

WHITE PAPER

자연어 검색

AXIS Camera Station Pro

4월 2026

요약

AXIS Camera Station Pro에는 여러 비디오 포렌식 검색 도구가 사전 설치되어 있습니다. 여기에는 타임라인 스크리빙, 데이터 검색, 사전 분류된 객체를 사용한 스마트 검색 및 자연어 검색이 포함됩니다.

자연어 검색을 사용하면, 움직이는 객체를 사용자 고유의 표현을 사용해 설명하여 검색할 수 있습니다. 다양한 설명 속성을 활용하여 상세한 검색 필터를 자유롭게 생성할 수 있어 관련 영상을 더 빠르게 찾을 수 있습니다.

자연어 검색 기능은 Axis가 감시 사용 사례에 맞게 최적화된 사전 훈련된 오픈 소스 기반 모델이 제공하는 텍스트-이미지 매칭을 기반으로 합니다. 검색은 한 대의 카메라 또는 여러 대의 카메라에 동시에 적용할 수 있습니다.

자연어 검색의 수치적 표현은 감지된 움직이는 객체 이미지의 수치적 표현과 비교됩니다. 검색 결과는 카메라 이름, 시간 및 날짜를 포함한 썸네일로 표시되며, 검색어와의 관련성에 따라 정렬됩니다.

자연어 검색을 통해, Axis는 AI를 사용하여 포렌식 검색 솔루션의 정확성과 효율성을 높이고 궁극적으로 인간의 의사 결정을 향상시킵니다. 법적 및 윤리적 기준을 준수하기 위해, 검색 기능에는 검색어에 불쾌감을 주는 단어를 사용하는 것을 제한하는 Axis가 개발한 별도의 조정 기능이 포함되어 있습니다. 또한 모든 검색은 기록되어 관리자가 확인할 수 있어, 오용 사례 발생 시 후속 조치와 시정 조치를 취할 수 있습니다.

목차

1	서론	4
2	배경: AXIS Camera Station Pro의 포렌식 검색	4
3	자연어 검색은 어떻게 작동합니까?	4
4	검색어 작성	5
4.1	프롬프트 지침	5
4.2	검색어 조정	5
4.3	검색어 로깅	6
5	텍스트-이미지 매칭	6
5.1	감시에 최적화된 파운데이션 모델	6
5.2	특징 벡터	6
6	프로세스 개요	7
7	책임감 있는 AI 사용	8

1 서론

AXIS Camera Station Pro의 자연어 검색 도구를 사용하면, 사전 정의된 필터 없이 원하는 단어를 입력하여 비디오 녹화 영상을 검색할 수 있습니다.

이 백서에서는 검색 방법의 작동 원리를 설명하고 검색 방법을 사용하는 방법에 대한 몇 가지 가이드 라인을 제시합니다. 법적 및 윤리적 기준을 준수하기 위해 마련된 중재 기능과 검색어 로깅에 대해서도 설명합니다.

2 배경: AXIS Camera Station Pro의 포렌식 검색

AXIS Camera Station Pro에는 타임라인 스크러빙, 데이터 검색은 물론 사전 분류된 객체를 사용한 스마트 검색, 시각적 유사성 및 자연어 검색을 포함한 여러 비디오 포렌식 검색 도구가 사전 설치되어 있습니다.

스마트 검색 기능은 Axis 장치에서 생성된 장면 메타데이터를 사용합니다. 메타데이터에는 움직이는 객체에 대한 객체 유형(사람, 차량 유형 또는 알 수 없는 객체)과 함께 의류 및 차량 색상, 번호판, 속도, 위치, 타임스탬프와 같은 속성이 포함됩니다.

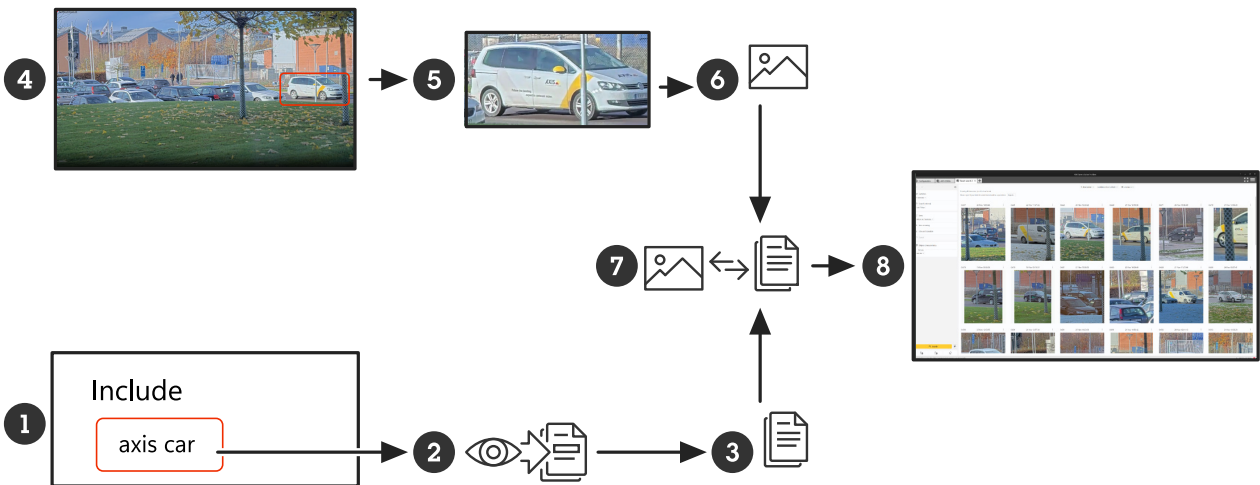
분석 기능이 제한된 장치의 경우, 검색 기능은 장치의 모션 디텍션과 AXIS Camera Station Pro 서버에서 수행된 객체 분류를 기반으로 합니다. 따라서 AXIS Camera Station Pro의 포렌식 검색은 에지 장치의 기능을 최대한 활용하되 필요한 경우 서버의 데이터로 보완하는 하이브리드 솔루션입니다.

기존에는 장면 메타데이터를 사용한 검색은 미리 정의된 검색 필터를 사용하여 수행해야 했습니다. 이를 통해 객체 유형(예: "차량"), 차량 유형(해당되는 경우 "자동차"), 색상(예: "파란색") 등을 포함하여 고정된 객체 설명자를 목록에서 선택할 수 있습니다. 대신 새로운 자연어 검색 방식을 사용하면 고유한 검색 필터를 만들 수 있습니다.

사전 분류된 검색은 높은 정확도의 결과를 제공하지만, 미리 정의되지 않은 새로운 객체 유형은 감지할 수 없습니다. 이러한 제한을 해결하기 위해 자연어 검색은 사용자가 직접 원하는 단어를 입력하여 유연하게 검색할 수 있도록 지원합니다. 자연어와 연관어를 사용해 움직이는 객체를 더 자세하게 설명하여 검색을 세밀하게 조정하고 더 많은 결과를 얻을 수 있습니다.

3 자연어 검색은 어떻게 작동합니까?

자연어 검색의 수치적 표현은 감지된 움직이는 객체 이미지의 수치적 표현과 비교됩니다. 이 텍스트-이미지 매칭의 결과는 일치도 순으로 표시되고 정렬됩니다. 검색 결과는 카메라 이름, 시간 및 날짜를 포함한 썸네일로 표시되며, 검색어와의 관련성에 따라 정렬됩니다.



자연어 검색 프로세스의 간소화된 개요. 백그라운드 처리가 활성화되어 있으면 검색하지 않을 때도 6 단계가 지속적으로 수행됩니다.

- 1 자연어 검색어를 입력합니다.
- 2 조정 모듈은 유해하고 비윤리적인 단어의 사용을 방지합니다.
- 3 파운데이션 모델은 검색어의 숫자 표현(특징 벡터)을 생성합니다.
- 4 카메라가 장면의 움직임을 감지합니다.
- 5 카메라가 움직이는 객체를 표현하기 위해 잘린 이미지 하나를 선택합니다.
- 6 파운데이션 모델은 객체의 모양, 패턴, 색상 등을 분석한 후 특징 벡터를 생성합니다.
- 7 두 기능 벡터를 비교합니다.
- 8 비교 결과는 일치도 순으로 정렬되어 썸네일로 표시됩니다.

자연어 검색은 한 대의 카메라 또는 여러 대의 카메라에 동시에 적용할 수 있습니다.

자연어 검색의 범위를 줄이기 위해서는 유사도 검색이나 시간 기반 검색과 같은 다른 스마트 검색 기능을 하나씩 차례로 적용하여 조합할 수 있습니다.

4 검색어 작성

움직이는 모든 객체와 모든 유형의 차량을 검색할 수 있습니다. 정지된 객체를 검색하면, 대부분의 경우 검색이 이루어지지 않습니다. 최상의 결과를 얻으려면 가이드라인을 따르십시오.

비윤리적인 검색 행위를 방지하기 위해 검색 구문이 조정되고 기록됩니다.

4.1 프롬프트 지침

- 이미지를 설명할 때와 마찬가지로 상황을 설명합니다. 모델에는 정지 이미지가 제공되므로 액션(예: 넘어지기, 달리기, 도둑질)을 검색하려면 더 많은 컨텍스트가 필요하기 때문에 액션 검색이 어려울 수 있습니다.
- “빨간 스웨터를 입은 사람” 또는 “노란색 픽업 트럭” 등 몇 가지 주요 설명자를 사용하여 객체를 설명합니다. 다른 멀티모달 모델과 마찬가지로 자연어 검색 모델은 물체 및 색상과 같은 설명자에 대해서는 잘 작동하지만, 숫자(“세 사람”), 속어 또는 감정 표현(“화난 표정의 남자”)에는 덜 적합합니다. 객체 설명은 주관적이거나 너무 모호하거나 너무 구체적인 디테일을 포함해서는 안 됩니다.
- *and*를 사용하여 여러 객체 속성을 결합: “빨간 모자를 쓰고 배낭을 멘 사람”
- 텍스트, 텍스트 로고 또는 브랜드 이름을 설명합니다(예: “Axis라는 텍스트가 있는 밴”).
- 환경을 설명하는 데 초점을 맞추지 마십시오. 객체의 잘린 이미지를 처리하기 때문에 모델이 객체의 주변을 보지 못할 수도 있습니다. 따라서 광범위한 장면이나 환경 설명자(예: “도시”, “도심”, “공원”, “정원”, “호수” 또는 “해변”)는 좋은 결과를 제공하지 않을 수 있습니다.
- 결과가 마음에 들지 않으면 다른 문구로 실험합니다.
- 자연어 프롬프트는 영어만 지원합니다.

4.2 검색어 조정

Axis는 효과적인 필터링을 위해 일반적인 관행에 기반한 검색어 조정을 구현했습니다. 모더레이션 모델은 검색 쿼리를 확인하여 부적절한 단어의 사용을 제한하는 자연어 처리 모델입니다. 이 모델은 전체 텍스트 문자열에서 유해하거나 부적절하거나 유해한 콘텐츠를 검사합니다.

뿐만 아니라, Axis는 금지된 검색 카테고리 및 단어의 사용자 정의 목록을 포함한 독자적인 조치를 통해 이러한 기능을 강화했습니다. 이 목록은 EU 인공지능법(EU Artificial Intelligence Act)의 취지를 따르면서도 더 넓은 범위를 포괄합니다. 이 목록에는 일반적으로 다음과 관련된 단어들도 포함됩니다.

- 비속어(민족적 및 종교적 비속어나 모욕적인 단어 포함)
- 민족(국적, 국가)
- 정치적, 종교적, 철학적 신념(종교적 의복 포함)
- 감정(예: “화남” 등)

- 주관적인 부정적 묘사 및 분류(예: "범죄자" 등)
- 성적 지향, 콘텐츠 또는 암시

EU 인공지능법은 개인의 인증, 정치적 견해, 노동조합 가입 여부, 종교적 또는 철학적 신념, 성생활 또는 성적 지향을 추론하기 위해 생체 인식 분류 시스템을 사용하는 것을 금지하고 있습니다.

검색 쿼리에 해당 목록의 단어나 문구가 포함된 경우, 안전한 검색 환경을 유지하고 윤리적인 결과를 보장하기 위해 검색을 거부합니다. 특정 단어가 차단되는 것에 동의하지 않거나 차단이 필요한 단어를 제안하고 싶은 경우 Axis에 익명으로 사용자 피드백을 제공할 수 있습니다.

4.3 검색어 로깅

AXIS Camera Station Pro는 사용자 운영에 대한 감사 추적을 유지합니다. 감사 추적은 특정 운영과 사용자 신원을 추적할 뿐만 아니라 해당 운영에 사용된 모든 데이터도 보관합니다. 즉, 검색 프롬프트를 포함한 모든 사용자 검색이 기록됩니다. 관리자는 로그를 사용하여 사용자의 부적절한 검색 활동을 식별하고, 비윤리적인 검색 프롬프트에 플래그를 지정하고, 시정 조치를 취할 수 있습니다.

비디오 데이터는 Axis와 공유되지 않습니다. 데이터는 서버에 남아 있습니다.

5 텍스트-이미지 매칭

자연어 검색을 사용해 비디오 메타데이터를 검색할 수 있게 되면 미리 정의된 속성 목록에서 거의 무제한의 검색 기준으로 검색 능력이 크게 확장됩니다. AXIS Camera Station Pro에서 이 기능은 수십억 개의 이미지-텍스트 쌍으로 학습된 오픈 소스 파운데이션 모델을 토대로 하며, Axis가 감시 사용 사례에 맞춰 성능 향상을 위해 미세 조정을 수행했습니다.

텍스트-이미지 매칭 및 특징 벡터 생성 속도는 서버 성능의 영향을 직접적으로 받으며, 하드웨어 사양이 높을수록 더 빠른 검색 결과를 제공합니다. AXIS Camera Station S1228 AI-Optimized Server는 GPU 가속 기능을 갖추고 있으며, 스마트 검색 내의 자연어 검색 성능을 극대화하기 위해 맞춤 설계되었습니다.

5.1 감시에 최적화된 파운데이션 모델

파운데이션 모델은 텍스트-이미지 쌍의 대규모 데이터셋으로 학습된 텍스트-이미지 모델입니다. 이 모델은 텍스트와 관련 이미지를 매칭하는 제로 샷 모델입니다. 제로 샷 모델은 사전 학습 데이터 없이도 객체나 개념을 인식하고 분류할 수 있는 인공지능(AI) 모델의 일종입니다. 즉, 이 모델은 이전에 해당 작업의 예제를 본 적이 없어도 작업을 수행할 수 있습니다. 이러한 능력은 Axis가 자연어와 이미지를 매칭하기 위해 최적의 성능을 제공하는 데 매우 중요합니다.

이 모델은 대량의 텍스트-이미지 조합으로 학습되었으며, 25억 개 이상의 매개변수를 가진 신경망에서 작동합니다. Axis에서는 자체적인 고유한 학습 자료를 사용하여 이 모델을 더욱 조정했으며, 이를 통해 일반적인 감시 카메라 보기와 객체가 포함된 이미지를 해석하는 능력을 향상시켰습니다. 즉, Axis는 이 모델을 감시 사용 사례에 맞게 최적화했습니다.

5.2 특징 벡터

자연어 검색을 하면 파운데이션 모델이 검색어의 특징 벡터를 생성합니다.

파운데이션 모델은 장면에서 추적되는 모든 객체에 대한 설명을 생성하며, 해당 객체를 표현하는 특징 벡터를 만듭니다. 각 객체는 단 하나의 특징 벡터로 표현되며, 이 벡터는 데이터베이스에 저장됩니다. 백그라운드 처리가 활성화되면 특징 벡터가 지속적으로 생성됩니다. 이에 따라 특징 벡터가 사전에 계산되어 데이터베이스에 준비되어 있으므로, 검색 속도가 빠릅니다.

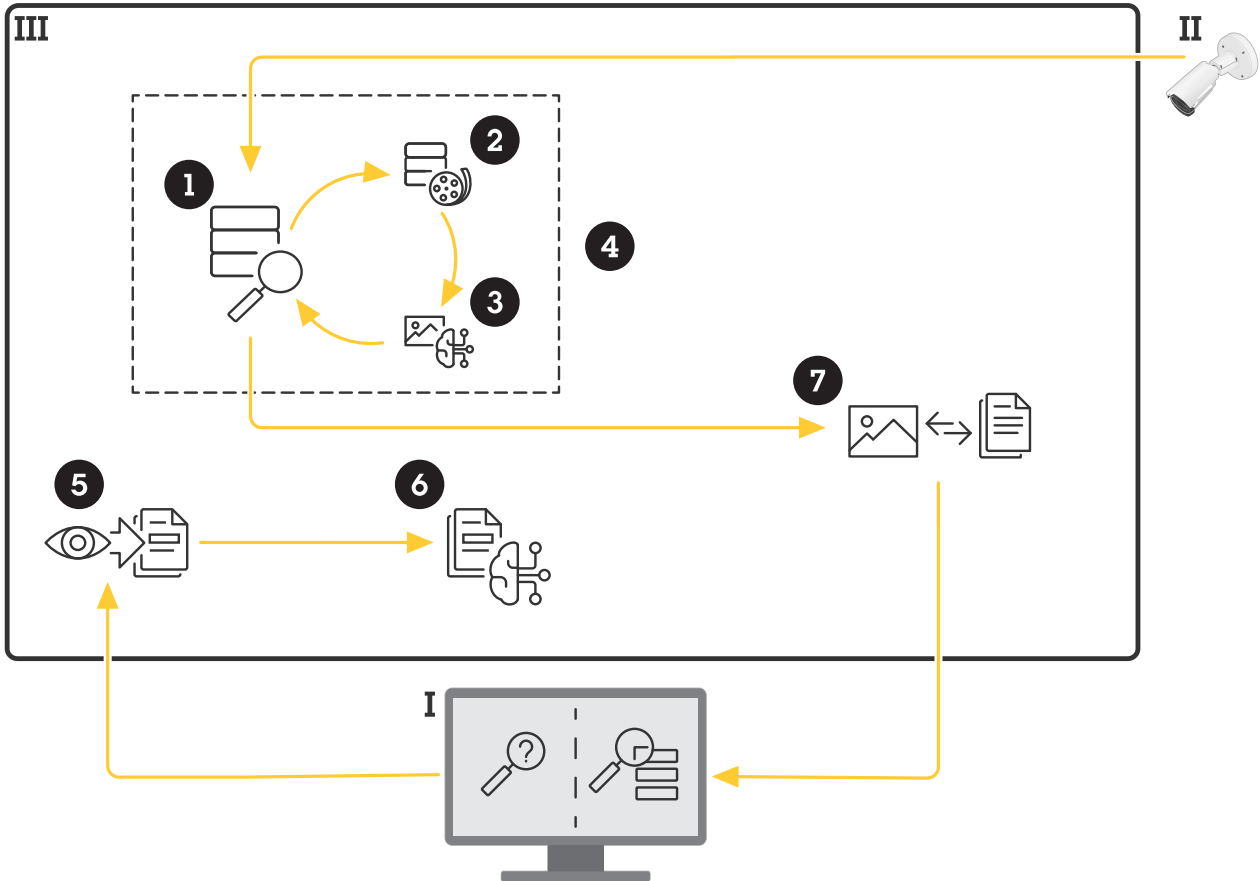
검색어와 감지된 객체를 나타내는 모든 가용 특징 벡터 간의 유사도 거리를 계산하기 위해, 두 가지 유형의 특징 벡터가 벡터 비교 엔진에 입력됩니다.

특징 벡터는 텍스트나 이미지를 숫자로 표현한 것입니다. 따라서 사람이나 객체의 특징 벡터는 사람이나 객체의 외형을 추상적으로 표현한 것에 불과합니다. 특징 벡터에는 머리카락이나 의상 색상과 같이 특정 개인을 식별하거나 매핑하는 데 사용할 수 있고 사람이 해석 가능한 특징 정보가 포함되지 않습니다. 특징 벡터는 다른 특징 벡터와 비교하는 용도로만 사용할 수 있습니다.

6 프로세스 개요

프로세스 개요도는 주요 프로세스 단계와 해당 단계가 수행되는 위치를 보여줍니다.

백그라운드 처리를 활성화한 경우, 다이어그램의 루프(4)는 검색할 때뿐만 아니라 감지된 모든 움직임은 객체의 특징 벡터를 지속적으로 생성하는 프로세스입니다.



자연어 검색 프로세스의 주요 위치(I~III) 및 주요 단계(1~7)

(I) AXIS Camera Station Pro 클라이언트: 여기에 검색어를 입력하면 정렬된 검색 결과를 받을 수 있습니다.

(II) 카메라: 자연어 검색은 AXIS OS 5.51 이상을 탑재한 Axis 카메라에서 작동하지만, 카메라가 좋을수록 더 좋은 결과를 얻을 수 있습니다. 구형 장치는 모션 디텍션만을 기반으로 덜 세분화된 메타데이터를 제공합니다. 최신 장치는 객체 분류를 포함하는 AXIS Scene Metadata의 메타데이터를 생성합니다. 카메라의 움직임은 객체 감지 기능은 감지된 각 객체의 대표 이미지를 하나씩 찾아 서버에서 분석할 이미지의 수를 줄이는 데 사용됩니다.

(III) AXIS Camera Station Pro 서버: 여기에서 카메라의 모든 메타데이터와 비디오 데이터가 처리되고 저장됩니다. 자연어 검색을 하기 전에 서버는 (감지된 각 움직임은 객체에 대해) 비디오를 디코딩하고 감지된 객체의 이미지를 추출해야 합니다. 그런 다음 파운데이션 모델이 이 이미지를 처리하여 특징 벡터를 생성합니다. 이러한 작업은 처리 용량 측면에서 비용이 많이 들기 때문에, 성능 향상을 위해 기능 벡터를 데이터베이스에 저장하여 향후 빠른 검색을 지원할 수 있도록 합니다. 서버에 여유 용

량이 있으면 가장 중요한 카메라의 영상을 백그라운드에서 처리하도록 설정하는 것을 강력히 권장합니다. 이를 통해 검색 속도를 크게 향상시킬 것이기 때문입니다.

(1) 데이터베이스 검색: 검색 데이터베이스에는 속성, 시간, 위치, 특징 벡터가 포함된 객체 분류 데이터를 포함하여 AXIS Scene Metadata의 전체 메타데이터 또는 서버에서 생성한 메타데이터가 저장됩니다.

(2) 녹화 영상 스토리지: 카메라의 영상이 저장되는 곳이며 파운데이션 모델이 이미지를 가져오는 곳입니다.

(3) 이미지 파운데이션 모델: 이 파운데이션 모델은 녹화 영상에 있는 각 객체 트랙에 대한 수치적 표현(특징 벡터)을 생성합니다.

(4) 백그라운드 처리: 백그라운드 처리가 활성화되어 있으면, 1~3단계는 검색하지 않을 때도 수행되는 연속 프로세스를 구성합니다.

(5) 검색어 조정: 조정 모델은 검색어를 검사하여 불쾌감을 주는 콘텐츠를 제한합니다.

(6) 텍스트 파운데이션 모델: 이 파운데이션 모델은 조정된 검색어 텍스트 문자열의 수치 표현(특징 벡터)을 생성합니다.

(7) 벡터 비교: 검색어 텍스트 문자열의 특징 벡터 표현을 비디오에서 감지된 객체 이미지의 특징 벡터 표현과 비교합니다.

7 책임감 있는 AI 사용

자연어 검색을 통해, Axis는 AI를 사용하여 포렌식 검색 솔루션의 정확성과 효율성을 높이고 궁극적으로 인간의 의사 결정을 향상시킵니다.

책임감과 신뢰성은 Axis의 AI 접근 방식에서 핵심 요소입니다. 여기에는 Axis가 만드는 AI 시스템이 윤리적 원칙을 준수하고, 법률을 준수하며, 위험을 효과적으로 관리하도록 보장하는 것이 포함됩니다. Axis는 고객이 운영의 신뢰성을 확신할 수 있도록 하는 다양한 도구를 제공합니다. AXIS Camera Station Pro의 자연어 검색 기능에는 텍스트 프롬프트 조절을 위해 미세 조정된 텍스트 분류 모델이 포함되어 있습니다. Axis는 비윤리적인 사용의 방지를 도울 수 있도록 검색어를 조정하기 위해 이 모델을 개발했습니다.

자연어 검색은 일주일에 한 번 Axis 클라우드 서비스에 연결하여 새로운 규정이나 요구 사항을 준수하기 위해 AI 모델에 업데이트가 필요한지 확인합니다. 연결에 실패하면 연결이 다시 설정될 때까지 자연어 검색 작업을 사용할 수 없습니다.

AI 적용 시 법적 및 윤리적 기준을 더욱 준수하기 위해, Axis 제품은 사용자 인증 자격 증명 및 접근 권한에 기반한 접근 제어를 제공합니다. 이를 통해 고객은 사용자가 운영 정책을 준수하도록 강제할 수 있습니다.

Axis Communications에 대하여

Axis는 보안, 안전, 운영 효율성 및 비즈니스 인텔리전스를 향상시켜 더 스마트하고 더 안전한 세상을 실현합니다. 네트워크 기술 회사이자 업계 선도 기업인 Axis는 영상 감시, 접근 제어, 인터콤 및 오디오 솔루션을 제공합니다. 이러한 솔루션은 지능형 분석 애플리케이션으로 보완되고 고품질 교육을 통해 지원됩니다.

50개 이상의 국가에서 약 5,000명의 Axis 임직원이 전 세계의 기술 및 시스템 통합 파트너와 협력하여 고객에게 최적의 솔루션을 제공하고 있습니다. Axis는 1984년에 설립되었으며 본사는 스웨덴 룬드에 있습니다.