

WHITE PAPER

Analitiche audio per la sicurezza

Febbraio 2024

Sommario

Le analitiche audio per la sicurezza possono rilevare modelli sonori ed evidenziare suoni inattesi in tempo reale. Ad esempio, possono identificare un'aggressione verbale per evitare che la situazione degeneri, rilevare vetri rotti per prevenire effrazioni o trasmettere avvisi se una persona è in difficoltà.

L'uso di sensori diversi, ad esempio video e audio (una telecamera e un microfono), aumenta l'attendibilità del rilevamento e contribuisce a ottenere informazioni più utili.

AXIS Audio Analytics è una piattaforma software per analitiche edge-based integrata nel software dei dispositivi Axis. Le analitiche captano e rilevano i suoni senza salvare il flusso audio originale. Questo accorgimento tutela la privacy e funziona perché le analitiche sono edge-based e forniscono solo metadati audio.

Sommario

1	Introduzione	4
2	Tecnologia	4
	2.1 Rilevamento di eventi sonori	4
	2.2 Acquisizione ed elaborazione dell'audio	4
	2.3 Analitiche edge-based o server-based	5
	2.4 Metadati	6
3	AXIS Audio Analytics	7
	3.1 Privacy	7
	3.2 Casi d'uso	8
4	Restrizioni legali	12
5	Dichiarazione di non responsabilità	12

1 Introduzione

Le analitiche audio per la sicurezza possono rilevare modelli sonori ed evidenziare suoni inattesi in tempo reale. Ad esempio, le analitiche possono identificare suoni associati ad aggressioni e vetri rotti o segnalare tempestivamente persone in difficoltà.

Le analitiche audio combinate con la videosorveglianza possono segnalare agli operatori potenziali eventi e guidarli verso le rispettive inquadrature delle telecamere. In questo modo, è possibile rilevare gli eventi in modo tempestivo, intervenire rapidamente e, in molti casi, impedire che la situazione degeneri.

Questo documento tecnico spiega come utilizzare le analitiche audio per la sicurezza. Descrive le tecnologie di acquisizione ed elaborazione dell'audio, con particolare attenzione alle analitiche edge in tempo reale: una di queste è la classificazione audio basata su IA direttamente sulla telecamera o sul microfono. Inoltre, spiega che le analitiche audio in modalità edge offrono varie opzioni per tutelare la privacy attraverso i metadati audio.

Il documento non fornisce consulenza legale. Prima di installare qualsiasi sistema di sorveglianza, è necessario informarsi sulle leggi e sui regolamenti in vigore nella propria area e per gli usi previsti. Spetta al proprietario del sistema garantire la conformità alle leggi, ai regolamenti e alle raccomandazioni territoriali.

2 Tecnologia

2.1 Rilevamento di eventi sonori

Un evento sonoro è un segmento audio che una persona identificherebbe come distintivo, ad esempio urla o vetri rotti. Questi suoni distintivi possono essere rilevati ed etichettati in modo simile a quanto avviene per le classi di oggetti nelle analitiche video.

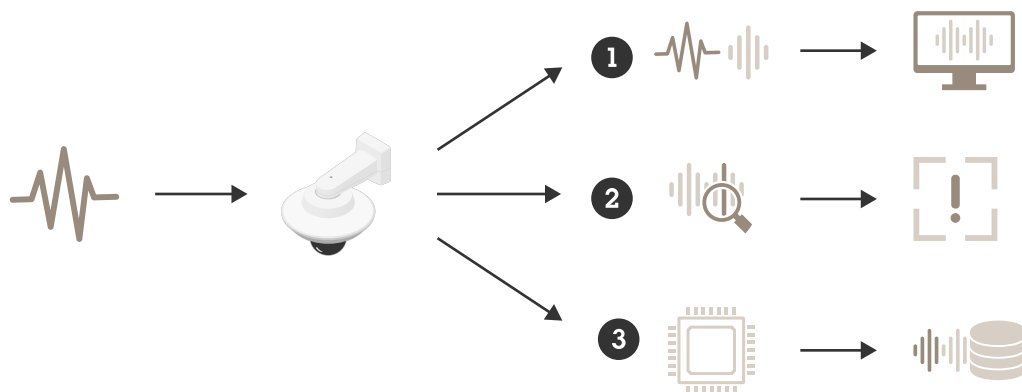
In genere, le analitiche addestrate con modelli acustici valutano una combinazione di caratteristiche come il livello di decibel e l'energia delle varie frequenze nel tempo. Quando rileva uno specifico modello sonoro, il sistema può inviare una notifica automatica al personale tramite un avviso visivo o attivando un allarme.

Se si utilizzano algoritmi basati su IA, possono essere addestrati con grandi quantità di dati. Ad esempio, un algoritmo può rilevare in modo affidabile le urla di una persona dopo l'addestramento con migliaia di suoni simili.

2.2 Acquisizione ed elaborazione dell'audio

Le analitiche audio utilizzano i dati acquisiti e ne analizzano le caratteristiche acustiche per generare un output non audio. Acquisire l'audio significa sostanzialmente digitalizzarlo e renderlo utilizzabile nel software. Queste operazioni vengono svolte captando le vibrazioni sonore nell'aria con un microfono, convertendo i segnali analogici in segnali digitali e trasferendoli a un processore. Se l'audio acquisito non viene trasferito su un supporto permanente, come una memoria flash o un disco rigido, non viene registrato. Sulle telecamere Axis, lo streaming audio e la registrazione sono disattivati per impostazione predefinita.

Dopo l'acquisizione iniziale, i dati vengono preparati per le fasi di elaborazione successive. I vari preparativi possono essere eseguiti in modo parallelo o esclusivo.



Preparazione dell'audio

- 1 *Trasformazione*
- 2 *Analitiche edge in tempo reale*
- 3 *Elaborazione e codifica per lo streaming o l'archiviazione: se si utilizza una telecamera Axis, l'audio non viene trasmesso né archiviato, a meno che non si attivi volontariamente lo streaming audio.*

- **Trasformazione.** L'audio viene reso astratto e convertito, ad esempio in informazioni visive come un grafico che mostra lo spettro sonoro. Questo processo non può essere invertito: non è possibile recuperare l'audio originale dal grafico dello spettro.
- **Analitiche edge in tempo reale.**

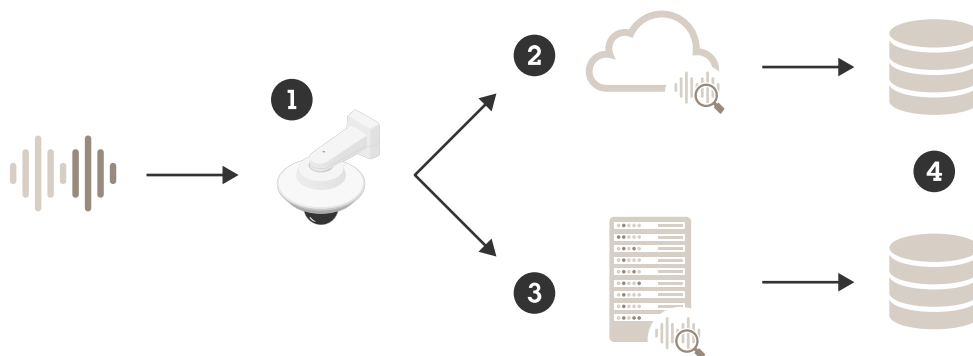
Se l'audio viene processato in modalità edge, è possibile utilizzare un **classificatore audio**. Il risultato è una serie di metadati che descrive le caratteristiche sonore. L'audio originale non può essere ricreato a partire dai metadati.

È possibile utilizzare un **rilevatore audio** per riconoscere modelli, livelli o frequenze e fornire informazioni di stato. Anche in questo caso, l'audio originale non è ripristinabile.
- **Elaborazione e codifica.** Nei casi in cui si utilizza l'audio originale (non trasformato o analizzato), in genere si eseguono alcune operazioni di elaborazione e codifica, in modo da preparare i dati audio per i casi d'uso previsti. Questi casi d'uso possono prevedere l'archiviazione dei dati audio in modalità edge, lo streaming a client esterni per un'ulteriore elaborazione (su server o cloud) o l'archiviazione esterna. Con una telecamera Axis è necessario innanzitutto attivare lo streaming audio, che è disattivato per impostazione predefinita per motivi di privacy (controllo della privacy audio).

2.3 Analitiche edge-based o server-based

La posizione del motore analitico all'interno del sistema è importante per molti motivi. Soprattutto per rispettare la privacy e i regolamenti sui dati personali, è importante sapere dove vengono analizzati i dati audio dall'algoritmo software. Esistono casi in cui i dati audio non possono essere inviati in rete; dunque, è fondamentale che i dati audio acquisiti (ma non archiviati) possano essere analizzati localmente. Se

gli algoritmi sono molto complessi e non possono essere eseguiti in modalità edge, può essere necessario inviare i dati audio digitali in cloud o a un server.



Ambienti di esecuzione delle analitiche

- 1 Edge
- 2 Cloud
- 3 Server
- 4 Archiviazione: è possibile archiviare l'audio originale solo se lo streaming audio è attivato.

- **Analitiche edge.** Quando le analitiche vengono eseguite in modalità edge, nessun flusso audio deve lasciare la telecamera. La telecamera invia solo l'output delle analitiche, ovvero metadati o trigger. Le analitiche di AXIS Audio Analytics sono edge-based.
- **Analitiche su server.** Quando le analitiche vengono eseguite su un server, i dati audio devono essere trasmessi dalla telecamera. Se pre-elaborati sulla telecamera, questi possono limitarsi a metadati astratti o spersonalizzati. Normalmente, un server fa parte di un sistema chiuso (controllato dal proprietario del sistema), quindi è possibile trattare l'audio trasmesso se può causare problemi di privacy. Tuttavia, occorre sempre seguire le regole e le normative applicabili.
- **Analitiche su cloud.** I dati audio possono anche essere trasmessi a un server che opera in un contesto cloud. Come nel caso delle analitiche su server, i dati audio possono essere pre-elaborati sotto forma di metadati. L'uso del cloud è spesso decentralizzato, quindi è ancora più importante curare la privacy e garantire il rispetto delle regole.

2.4 Metadati

Le analitiche audio generano un flusso costante di metadati sui livelli audio. Le analitiche generano anche metadati sugli eventi rilevati dai classificatori e dal rilevamento audio adattivo. Le analitiche eseguite in modalità edge analizzano i dati audio all'interno della telecamera. Non occorre trasmettere il flusso audio vero e proprio: è possibile trasmettere solo i metadati, che danno informazioni su ciò che accade nella scena.

Il flusso di metadati consente l'ispezione visiva dell'involucro audio, rappresentandolo ad esempio su una dashboard per allineare e visualizzare insieme gli eventi audio e gli eventi video. In questo modo, i metadati consentono di cercare in modo rapido ed efficiente eventi specifici e suoni inattesi, facendo risparmiare molte ore agli investigatori quando eseguono ricerche in grandi quantità di filmati.

3 AXIS Audio Analytics

AXIS Audio Analytics è una piattaforma software che introduce costantemente analitiche e funzionalità nuove o aggiornate. Le analitiche sono integrate nel software AXIS OS e sono incluse gratuitamente su telecamere e altri dispositivi Axis.

Le analitiche all'interno della piattaforma sono edge-based, con algoritmi eseguiti direttamente sulla telecamera. In questo modo si ha una scalabilità ottimale, un basso traffico dati e si tutela la privacy. Viene salvato solo l'output delle analitiche (metadati o trigger). L'audio non viene registrato o trasmesso dalla telecamera e i suoni originali non possono essere ricreati a partire dai metadati.

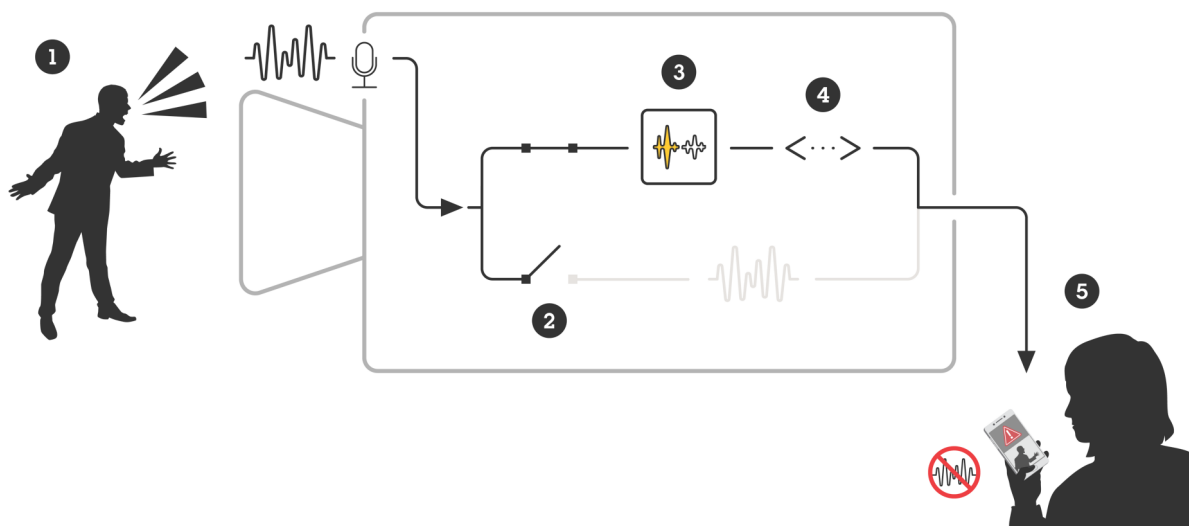
- **Classificazione audio (urla e grida).** Si tratta di un classificatore basato su IA che rileva e analizza suoni specifici come urla e grida. In futuro verranno aggiunte altre classificazioni audio, ad esempio di vetri rotti. Il prodotto delle analitiche di classificazione sono metadati che descrivono le caratteristiche sonore.
- **Rilevamento audio adattivo.** Si tratta di un rilevatore che crea un evento quando si ha una variazione improvvisa del livello audio. Rileva picchi sonori di qualsiasi tipo, con il vantaggio di adattarsi al rumore ambientale anche al variare del livello di rumorosità.

3.1 Privacy

In generale, le analitiche audio non registrano l'audio in ingresso e non lo trasmettono dalla telecamera. Semplicemente, elaborano i suoni per consentire la ricerca di eventi, modelli o livelli sonori specifici in un sistema ricevente, come una dashboard per ulteriori indagini o un software di gestione video per avvisare gli operatori. Nessun dato audio può essere ricostruito e nessuna conversazione privata può essere registrata, perché le analitiche sono edge-based e forniscono metadati audio.

AXIS Audio Analytics non registra e non trasmette l'audio, ma trasmette solo metadati. Per motivi di privacy, lo streaming audio è disattivato per impostazione predefinita sui dispositivi Axis (controllo della privacy audio). Dunque, l'audio non viene trasmesso, non viene registrato e non può essere ricreato. In caso di necessità è possibile attivare lo streaming audio; anche se è disattivato, però, si riceve una notifica quando accade qualcosa di rilevante.

La figura mostra il funzionamento di AXIS Audio Analytics con controllo della privacy audio quando rileva i suoni e utilizza i metadati per creare un avviso.



Controllo della privacy audio e AXIS Audio Analytics sulle telecamere Axis, da un urlo (1) a un avviso (5).

- 1 I suoni vengono captati dal microfono.*
- 2 Per impostazione predefinita, lo streaming audio è disattivato.*
- 3 Il classificatore audio di AXIS Audio Analytics rileva un urlo nell'audio in ingresso.*
- 4 Il classificatore audio genera metadati che comprendono una notifica di evento.*
- 5 Gli interessati ricevono un avviso in base alla notifica dell'evento e ai metadati e possono verificare il flusso video. Non è disponibile alcun flusso audio.*

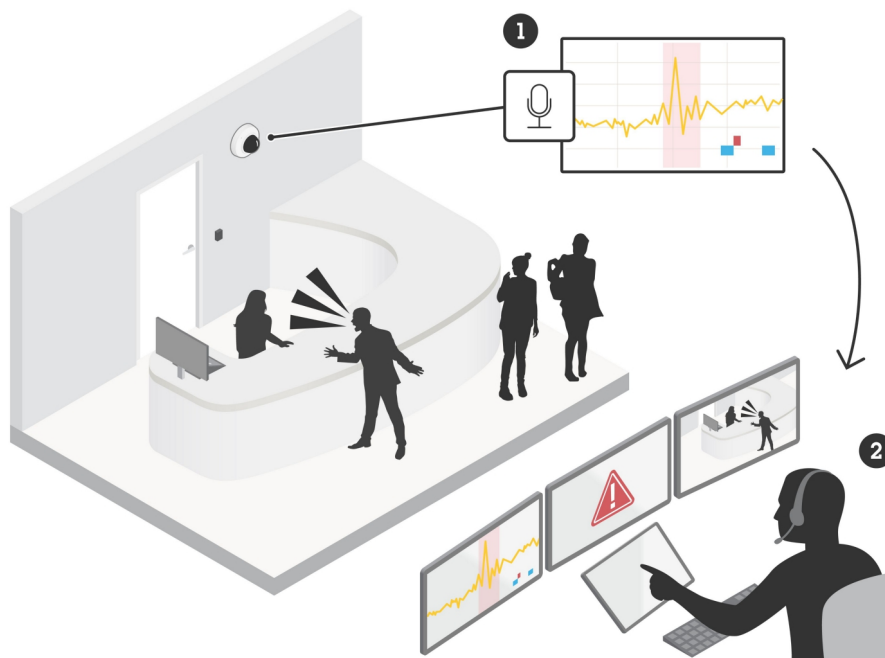
3.2 Casi d'uso

Anche se le analitiche basate su IA hanno buone capacità di filtrare i rumori irrilevanti, possono segnalare falsi allarmi se il rumore di fondo è molto alto. La pioggia che batte sulle finestre, i tuoni, le sirene, la musica o le scene con molte persone che parlano possono attivare falsi allarmi. Pertanto, gli ambienti d'uso più tipici sono aree tranquille come banche e reception, ma anche spazi interni negli orari di chiusura come negozi, ristoranti, scale o uffici.

3.2.1 Rilevamento audio con avvisi

In banca o in reception, le analitiche di classificazione audio possono monitorare l'area e rilevare suoni associati ad aggressioni o vetri rotti. Al rilevamento, i dati degli eventi del sistema e i metadati audio

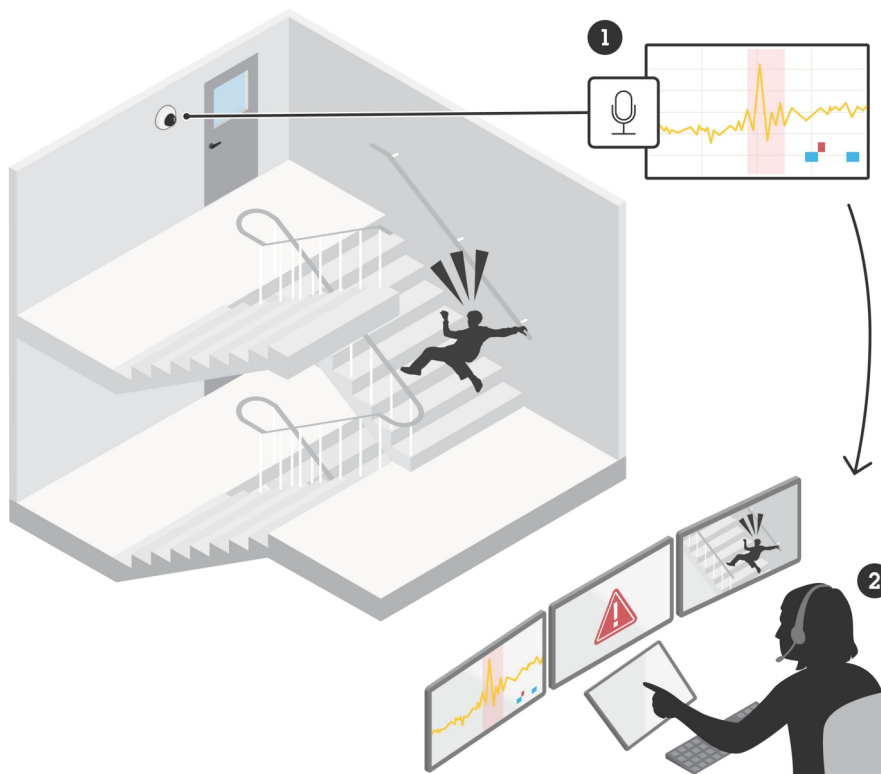
inviando notifiche automatiche al personale tramite un avviso visivo o attivando un allarme. In questo modo, si riceve un avviso tempestivo che consente di reagire e intervenire rapidamente.



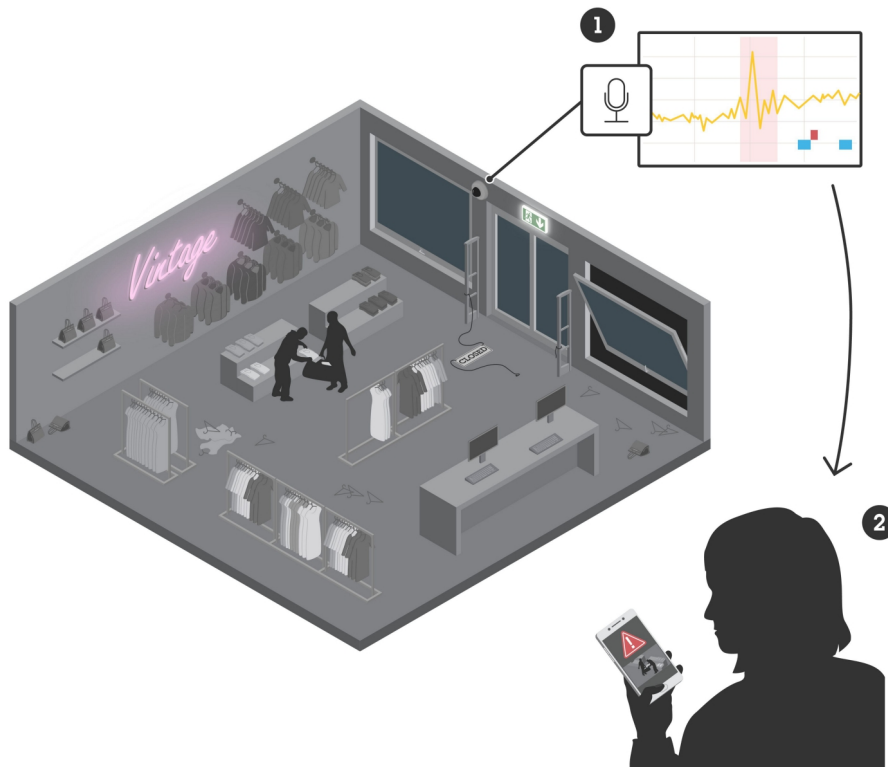
- 1 Una telecamera con analitiche di classificazione audio rileva urla o grida alla reception.
- 2 Un operatore riceve un avviso e può controllare il video prima di intraprendere ulteriori azioni.

Le analitiche con rilevamento audio adattivo possono essere utilizzate per rilevare suoni inattesi al di fuori degli orari di lavoro. Le analitiche analizzano i suoni ambientali e rispondono quando rilevano voci,

finestre rotte o altri rumori improvvisi e di breve durata. Quando rilevano eventi, le analitiche inoltrano i metadati per avvisare gli operatori.



- 1 Una telecamera con analitiche di classificazione audio rileva suoni inattesi dopo l'orario di lavoro.
- 2 Un operatore riceve un avviso e può controllare il video prima di intraprendere ulteriori azioni.



- 1 Una telecamera con analitiche di classificazione audio rileva i suoni in un negozio dopo l'orario di lavoro.
- 2 Il titolare del negozio riceve un avviso e può controllare il video prima di intraprendere ulteriori azioni.

3.2.2 Combinazione di sensori per sfruttare al massimo il sistema di sorveglianza

Spesso, i sistemi di sorveglianza integrano vari tipi di sensori. Uno di questi, ovviamente, è il sensore d'immagine della telecamera, che registra l'aspetto visivo di una scena. Inoltre si utilizzano spesso sensori non ottici, come quelli di movimento basati sulla tecnologia radar o sull'emissione di radiazioni a infrarossi. I sensori non ottici integrano il sistema di telecamere aggiungendo informazioni di altro tipo.

Aggiungendo sensori audio (microfoni) a un sistema di sorveglianza, si dispone di un'arma in più in gran parte dei casi d'uso. L'aggiunta di funzionalità e analitiche audio a un sistema non audio consente l'interazione multisensore. Se si utilizzano analitiche video, le analitiche audio possono aumentare l'attendibilità del rilevamento. Questo è vero soprattutto se le analitiche video devono affrontare condizioni difficili, come una scarsa illuminazione.

È possibile configurare il sistema nel software di gestione video in modo che attivi azioni solo se reagiscono sia le analitiche video che le analitiche audio, ad esempio quando le prime rilevano una persona nel campo visivo della telecamera e le seconde rilevano un urlo. In alcuni ambienti, questa combinazione offre il giusto livello di sicurezza.

3.2.3 Input per le dashboard

I metadati audio possono essere inviati a dashboard analitiche, ad esempio VSaaS (Video Surveillance as a Service), IoT (Internet of Things) o piattaforme di business intelligence, che raccolgono e presentano visivamente i metadati. Le dashboard analizzano le tendenze storiche e in tempo reale per generare una panoramica istantanea e informazioni utili. Le analisi statistiche basate sul flusso o sull'esperienza dei clienti consentono di prendere decisioni basandosi sui dati, in modo da migliorare le operazioni.

Con le dashboard è possibile vedere i risultati senza neanche ascoltare l'audio o recuperare l'audio originale. È invece possibile ottenere informazioni utili, ad esempio dal conteggio degli eventi, senza preoccuparsi della privacy dei dati. Nota: le leggi potrebbero prevedere restrizioni diverse a seconda che l'audio sia registrato o solo acquisito.

4 Restrizioni legali

Molte persone si preoccupano dell'uso dei microfoni nella videosorveglianza. In genere si tratta di preoccupazioni legate alla registrazione di normali conversazioni insieme al materiale video. In genere, però, le analitiche non registrano e non trasmettono l'audio. Le leggi che regolamentano la sorveglianza variano per regione e Paese, quindi occorre sempre sapere quali operazioni sono consentite prima di usare l'audio nel sistema di sorveglianza.

L'acquisizione e la registrazione dell'audio possono essere vietati o richiedere accorgimenti particolari per diversi motivi, a causa delle leggi nazionali o di varie regole e norme locali. Mentre in un'area o un ambiente l'acquisizione dell'audio potrebbe essere consentita, la registrazione potrebbe essere vietata. Anche le aziende possono vietare la sorveglianza audio nei loro locali.

5 Dichiarazione di non responsabilità

Il presente documento e i suoi contenuti sono forniti per gentile concessione di Axis. Tutti i diritti relativi al documento o i diritti di proprietà intellettuale correlati (compresi a titolo esemplificativo marchi commerciali, denominazioni commerciali, logotipi e marchi simili) sono protetti per legge. Tutti i diritti, i titoli e/o gli interessi in relazione al documento o i diritti di proprietà intellettuale sono e rimangono assegnati ad Axis Communications AB.

Il documento viene fornito nello stato in cui si trova, senza garanzie di alcun tipo e a solo scopo informativo. Le informazioni fornite nel presente documento non costituiscono e non intendono costituire consulenza legale. Il documento non ha lo scopo di creare né crea obblighi legali per Axis Communications AB e/o le sue affiliate. Gli obblighi di Axis Communications AB e/o delle sue affiliate in relazione ai prodotti Axis sono soggetti esclusivamente ai termini e alle condizioni contrattuali tra Axis e l'entità che ha acquistato i suddetti prodotti direttamente da Axis.

PER EVITARE OGNI DUBBIO, TUTTI I RISCHI INERENTI L'USO, I RISULTATI E L'ESECUZIONE DEL PRESENTE DOCUMENTO SONO ASSUNTI DALL'UTILIZZATORE DEL DOCUMENTO. AXIS DECLINA ED ESCLUDE, NELLA MASSIMA MISURA CONSENTITA PER LEGGE, TUTTE LE GARANZIE STATUTARIE, ESPLICITE O IMPLICITE, COMPRESA A TITOLO ESEMPLIFICATIVO GARANZIE IMPLICITE DI COMMERCIALITÀ, IDONEITÀ PER UNO SCOPO PARTICOLARE, TITOLARITÀ E NON VIOLAZIONE E RESPONSABILITÀ SUI PRODOTTI, O GARANZIE DERIVANTI DA QUALSIASI PROPOSTA, SPECIFICA O ESEMPIO IN RELAZIONE AL PRESENTE DOCUMENTO.

Informazioni su Axis Communications

Axis consente un mondo più intelligente e più sicuro creando soluzioni per migliorare la sicurezza e le prestazioni aziendali. Come società di tecnologie di rete e leader nel settore, Axis offre soluzioni nella videosorveglianza, controllo degli accessi, interfono e sistemi audio. Queste sono ottimizzate da applicazioni di analisi intelligente e supportate da formazione di alta qualità.

Axis ha circa 4.000 impiegati dedicati in più di 50 paesi e collabora con partner di tecnologia e integrazione di sistema in tutto il mondo per offrire soluzioni di clienti. Fondata nel 1984, Axis è con sede a Lund, in Svezia