

LIVRE BLANC

# Analyse audio à des fins de sécurité et de protection

Février 2024

## **Avant-propos**

L'analyse audio à des fins de sécurité et de protection peut détecter les motifs sonores et repérer les bruits inhabituels dans l'audio en direct. Les fonctions d'analyse audio peuvent par exemple identifier les agressions verbales pour éviter toute escalade vers des violences physiques, détecter les bris de verre pour prévenir les effractions ou fournir des alertes précoces de personnes en détresse.

Le recours à plusieurs types de capteurs, notamment audio et vidéo (caméra et microphone), accroît la fiabilité des résultats détectés et contribue à apporter des éclairages plus exploitables.

AXIS Audio Analytics est une plateforme logicielle d'analyse en périphérie de réseau, intégrée au logiciel des dispositifs Axis. Ces fonctions d'analyse capturent et détectent les bruits sans enregistrer le flux audio d'origine. Cette méthode préserve la confidentialité du fait que les analyses s'exécutent en périphérie de réseau pour fournir des métadonnées audio.

# Table des matières

1	Introduction	4
2	Technologie	4
	2.1 Détection d'événements sonores	4
	2.2 Capture et traitement du son	4
	2.3 Analyses en local ou sur serveur	5
	2.4 Métadonnées	6
3	AXIS Audio Analytics	7
	3.1 Confidentialité	7
	3.2 Scénarios d'utilisation	8
4	Restrictions juridiques	12
5	Clause de non-responsabilité	12

# 1 Introduction

L'analyse audio à des fins de sécurité et de protection peut détecter les motifs sonores et repérer les bruits inhabituels dans l'audio en direct. Par exemple, des fonctions d'analyse peuvent identifier les bruits associés à une agression, détecter les bris de verre ou fournir des signes précoces de personnes en détresse.

Associées à la vidéosurveillance, les fonctions d'analyse audio peuvent alerter les opérateurs d'incidents potentiels en cours et les orienter vers les vues des caméras correspondantes. Ces systèmes favorisent une détection précoce et une intervention rapide, évitant les escalades dans de nombreux cas.

Ce livre blanc explique comment utiliser les fonctions d'analyse audio à des fins de sécurité et de protection. Nous présentons les technologies de capture et de traitement du son, notamment sous l'angle de l'analyse locale en temps réel, au travers d'une classification des bruits par IA directement dans la caméra ou le microphone. Nous présentons également les différentes options pour préserver la confidentialité, par la transmission des métadonnées audio issues de l'analyse des bruits en périphérie de réseau.

Ce document n'aborde pas les questions d'ordre juridique. Avant d'installer un système de surveillance, vous devez examiner les lois et règlements qui s'appliquent à votre scénario d'utilisation dans votre région. Il incombe au responsable du système de vérifier qu'il respecte les législations, réglementations et recommandations régionales.

## 2 Technologie

### 2.1 Détection d'événements sonores

Un événement sonore est un segment audio généralement identifiable en tant que concept distinctif, par exemple cris ou bris de verre. Il est possible de détecter ces types de concepts sonores et de les étiqueter de la même manière que des classes d'objet en analyse vidéo.

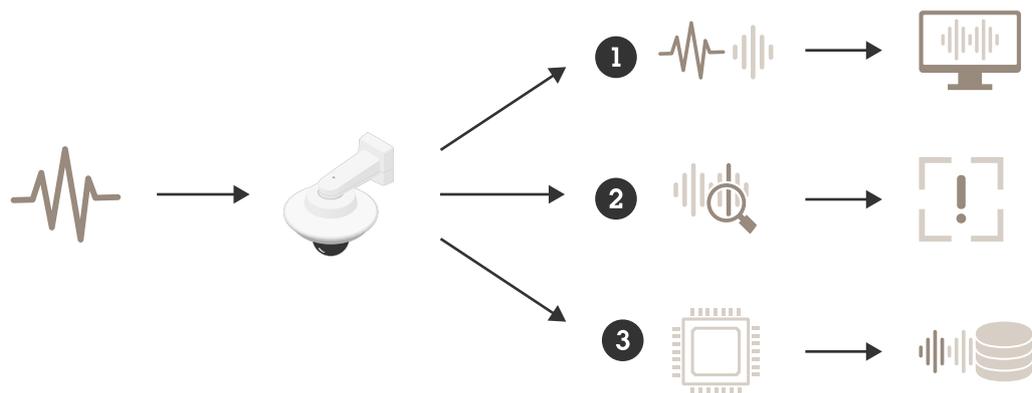
Les fonctions d'analyse entraînées pour reconnaître certains motifs sonores écoutent généralement une série de caractéristiques, du niveau sonore à l'évolution de l'énergie des différentes fréquences en fonction du temps. En cas de détection d'un motif sonore particulier, le système peut envoyer une notification automatique au personnel, sous la forme d'une alerte visuelle ou du déclenchement d'une alarme.

S'ils font appel à l'IA, les algorithmes peuvent être entraînés avec de gros volumes de données. Par exemple, un algorithme peut détecter des cris humains avec fiabilité après avoir « appris » des milliers de ces bruits.

### 2.2 Capture et traitement du son

Les fonctions d'analyse audio analysent les caractéristiques acoustiques pertinentes des données audio capturées pour produire des résultats non sonores. Capturer l'audio signifie simplement le numériser et le rendre disponible dans un logiciel pour utilisation. Cette action est réalisée par la détection des vibrations sonores dans l'air à l'aide d'un microphone, puis la conversion de ces signaux analogiques en signaux numériques et leur transmission à un processeur. Si l'audio capturé n'est pas placé sur un support permanent tel qu'une clé USB ou un disque dur, il n'est pas enregistré. Dans les caméras Axis, l'enregistrement et la diffusion du flux audio sont désactivés par défaut.

Après la capture initiale de l'audio, les informations capturées sont préparées pour les étapes suivantes de traitement. Les différentes préparations peuvent être réalisées en parallèle ou en exclusivité.



### Préparation de l'audio

- 1 Transformation
- 2 Analyses en périphérie de réseau en temps réel
- 3 Traitement et encodage pour streaming ou stockage : si vous utilisez une caméra Axis, le son n'est ni diffusé, ni stocké si vous n'activez pas volontairement le streaming audio.

- **Transformation.** Le son est rendu abstrait et converti, par exemple, en informations visuelles telles qu'un graphique montrant le spectre sonore. Ce procédé ne peut être annulé : vous ne pouvez pas récupérer le son d'origine à partir du graphique du spectre.

- **Analyses en périphérie de réseau en temps réel.**

Un **classificateur de son** peut être utilisé si le son est traité en périphérie de réseau. Il en résultera des métadonnées qui décrivent les caractéristiques du son. Le son d'origine ne peut être recréé à partir de ses métadonnées.

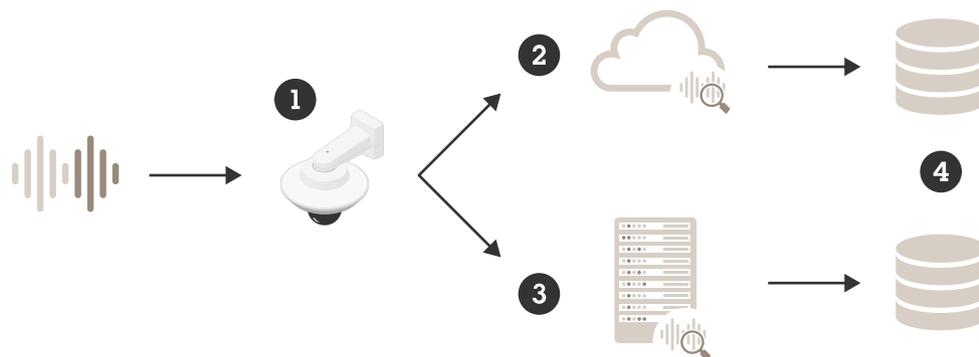
Un **détecteur de bruit** peut être utilisé pour reconnaître des motifs acoustiques, des volumes sonores ou des fréquences et fournir un état de la situation. À nouveau, le son d'origine ne peut être restauré.

- **Traitement et cryptage.** Pour les cas dans lesquels l'audio d'origine sera utilisé (non transformé ou analysé), un traitement et un cryptage sont normalement réalisés pour préparer les données audio en vue des cas d'usage prévus. Ces cas d'usage peuvent impliquer le stockage de données audio en périphérie de réseau, la diffusion à des clients extérieurs pour traitement complémentaire (sur serveur ou cloud) ou le stockage externe. Avec une caméra Axis, vous devez d'abord activer manuellement le streaming audio, qui est désactivé par défaut pour des raisons de confidentialité (contrôle de confidentialité audio).

## 2.3 Analyses en local ou sur serveur

L'emplacement du moteur d'analyse dans le système est important pour de nombreuses raisons. En particulier pour la gestion des questions de confidentialité et la conformité avec les réglementations relatives aux données à caractère personnel, il est important de savoir où le logiciel d'algorithmes analyse les données audio. Il existe des situations dans lesquelles les données audio ne peuvent être envoyées sur le réseau et il est essentiel que les données audio capturées (mais non stockées) puissent être analysées localement. Si des algorithmes à très forts besoins en calcul sont nécessaires, de telle sorte qu'ils ne

peuvent être exécutés en périphérie de réseau, il pourrait être nécessaire d'envoyer les données audio numériques sur le cloud ou un serveur.



#### Emplacements possibles d'exécution des fonctions d'analyse

- 1 Périphérie de réseau
- 2 Cloud
- 3 Serveur
- 4 Espace de stockage : il est possible de stocker le son d'origine uniquement si le flux audio est activé.

- **Analyses en périphérie de réseau.** Lorsque les analyses s'exécutent en périphérie de réseau, il n'est pas nécessaire que la caméra transmette le flux audio. Seul le résultat des analyses pratiquées, c'est-à-dire les métadonnées ou les déclencheurs, est transmis. Les fonctions d'analyse d'AXIS Audio Analytics sont exécutées localement.
- **Analyses sur serveur.** Lorsque les fonctions d'analyse s'exécutent sur serveur, la caméra doit transmettre les données audio. Si les données audio sont prétraitées dans la caméra, elles peuvent se limiter à des métadonnées abstraites ou non identifiables. Comme un serveur fait normalement partie d'un système fermé (sous le contrôle d'un responsable du système), il est possible de résoudre les questions d'anonymat des données audio transmises. Néanmoins, il faut s'assurer que les lois et réglementations en vigueur sont respectées.
- **Analyses sur le cloud.** Il est également possible de transmettre les données audio à un serveur dans le cloud. Comme dans le cas des analyses sur serveur, les informations audio peuvent être prétraitées en métadonnées. L'utilisation du cloud est souvent décentralisée, il est par conséquent d'autant plus important de s'occuper des problèmes de confidentialité et de veiller à la conformité avec la réglementation.

## 2.4 Métadonnées

Les fonctions d'analyse audio produisent un flux continu de métadonnées relatives au niveau sonore. Elles génèrent également des métadonnées d'après des événements détectés par des fonctions de classification et de détection audio adaptative. Les fonctions d'analyse exécutées en périphérie de réseau analysent les informations audio dans la caméra. Elles n'ont donc pas à transmettre le flux audio et peuvent se limiter au transfert des métadonnées, qui fournissent des renseignements sur la situation en cours dans la scène.

Le flux de métadonnées permet une inspection visuelle de l'enveloppe sonore, représentée par exemple dans un tableau de bord qui associe les événements audio et vidéo et les affiche en parallèle. De cette manière, les métadonnées peuvent assurer une recherche efficace et rapide d'événements particuliers et de bruits sortant de l'ordinaire. Cette méthode peut épargner aux enquêteurs des heures de recherche sur les gros volumes de vidéo.

## 3 AXIS Audio Analytics

AXIS Audio Analytics est une plateforme logicielle qui recevra régulièrement des fonctions d'analyse nouvelles ou améliorées. Les fonctions d'analyse sont intégrées au logiciel AXIS OS des dispositifs et fournies gratuitement avec les caméras et d'autres dispositifs Axis.

Les fonctions d'analyse de la plateforme sont exécutées localement et les algorithmes opèrent directement dans la caméra. Cette méthode se traduit par une évolutivité optimale, un faible trafic de données et le maintien de la confidentialité. Seul le résultat de l'analyse (métadonnées ou déclencheurs) est enregistré. Aucun son n'est enregistré ou diffusé depuis la caméra, et le son d'origine n'est pas récupérable à partir des métadonnées.

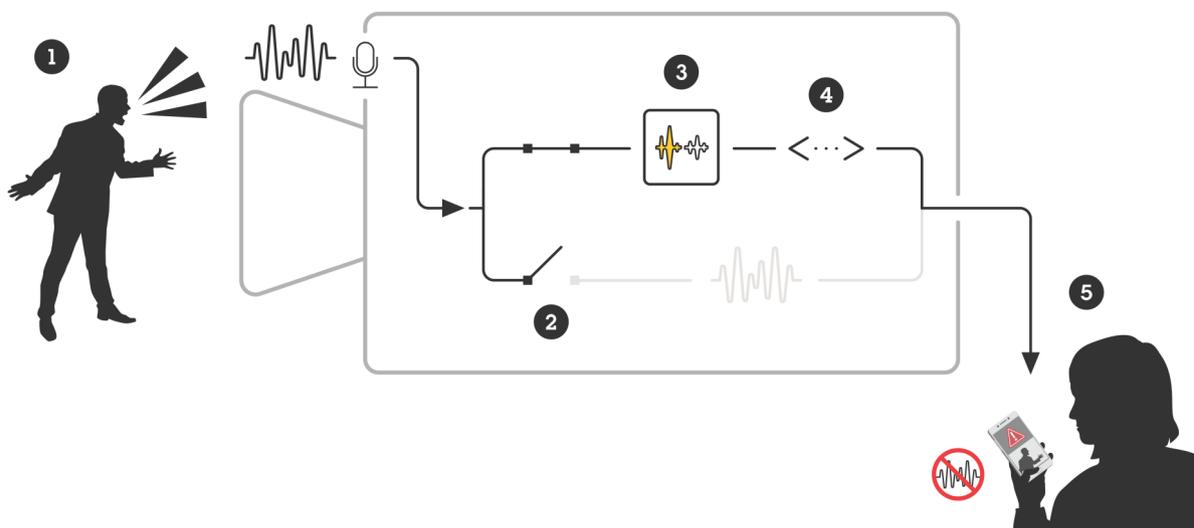
- **Classification audio (cris et éclats de voix).** Cette fonction de classification sonore basée sur l'IA détecte et analyse des bruits particuliers tels que cris et éclats de voix. D'autres classifications de bruits, par exemple les bris de verre, seront ajoutées ultérieurement. Ces fonctions de classification produisent des métadonnées qui décrivent les caractéristiques du bruit.
- **Détection audio adaptative.** Un détecteur audio crée un événement lorsque le volume sonore augmente brusquement. Il détecte les pics sonores de toutes sortes, avec pour avantage de s'adapter au bruit ambiant même lorsqu'il varie.

### 3.1 Confidentialité

En général, les fonctions d'analyse audio n'enregistrent pas le son et ne le transmettent pas hors de la caméra. Elles se limitent à traiter les bruits pour faciliter la recherche d'événements, de motifs acoustiques ou de volumes sonores spécifiques dans un système récepteur, par exemple un tableau de bord pour investigations plus poussées ou un logiciel de gestion vidéo pour alerter les opérateurs. Il est impossible de reconstruire les données audio ou d'enregistrer des conversations privées. Et pour cause, ces fonctions d'analyse résident en périphérie de réseau et fournissent des métadonnées audio.

AXIS Audio Analytics n'enregistre pas de son et n'en transmet pas ; la fonction transmet uniquement des métadonnées. Pour des raisons de confidentialité, toutes les fonctions de streaming audio sont également désactivées par défaut dans les dispositifs Axis (contrôle de confidentialité audio). Concrètement, le son n'est ni diffusé en streaming, ni enregistré, et il n'est pas possible de le recréer. Vous pouvez activer le streaming audio si nécessaire, mais même s'il est désactivé, vous êtes averti lorsqu'un événement pertinent se produit.

L'image illustre le principe de fonctionnement d'AXIS Audio Analytics avec le contrôle de confidentialité audio pour détecter les bruits et utiliser les métadonnées pour générer une alerte.



*Contrôle de confidentialité audio et AXIS Audio Analytics dans les caméras Axis, de l'émission d'un cri (1) à une alerte (5).*

- 1 Les bruits sont détectés par le microphone.*
- 2 Par défaut, le streaming audio est désactivé.*
- 3 La fonction de classification audio d'AXIS Audio Analytics détecte un cri ou un éclat de voix dans l'entrée audio.*
- 4 La fonction de classification audio génère des métadonnées, notamment une notification d'événement.*
- 5 Les personnes concernées reçoivent une alerte basée sur la notification d'événement et les métadonnées. Elles peuvent vérifier la situation en visionnant le flux vidéo. Aucun flux audio n'est disponible.*

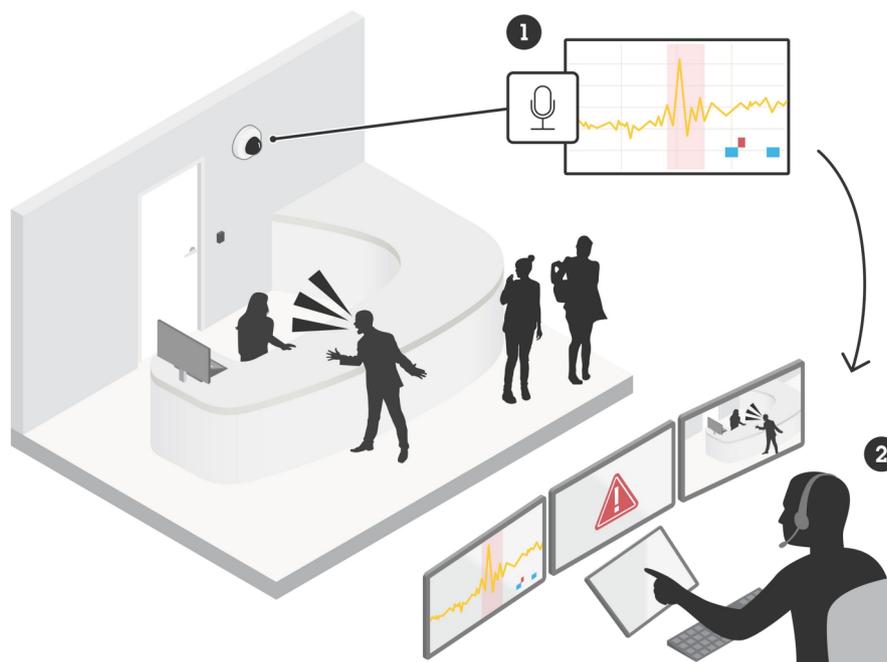
## 3.2 Scénarios d'utilisation

Bien que les fonctions d'analyse basées sur l'IA puissent potentiellement ignorer les bruits sans intérêt, elles peuvent produire des fausses alertes si le bruit ambiant est excessif. La pluie frappant les vitres, le tonnerre, les sirènes, la musique ou les discussions animées peuvent déclencher de fausses alertes. Par conséquent, les scénarios d'utilisation typiques sont notamment les lieux calmes, comme les banques ou les halls d'accueil, mais aussi une diversité d'espaces intérieurs après les horaires d'ouverture, notamment magasins, restaurants, cages d'escalier ou bureaux.

### 3.2.1 Détection de bruit avec alertes

Dans une banque ou à un comptoir d'accueil, une fonction de classification des bruits peut écouter la zone et détecter les sons associés à une agression ou à un bris de verre. Après détection, les données d'événement et les métadonnées audio du système envoient des notifications automatiques au personnel

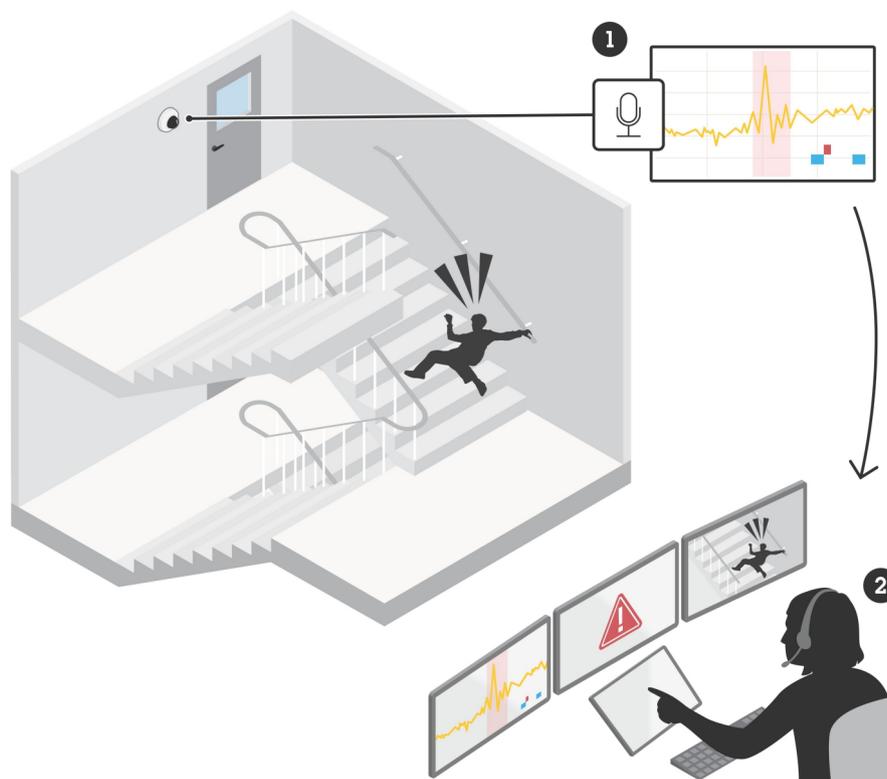
par alerte visuelle ou déclenchement d'une alarme. Cet avertissement rapide permet des réponses et une intervention rapides.



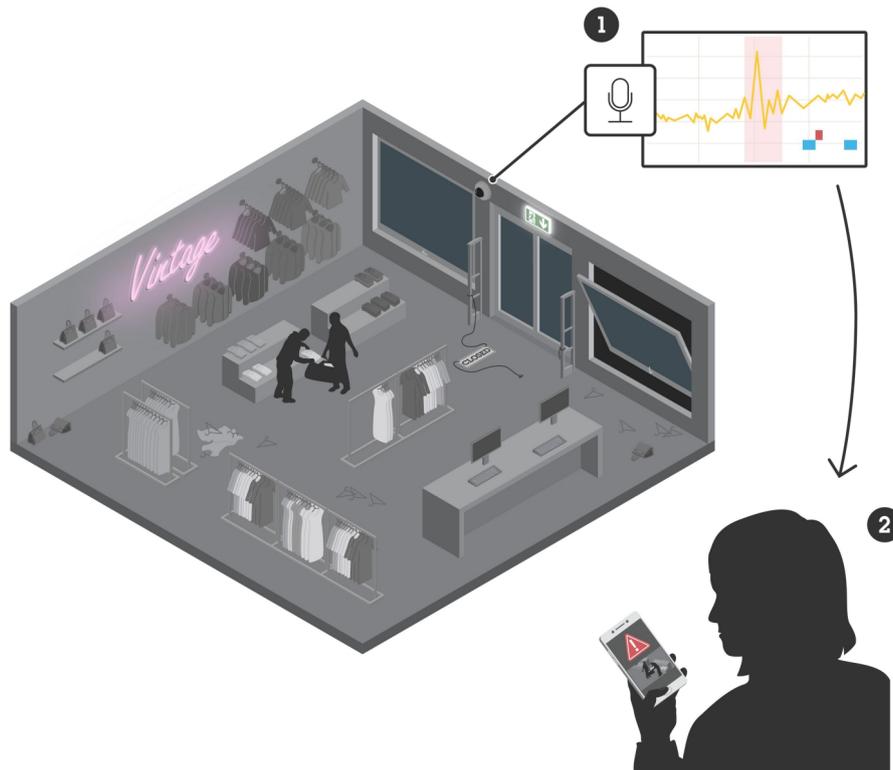
- 1 Une caméra exécutant une fonction de classification des bruits détecte des cris ou des éclats de voix au comptoir d'accueil.
- 2 Un opérateur reçoit une alerte et peut vérifier la situation en observant le flux vidéo avant de prendre les mesures qui s'imposent.

Les fonctions de détection audio adaptatives peuvent servir à détecter les bruits inhabituels en dehors des horaires d'ouverture. Ce type de fonction analyse les bruits ambiants et réagit lorsqu'elle détecte des voix,

des bris de vitre ou d'autres bruits soudains de courte durée. Après détection d'un événement, la fonction d'analyse transmet les métadonnées pour en aviser les opérateurs.



- 1 Une caméra exécutant une fonction de classification des bruits détecte des bruits inhabituels après les horaires d'ouverture.
- 2 Un opérateur reçoit une alerte et peut vérifier la situation en observant le flux vidéo avant de prendre les mesures qui s'imposent.



- 1 Une caméra exécutant une fonction de classification des bruits détecte des bruits dans une boutique après les horaires d'ouverture.
- 2 Le gérant de la boutique reçoit une alerte et peut vérifier la situation en observant le flux vidéo avant de prendre les mesures qui s'imposent.

### 3.2.2 Association de capteurs pour tirer le maximum de votre système de surveillance

Les systèmes de surveillance intègrent souvent plusieurs types de capteurs. Le capteur d'image de la caméra est l'un de ces capteurs. Bien entendu, il enregistre l'aspect visuel d'une scène. Des capteurs non visuels sont également généralement utilisés, tels que les détecteurs de mouvement basés sur la technologie du radar ou les émissions de radiation infrarouge. Les capteurs non visuels complètent le système de caméras en y associant d'autres informations.

En employant également des capteurs audio (microphones) dans une installation de surveillance, la grande majorité de tous les cas d'usage possibles sont renforcés. L'ajout de fonctionnalités audio et de fonctions d'analyse acoustique à un système sans audio permet de faire interagir plusieurs capteurs. Si vous utilisez des fonctions d'analyse vidéo, l'ajout de fonctions d'analyse audio peut accroître la fiabilité de la détection. C'est particulièrement le cas lorsque par exemple une luminosité insuffisante complique le travail des fonctions d'analyse vidéo.

Vous pouvez configurer le système, par exemple dans le logiciel de gestion vidéo, pour qu'il déclenche des actions uniquement lorsque les fonctions d'analyse vidéo et les fonctions d'analyse audio ont toutes deux

réagi : par exemple, l'analyse audio détecte un cri et l'analyse vidéo détecte un individu dans le champ de vision de la caméra. Dans certains environnements, cette association offre le juste niveau de sécurité.

### **3.2.3 Source d'informations pour les tableaux de bord**

Les métadonnées audio peuvent servir d'informations d'entrée pour les tableaux de bord analytiques des plateformes VSaaS (VideoSurveillance-as-a-Service), IoT (Internet of Things) ou de business intelligence, qui recueillent et représentent visuellement les métadonnées. Ces plateformes analysent les données en temps réel et les tendances pour produire instantanément une vue d'ensemble et des éclairages concrets. Les analyses statistiques basées sur les flux de clients ou l'expérience client facilitent les prises de décision pilotées par les données pour améliorer les opérations.

Avec les tableaux de bord, vous pouvez visualiser les résultats sans jamais écouter l'audio associé, et même sans aucune possibilité de récupérer le son d'origine. En revanche, vous pouvez obtenir des renseignements concrets, tirés par exemple du comptage des événements, sans porter préjudice à la confidentialité des données. Notez que les restrictions juridiques peuvent varier selon que le son est enregistré ou seulement capturé.

## **4 Restrictions juridiques**

Nombreux sont ceux qui s'inquiètent du recours aux microphones en vidéosurveillance. Ces préoccupations sont généralement liées à l'enregistrement de conversations avec la vidéo. Cependant, avec les fonctions d'analyse audio, le son n'est généralement ni enregistré, ni transmis. Les lois qui régissent la surveillance varient selon la région et le pays. Vous devez donc bien vous informer des pratiques autorisées avant d'utiliser l'audio dans votre système de surveillance.

La législation nationale et divers types de règlements locaux peuvent interdire la capture et l'enregistrement du son ou leur imposer des contraintes particulières pour plusieurs raisons. Une région ou un environnement qui autorise la capture audio pourrait très bien interdire les enregistrements audio. Les entreprises peuvent également proscrire la surveillance audio dans leurs locaux.

## **5 Clause de non-responsabilité**

Le présent document et son contenu sont offerts par Axis et tous les droits sur le document ou tout droit de la propriété intellectuelle s'y afférant (y compris mais sans s'y limiter les marques déposées, les noms commerciaux, les logos et les symboles similaires stipulés aux présentes) sont protégés par la loi et tous les droits, titre et/ou intérêt au présent document ou tout droit de la propriété intellectuelle s'y afférant sont et resteront acquis à Axis Communications AB.

Veillez noter que le présent document est fourni « en l'état » sans garantie d'aucune nature à des fins d'informations uniquement. Les informations fournies dans le présent document ne constituent pas des conseils de nature légale et ne sont pas destinées à constituer des conseils de nature légale. Le présent document n'est pas destiné et ne devra pas être destiné à créer d'obligation légale de quelque nature que ce soit pour Axis Communications AB et/ou ses filiales. Les obligations d'Axis Communications AB et/ou de ses filiales relatives aux produits Axis sont soumises exclusivement aux conditions générales de l'accord entre Axis et l'entité qui s'est portée acquéreur desdits produits directement auprès d'Axis.

**POUR LEVER TOUTE AMBIGUÏTÉ, L'UTILISATEUR DU PRÉSENT DOCUMENT ASSUME L'INTÉGRALITÉ DU RISQUE LIÉ À L'UTILISATION, AUX RÉSULTATS ET AUX PERFORMANCES LIÉS À CE DOCUMENT, ET AXIS REJETTE ET EXCLUT, DANS LA LIMITE MAXIMALE AUTORISÉE PAR LA LOI, TOUTE GARANTIE, QU'ELLE SOIT STATUTAIRE, EXPRESSE OU IMPLICITE, Y COMPRIS, MAIS SANS S'Y LIMITER, TOUTE GARANTIE IMPLICITE**

DE QUALITÉ MARCHANDE, D'ADÉQUATION À UN USAGE PARTICULIER, DE TITRE ET D'ABSENCE DE CONTREFAÇON ET DE RESPONSABILITÉ DU FAIT DES PRODUITS, OU TOUTE GARANTIE DÉCOULANT DE TOUTE PROPOSITION, SPÉCIFICATION OU ÉCHANTILLON EN RAPPORT AVEC CE DOCUMENT.

# À propos d'Axis Communications

En concevant des solutions qui améliorent la sécurité et les performances de l'entreprise, Axis crée un monde plus clairvoyant et plus sûr. En tant qu'entreprise de technologie de réseau et leader de l'industrie, Axis propose des solutions de vidéosurveillance, de contrôle d'accès, d'interphonie et de systèmes audio. Les performances de ces solutions sont améliorées grâce à des applications d'analyse intelligentes et une formation de haute qualité.

Axis emploie près de 4 000 personnes dans plus de 50 pays et collabore avec des partenaires technologiques et d'intégration de systèmes dans le monde entier pour fournir des solutions clients adaptées. Axis a été fondée en 1984 et le siège social se trouve à Lund, en Suède.