

# Cómo aprovechar el poder de los metadatos de las escenas

Más información, mayor eficiencia y un mejor conocimiento de la situación

Enero 2024

# Resumen

En el contexto de la videovigilancia, los metadatos aportan una descripción textual del contenido del vídeo, como los objetos de interés visibles, o aportan una descripción detallada de la escena en sí. Esta descripción también puede incluir atributos asociados a los objetos o a la escena, como los colores de los vehículos o las prendas de ropa, ubicaciones exactas o la dirección de movimiento. Los metadatos se generan en tiempo real, directamente en la cámara o a través de otro componente del sistema capaz de ejecutar analítica de vídeo.

Los metadatos también aportan contexto asociado a los eventos y agilizan la clasificación y la búsqueda en grandes volúmenes de imágenes. El resultado es la posibilidad de usar funciones relacionadas con tres grandes áreas:

- **Búsquedas forenses postevento.** Para realizar búsquedas de objetos o eventos de interés a partir de diferentes parámetros y acotar la búsqueda a un número reducido de candidatos. Los datos sobre clasificación de objetos abren la puerta a búsquedas con un gran abanico de detalles.
- **Uso en tiempo real.** Para ayudar a los operadores a responder con celeridad ante posibles cambios en una situación o aportar información útil para la toma de decisiones o para la automatización de acciones.
- **Identificación de tendencias, patrones y datos.** El IoT y las plataformas de eficiencia operativa pueden utilizar los metadatos para obtener información sobre el número de visitantes, la velocidad, los flujos de tráfico y otros tipos de datos automáticos.

Algunas cámaras pueden descodificar audio para recuperar metadatos de audio. Los patrones de sonido específicos pueden detectarse y etiquetarse de una forma similar a como se detectan y etiquetan las clases de objetos en el vídeo. Un sistema de reconocimiento de audio, por ejemplo, podría identificar situaciones de violencia verbal o detectar el ruido de cristales rotos.

La combinación de metadatos de diferentes fuentes, por ejemplo, de imagen, de audio, vinculadas a la actividad o relacionadas con procesos, aporta mucha más información que cada una de las fuentes por separado. Para una integración sencilla de los metadatos, es imprescindible usar protocolos abiertos y normas sectoriales.

# Índice

1	Introducción	4
2	¿Qué son los metadatos?	4
3	Generación de metadatos en el dispositivo	4
4	Aplicaciones	5
	4.1 Uso en tiempo real para actuar al instante	5
	4.2 Búsqueda forense	5
	4.3 Identificación de tendencias y patrones para obtener información	6
5	¿Dónde se utilizan los metadatos?	6
6	¿Cómo se transmiten los metadatos?	7
7	Metadatos de audio	9
8	Combinación de metadatos de varias fuentes	9

# 1 Introducción

Los metadatos son la base para obtener información a partir del vídeo, ya que asignan un sentido digital al contenido de vídeo mediante la descripción de los principales detalles de la escena. Los metadatos permiten localizar y evaluar rápidamente la información clave en enormes volúmenes de vídeo y tomar las medidas oportunas. Y esto explica que su papel sea cada día más decisivo para garantizar la eficiencia de las operaciones de seguridad y comerciales.

Este documento técnico gira en torno a los metadatos, tanto en el contexto de la vigilancia como desde el punto de vista de la eficiencia operativa. Explica con detalle las ventajas de los metadatos y su uso en sistemas de gestión de vídeo y otras aplicaciones.

## 2 ¿Qué son los metadatos?

Los metadatos son datos sobre otros datos. En el contexto de la videovigilancia, los metadatos ofrecen una descripción textual del contenido del vídeo, por ejemplo los objetos de interés visibles, o aportan una descripción detallada de la escena en sí. Esta descripción puede incluir atributos asociados a los objetos o a la escena, como los colores de los vehículos o las prendas de ropa, ubicaciones exactas o la dirección de movimiento. Los metadatos se generan en tiempo real, directamente en la cámara o a través de otro componente del sistema capaz de ejecutar analítica de vídeo.

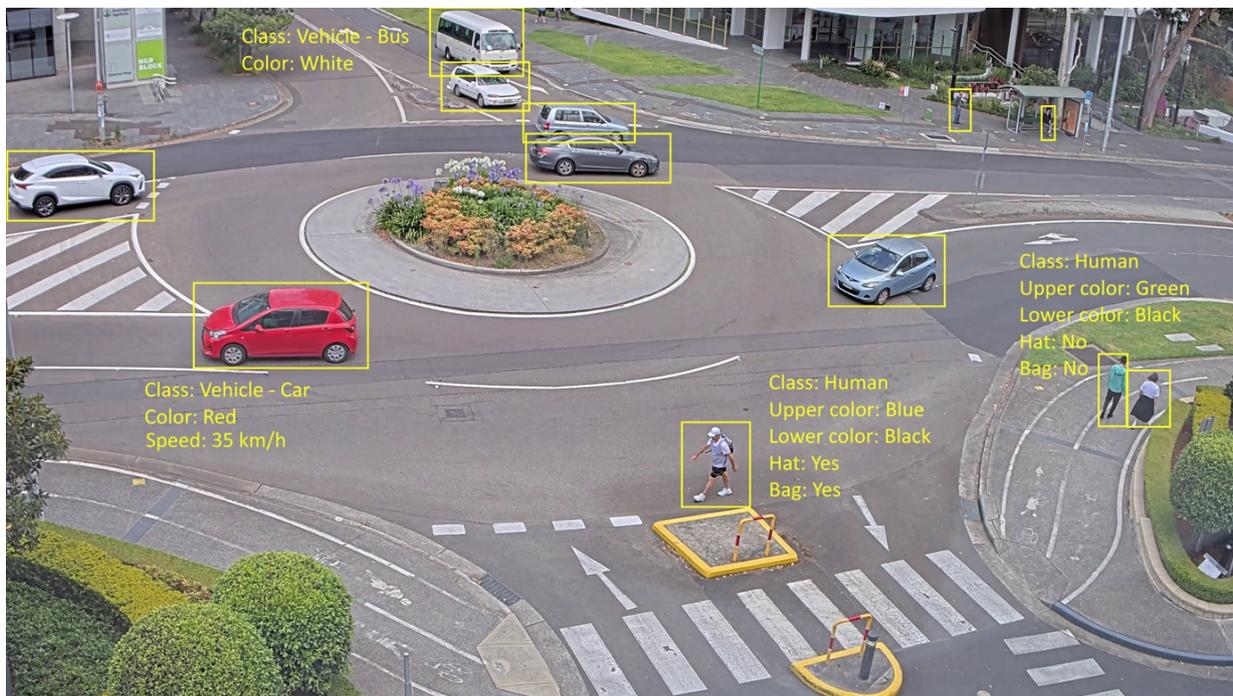


Figure 1. Ejemplo de un fotograma de vídeo en el que se detectan objetos de interés y se analizan para crear metadatos.

## 3 Generación de metadatos en el dispositivo

La analítica de vídeo avanzada normalmente estaba vinculada a un servidor, ya que requería más capacidad de procesamiento de la que podía ofrecer un dispositivo local. En los últimos años, la mejora de los algoritmos y el aumento de la potencia de procesamiento de los dispositivos locales han abierto la puerta

a la ejecución de analítica avanzada en el dispositivo. Por tanto, los metadatos pueden generarse en el dispositivo y otras soluciones de analítica pueden utilizarlos directamente en el dispositivo. Además, la transmisión de vídeo y el flujo de metadatos también pueden enviarse al software de gestión de vídeo u otra aplicación para su procesamiento.

La analítica en el extremo tiene acceso a vídeo sin comprimir y con una latencia muy baja, lo que se traduce en aplicaciones rápidas en tiempo real y, además, ahorra el coste extra y la complejidad de transferir todo el vídeo para su procesamiento en otro punto del sistema. Y no solo eso: la analítica en el extremo también tiene unos costes de hardware e implantación inferiores, ya que el sistema no necesita tantos recursos de servidores.

La generación de metadatos en el extremo permite también extraer datos del vídeo sin perder información a causa de la compresión o la transmisión. El resultado son unos metadatos más precisos y unos análisis más exactos del contenido de vídeo. Y es que, a mayor calidad de imagen, mejor es la calidad de los metadatos.

## 4 Aplicaciones

Los metadatos no solo aportan detalles sobre los objetos de una escena, sino también el contexto asociado a los eventos, y además agilizan la clasificación y la búsqueda en grandes volúmenes de imágenes. El resultado es la posibilidad de usar funciones en tres grandes áreas: búsquedas forenses postevento, aplicaciones en tiempo real o identificación de tendencias, patrones y datos.

### 4.1 Uso en tiempo real para actuar al instante

El uso de los metadatos en tiempo real puede ayudar a los operadores a responder con celeridad ante posibles cambios en una situación. También pueden resultar útiles para ayudar en la toma de decisiones o automatizar acciones. Un sistema de analítica local en tiempo real basado en metadatos de alta calidad puede ayudarle a proteger personas, sitios y edificios, y a reducir el riesgo de daños intencionados o accidentales. La posibilidad de detectar, verificar y evaluar rápidamente las amenazas permite darles respuesta de una forma eficaz.

### 4.2 Búsqueda forense

Los metadatos abren la puerta a unas búsquedas rápidas y eficientes de objetos o eventos de interés. Los investigadores pueden ahorrarse muchas horas de trabajo, especialmente al buscar en grandes volúmenes de vídeo de diferentes fuentes. Puede realizar búsquedas de objetos, como personas o vehículos, a partir de diferentes parámetros y acotar así la búsqueda a un número reducido de candidatos. Los parámetros de búsqueda pueden incluir, por ejemplo, el movimiento, el periodo de tiempo o las características del objeto.

*Tabla 4.1 El uso de diferentes categorías de metadatos agiliza las búsquedas.*

Categoría de metadatos	¿Qué se detecta?	Ejemplos
Movimiento	Cómo se mueve un objeto	Dirección, velocidad, otro comportamiento
Periodo de tiempo	Cuándo aparece un objeto	Día de la semana, hora del día, tiempo de permanencia
Lugar	Dónde está el objeto	Lugar, campo de visión de la cámara

Tabla 4.1. El uso de diferentes categorías de metadatos agiliza las búsquedas. (Continuación)

Categoría de metadatos	¿Qué se detecta?	Ejemplos
Clasificación del objeto	Qué tipo de objeto es	Persona, vehículo (coche, autobús, camión, bicicleta/motocicleta)
Atributos del objeto	Qué características tiene	Ropa, accesorios como sombreros o bolsas, características físicas como el color de la ropa

Aunque solo tenga acceso a una categoría de metadatos, como el periodo de tiempo, esta información puede resultar vital para obtener los resultados que necesita.

Los metadatos sobre movimiento permiten realizar búsquedas basadas en la velocidad relativa de un objeto y su dirección de movimiento. Los datos sobre clasificación de objetos abren la puerta a búsquedas con un mayor abanico de detalles. Las cámaras con unidades de procesamiento con deep learning (DLPU) normalmente pueden obtener metadatos enriquecidos, con una clasificación de objetos más detallada, lo que permite buscar, por ejemplo, un camión verde o una persona con una chaqueta azul.

### 4.3 Identificación de tendencias y patrones para obtener información

El IoT y las plataformas de eficiencia operativa pueden utilizar los metadatos para obtener información sobre el número de visitantes, la velocidad, los flujos de tráfico y otros tipos de datos automáticos. Los datos, una vez analizados, pueden dar pie a determinadas acciones.

## 5 ¿Dónde se utilizan los metadatos?

El uso de metadatos para entender las características y el contenido de una escena aporta numerosas ventajas. A continuación, detallamos dónde suelen utilizarse los metadatos.

**Aplicaciones en el extremo.** La analítica ejecutada en local en la cámara puede aplicar filtros lógicos y reglas a la información sobre los objetos de la escena. En este contexto, la analítica puede activar acciones basadas en umbrales definidos o comportamientos concretos, por ejemplo, el control de una cámara PTZ cuando se produce una detección y el movimiento de una persona en la escena.

**Sistemas de gestión de vídeo (VMS).** En el ámbito de la videovigilancia, los metadatos se han utilizado a menudo en un software de gestión de vídeo para mostrar superposiciones visuales vinculadas a posibles objetos de interés en una escena. Con los avances en los algoritmos de detección y clasificación de objetos, ahora los operadores también pueden localizar objetos de interés a partir de características específicas, como el color de la ropa. La posibilidad de realizar consultas basadas en estos datos reduce considerablemente la necesidad de revisar manualmente ingentes cantidades de vídeo.

**Plataformas IoT.** Los metadatos pueden recopilarse y presentarse visualmente en plataformas de inteligencia empresarial para generar información útil mediante el análisis de tendencias históricas y en tiempo real. Los análisis estadísticos basados en la afluencia o la experiencia del cliente ayudan a tomar decisiones basadas en datos con el objetivo de mejorar las operaciones.

**Segundo nivel de analítica.** Algunas aplicaciones requieren una combinación de procesamiento en el extremo y en servidores para ejecutar análisis más avanzados. En estos casos, puede aplicarse un preprocesamiento en la cámara y un procesamiento posterior en un servidor. Un sistema híbrido de este

tipo abre la puerta a escalar la analítica sin que se disparen los costes, ya que solo se transmiten los vídeos y metadatos necesarios al servidor.

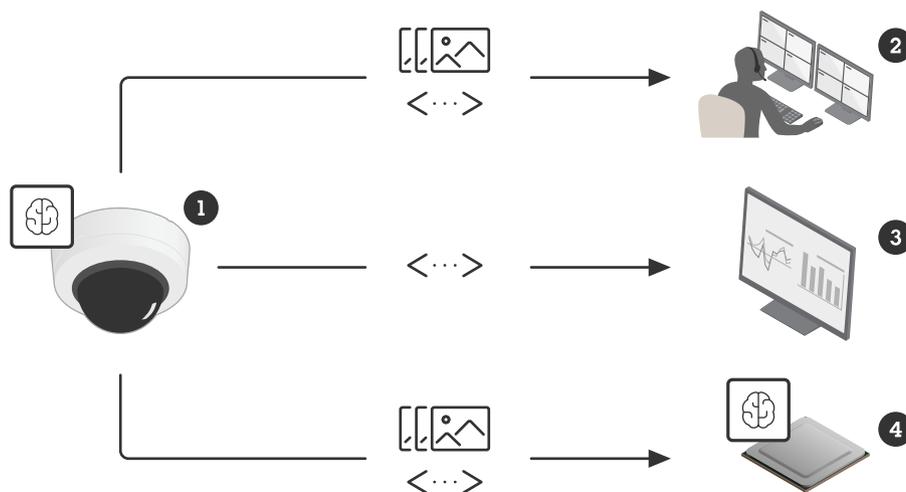


Figure 2. Dónde se utilizan los metadatos

- 1 Aplicaciones en el extremo
- 2 VMS
- 3 Plataformas IoT
- 4 Segundo nivel de analítica

## 6 ¿Cómo se transmiten los metadatos?

Los metadatos generados se pueden transmitir de diferentes maneras en función de su finalidad. En aplicaciones en tiempo real, la transmisión debe ser continua, porque son clave para ofrecer una respuesta adecuada y saber qué está ocurriendo en una escena. En aplicaciones donde no es necesario un seguimiento en tiempo real, es posible agrupar los metadatos, por ejemplo, teniendo en cuenta el

seguimiento de un objeto concreto en la escena, antes de transmitirlos al cliente. De este modo, hay que almacenar y procesar menos datos.

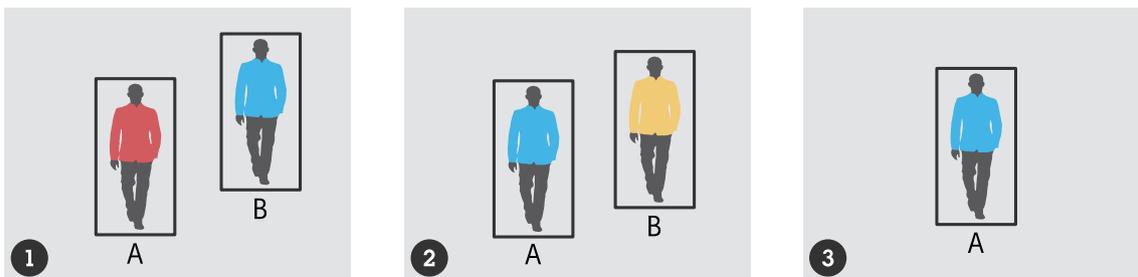


Figure 3. Esta figura ilustra la transmisión en directo de metadatos. Los fotogramas continuos de la cámara aportan información en tiempo real sobre la escena. Cada fotograma captura la escena en un momento concreto, desvinculado de los eventos anteriores.

- 1 El fotograma 1 detecta los objetos A y B. Clasifica el objeto A como una persona con ropa roja y el objeto B, como una persona con ropa azul.
- 2 En el fotograma 2, la cámara actualiza la clasificación: determina que el objeto A lleva ropa azul y que el objeto B lleva ropa amarilla. Aunque los objetos son los mismos que en el fotograma 1, sus atributos de color cambian y eso se refleja en los metadatos.
- 3 El fotograma 3 muestra la ausencia del objeto B: la cámara sigue solo el objeto A y lo continúa clasificando como una persona con ropa azul.

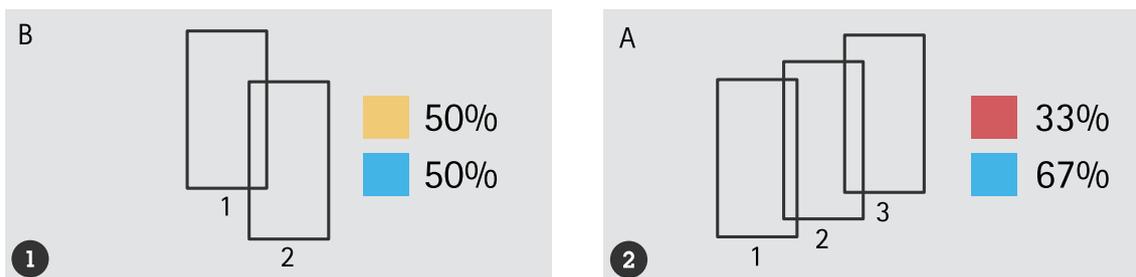


Figure 4. Esta figura ejemplifica la transmisión agrupada de metadatos: la cámara proporciona la información en un formato unificado a partir del seguimiento detectado de los objetos en la escena. Los fotogramas de cada objeto incluyen todos los detalles obtenidos a lo largo de todo el seguimiento del objeto.

- 1 En el primer fotograma, se muestran detalles sobre el objeto B, como la primera y la última detección, un resumen de la trayectoria y atributos detectados durante el seguimiento. Hay una probabilidad del 50% de que el objeto B lleve ropa amarilla y del 50% de que lleve ropa azul.
- 2 El segundo fotograma utiliza el mismo formato para el objeto A, con una probabilidad del 33% de que lleve ropa roja y del 67% de que lleve ropa azul.

La ventaja del método agrupado es que la cámara envía menos información al usuario, ya que solo se transmiten metadatos cuando hay objetos presentes en la escena y, además, se presentan agrupados para

facilitar la interpretación. El método de transmisión en tiempo real permite obtener una descripción completa de la escena en cada fotograma, aunque no haya actividad ni objetos, y el usuario tiene que interpretar esta información según sus necesidades concretas. Como apuntábamos antes, el método de transmisión en tiempo real es útil cuando se utiliza para tomar decisiones en el momento, mientras que el método de agrupación de metadatos es recomendable cuando no hay que actuar de forma inmediata.

Entender las ventajas y las limitaciones de cada opción es fundamental al diseñar la arquitectura del sistema. Por ejemplo, en una plataforma IoT que genere información a partir de los metadatos resulta útil recibir un resumen a posteriori de los objetos de la escena, porque estos servicios suelen tener ciertas limitaciones de almacenamiento y ancho de banda.

Además, los metadatos pueden transmitirse por diferentes protocolos de comunicación y formatos de archivo en función de las necesidades y preferencias específicas del destinatario.

## **7 Metadatos de audio**

Algunas cámaras pueden decodificar audio para recuperar metadatos de audio. La analítica de reconocimiento de audio puede detectar patrones acústicos y destacar sonidos de interés en audio en directo y grabado. Así, la combinación de sistemas de reconocimiento de audio con dispositivos de videovigilancia puede alertar a los operadores sobre posibles incidentes y mostrarles las vistas de las cámaras relacionadas. El sistema, por ejemplo, podría identificar situaciones de violencia verbal para que la situación no vaya a peor, detectar el ruido de cristales rotos para evitar intrusiones o alertar de crisis en pacientes. Como los operadores no solo ven lo que ocurre en la escena sino que también lo escuchan, los sistemas de reconocimiento de sonido facilitan la detección y la intervención rápida y, en muchos casos, evitan el empeoramiento de los conflictos. El reconocimiento del sonido aporta, además, un segundo filtro de verificación.

Una solución de analítica entrenada para reconocer patrones de sonido normalmente trata de identificar una combinación de características, que van desde el nivel de decibelios hasta la energía en las diferentes frecuencias a lo largo del tiempo. Los patrones de sonido específicos pueden detectarse y etiquetarse de una forma similar a como se detectan y etiquetan las clases de objetos en el vídeo.

## **8 Combinación de metadatos de varias fuentes**

Los metadatos pueden desplegar todo su potencial cuando se aplican a diferentes fuentes, por ejemplo de imagen, audio, vinculadas a la actividad o relacionadas con procesos. Fuentes como el seguimiento por RFID, las coordenadas GPS, las alertas de manipulación, las lecturas de contadores (como la temperatura o los niveles de concentración), la detección de ruido o los datos transaccionales de puntos de venta pueden resultar de gran utilidad para la gestión de cualquier local. Los datos de todas las fuentes pueden ordenarse juntos según sus marcas de tiempo.

La combinación de metadatos de diferentes fuentes permite obtener una información mucho más completa que al usar fuentes individuales por separado. Para una integración sencilla de los metadatos, es imprescindible usar protocolos abiertos y normas sectoriales.

# Acerca de Axis Communications

Axis contribuye a crear un mundo más inteligente y seguro a través de soluciones para mejorar la seguridad y el rendimiento empresarial. Como empresa de tecnología de red y líder del sector, Axis ofrece soluciones de videovigilancia, control de acceso y sistemas de audio e intercomunicación. Se ven reforzadas por aplicaciones de análisis inteligentes y respaldadas por formación de alta calidad.

Axis tiene alrededor de 4000 empleados dedicados en más de 50 países y colabora con socios de integración de sistemas y tecnología en todo el mundo para ofrecer soluciones personalizadas. Axis se fundó en 1984 y la sede está en Lund, Suecia