerciales. El Ayuntamiento planteó una preocupación legítima: ¿qué pasa si los proveedores suben los precios? Rediseñamos la arquitectura para que fuera agnóstica: capaz de trabajar con cualquier proveedor. Probamos que podíamos cambiar de OpenAI a Google a Meta con facilidad. Esto abre una ruta de independencia tecnológica.

**Tercero, añadir sistema de priorización (scoring).** Al inicio considerábamos que el objetivo era detectar y reportar todas las incidencias. Los técnicos señalaron que, sin priorización, el sistema no sería práctico. A partir de ahí desarrollamos el scoring.

Ninguno de estos tres cambios estaba en el diseño original. Surgieron del feedback continuo con los técnicos del Ayuntamiento. Trabajar codo con codo con quien va a usar el sistema fue clave.

#### **5.1.8. Centrarse en lo importante primero**

Al diseñar GreenView® hablamos de desarrollar 5 agentes: Inspector de Arbolado, Inspector de Alcorques, Inspector de Limpieza, Gestor de Inventario (actualizar atributos GIS) y Auditor de Operaciones (verificar trabajos de contratatas).

Durante el piloto descubrimos que no todos eran igual de urgentes. Los tres primeros (detección de incidencias) generan valor inmediato: cada detección es una posible actuación preventiva. Los otros dos (inventarios y auditoría) son valiosos, pero no son críticos para la operativa diaria inmediata.

Decidimos concentrar los recursos del piloto en perfeccionar los tres agentes de detección y dejar los otros para fases posteriores. Esto nos permitió profundizar en lo esencial y validar más rápido.

Los agentes de Inventario y Auditoría siguen en la hoja de ruta, pero para cuando el sistema esté consolidado y tengamos flujos de trabajo estables con las contratatas. Es mejor tener 3 componentes excelentes que generen valor real que 5 mediocres que añadan complejidad sin impacto claro.

### 5.1.9. Lo que quedó fuera del piloto respecto al DDE

Es importante ser transparentes sobre los aspectos del documento de definición de éxito original que no se completaron durante el piloto. Esto no representa un fallo del proyecto, sino una priorización estratégica que permitió concentrar recursos en validar lo esencial. Estas son algunas de las cuestiones que no se completaron:

- **Gemelo digital con ESRI:** El plan original contemplaba una integración completa con la plataforma ESRI y el desarrollo de un gemelo digital urbano avanzado. Esta funcionalidad no se implementó en el piloto. En su lugar, se validó la arquitectura de datos y la capacidad de exportación en formatos compatibles con GIS, dejando la integración completa para fases posteriores cuando el valor del sistema esté consolidado.
- **Integración con MINT:** El análisis de viabilidad de integración con la plataforma MINT se realizó mediante análisis técnico, sin establecer reuniones formales con el equipo responsable de MINT. Se validó la capacidad de exportación de datos en formatos compatibles, pero la conexión bidireccional vía API queda pendiente para futuras fases del proyecto.
- **Programa extensivo de capturas:** No ejecutamos un programa de capturas de campo. El piloto se centró en validar los algoritmos y flujos de trabajo con imágenes de Street View y capturas limitadas.
- **Agentes de Inventario y Auditoría:** Como se explicó en el punto 5.1.8, estos agentes quedaron fuera del alcance del piloto para concentrar esfuerzos en los agentes de detección de incidencias. Aunque no estaban en el DDE sí nos hubiera gustado implementarlo ya que durante el piloto se ha visto su valor.
- **Formación estructurada:** Aunque se realizaron sesiones de trabajo con los técnicos del Ayuntamiento, no se desarrolló un programa formal de formación y materiales de capacitación como estaba previsto en el DDE.

- **Sistema de alertas y notificaciones:** El DDE contemplaba un sistema automatizado de alertas. Esta funcionalidad queda pendiente para fases posteriores, cuando se defina el flujo operativo completo con las contratas y se determine cómo deben gestionarse las notificaciones en el día a día.

Estas priorizaciones surgieron de la dinámica del piloto y del feedback del Ayuntamiento. En lugar de intentar completar todas las funcionalidades del plan original de forma superficial, se optó por validar rigurosamente los componentes críticos: la detección mediante IA, el sistema de scoring, la validación humana, y la arquitectura agnóstica. Esta decisión ha resultado acertada: ahora sabemos que el núcleo del sistema funciona, y las integraciones técnicas pueden abordarse con mayor seguridad en fases posteriores.

## **5.2. Qué podría venir después**

El piloto ha demostrado que el sistema es suficientemente eficiente para su despliegue operativo. La capacidad de detección, el proceso de validación y la arquitectura técnica están validados y son funcionales. Hay camino de mejora, como en cualquier sistema, pero no son necesidades críticas.

El Ayuntamiento ya tiene mecanismos que funcionan para gestionar el arbolado: avisos del 010, inspecciones, actuaciones de emergencia. GreenView® añade algo distinto, poder detectar problemas antes de que alguien los reporte. No sustituye nada, añade una capa de anticipación. Eso es lo que hemos probado que funciona.

Si se siguiera usando, el sistema seguiría mejorando de forma natural. Cada validación de un técnico enriquece el conocimiento del sistema. Las capturas repetidas acumulan información temporal. Con datos de 2-3 años, se podrían identificar tendencias. Pero estas son evoluciones deseables, no requisitos previos. El sistema ya es operativamente viable.

La arquitectura agnóstica permite al Ayuntamiento mantener flexibilidad tecnológica. Si fuera necesario cambiar de proveedor de IA, la transición puede realizarse sin grandes complicaciones. La migración hacia modelos open-source es viable cuando se considere oportuno, aunque el sistema actual con modelos comerciales cumple con los requisitos operativos.

## **5.3. Conclusión final**

El piloto ha demostrado tanto la viabilidad técnica como económica del sistema. GreenView® detecta incidencias críticas de forma efectiva, los técnicos pueden validar las propuestas de la IA de manera ágil, y la arquitectura técnica soporta la operativa. Lo más importante: hemos probado que se pueden detectar problemas antes de que alguien los reporte, permitiendo una gestión proactiva del arbolado urbano.

El enfoque del sistema está en identificar incidencias críticas. Dado que el volumen de detecciones será considerable, hemos desarrollado un sistema de priorización (scoring) que permite gestionar eficientemente las alertas, concentrando los recursos en lo verdaderamente urgente. Este equilibrio entre cobertura y priorización es lo que hace viable el sistema operativamente.

Si el proyecto continúa, el sistema mejorará de forma natural con su uso. Cada validación que realiza un técnico enriquece el conocimiento del sistema mediante RAG y alimenta futuros

reentrenamientos. Con el tiempo, esto genera una base de datos única que permite identificar patrones, anticipar tendencias estacionales y refinar los criterios de detección específicos para Madrid.

Las posibilidades de evolución son amplias. Se pueden incorporar nuevos sistemas de captura —bicicletas, patines, mochilas— para llegar a zonas donde los vehículos no acceden: senderos de parques, áreas peatonales, zonas de acceso restringido. Los modelos open-source están evolucionando rápidamente y estimamos que en 1-2 años alcanzarán el rendimiento de los modelos Pro actuales, lo que permitirá optimizar costes manteniendo la calidad. El sistema puede ampliarse a otros casos de uso siguiendo la misma arquitectura validada, y el rendimiento seguirá mejorando mediante ajustes en las parametrizaciones del scoring según las prioridades municipales.

La integración futura con el gemelo digital abre oportunidades bidireccionales interesantes. Por un lado, GreenView® puede enriquecer el inventario municipal con datos actualizados automáticamente: estados de árboles, cambios en alcorques, evolución temporal de elementos verdes. Por otro lado, el sistema se volvería más eficiente al consultar el inventario existente: conocer las especies, tamaños históricos, actuaciones previas permite afinar las detecciones y reducir falsos positivos.

GreenView® complementa los mecanismos actuales del Ayuntamiento sin sustituirlos. Los técnicos mantienen el control sobre las decisiones, las prioridades y la evolución tecnológica del sistema. El piloto ha validado que esto funciona, y las bases están puestas para seguir construyendo sobre lo aprendido.