

BIAŁA KSIĘGA

# Uwolnić potencjał metadanych sceny

Wyższy poziom analizy i świadomości sytuacji

Styczeń 2024

# Streszczenie

W kontekście dozoru wizyjnego metadane opisują za pomocą tekstu zawartość materiału wizyjnego. Mogą to być np. widoczne w materiale interesujące obiekty lub ogólny opis samej sceny. Metadane mogą także obejmować atrybuty poszczególnych obiektów lub sceny, takie jak kolory pojazdów i ubrań, ich dokładne położenie albo kierunek ruchu. Metadane są tworzone w czasie rzeczywistym – bezpośrednio w kamerze albo przez inny komponent systemu mogący wykonywać analizy wideo.

Metadane zapewniają również kontekst zdarzeń oraz umożliwiają szybkie sortowanie i przeszukiwanie dużej ilości materiału wizyjnego. Otwiera to drogę do licznych zastosowań, które można z grubsza podzielić na trzy obszary:

- **Prace wyjaśniające po zdarzeniu.** Polegają na wyszukiwaniu interesujących obiektów lub zdarzeń na podstawie różnych parametrów, pozwalających zawęzić wyszukiwanie do ograniczonej liczby potencjalnych kandydatów. Dane dotyczące klasyfikacji obiektów umożliwiają wyszukiwanie uwzględniające szeroki zakres szczegółów.
- **Wykorzystywanie w czasie rzeczywistym.** Pomaga operatorom szybko reagować na zmieniające się warunki, wspiera proces decyzyjny lub umożliwia automatyczne wykonywanie działań.
- **Identyfikacja trendów, wzorców i spostrzeżeń.** Platformy z zakresu IoT i efektywności operacyjnej na potrzeby raportowania statystycznego mogą wykorzystywać metadane do celów zliczania gości, mierzenia prędkości, gromadzenia danych dotyczących ruchu i innych rodzajów automatycznego zbierania danych.

Niektóre kamery potrafią dekodować dźwięk i pozyskiwać metadane na jego temat. Określone wzorce dźwiękowe mogą być wykrywane i oznaczane w podobny sposób jak klasy obiektów w materiale wizyjnym. System rozpoznawania dźwięków umożliwia na przykład identyfikację agresji słownej lub wykrywanie przypadków stłuczenia szkła.

Połączenie metadanych pochodzących z wielu źródeł, takich jak źródła wizyjne, dźwiękowe oraz związane z aktywnością i procesami, pozwala pozyskiwać znacznie bogatszą wiedzę niż w przypadku każdego z tych źródeł osobno. Do płynnej integracji metadanych niezbędne są otwarte protokoły i standardy.

# Spis treści

1	Wprowadzenie	4
2	Co to są metadane?	4
3	Generowanie metadanych w urządzeniach brzegowych	4
4	Zastosowania	5
	4.1 Wykorzystywanie w czasie rzeczywistym do natychmiastowych działań	5
	4.2 Prace wyjaśniające	5
	4.3 Identyfikowanie trendów i wzorców w celu pozyskiwania wiedzy	6
5	Gdzie wykorzystuje się metadane?	6
6	Jak są dostarczane metadane?	7
7	Metadane dotyczące dźwięku	9
8	Łączenie metadanych z wielu źródeł	9

# 1 Wprowadzenie

Metadane stanowią podstawę do gromadzenia wiedzy z materiału wizyjnego. Przypisują one cyfrowe znaczenie treści materiału przez opisywanie najważniejszych elementów sceny. Dzięki metadanom można szybko znajdować i oceniać ważne elementy materiału wizyjnego oraz podejmować działania na ich podstawie. Z tego powodu metadane coraz częściej stanowią niezbędny składnik efektywnego systemu bezpieczeństwa i wsparcia działalności.

W tej białej księdze omówiono znaczenie metadanych w kontekście zarówno dozoru, jak i efektywności operacyjnej. Przedstawiono także zalety metadanych i opisano ich zastosowania w systemach zarządzania materiałem wizyjnym oraz innych aplikacjach.

## 2 Co to są metadane?

Metadane to dane na temat innych danych. W kontekście dozoru wizyjnego metadane opisują za pomocą tekstu zawartość materiału wizyjnego, np. widoczne w nim obiekty, lub dostarczają ogólny opis samej sceny. Mogą one obejmować atrybuty poszczególnych obiektów lub sceny, takie jak kolory pojazdów i ubrań, ich dokładne położenie albo kierunek ruchu. Metadane są tworzone w czasie rzeczywistym: albo bezpośrednio w kamerze, albo przez inny komponent systemu mogący wykonywać analizy wideo.

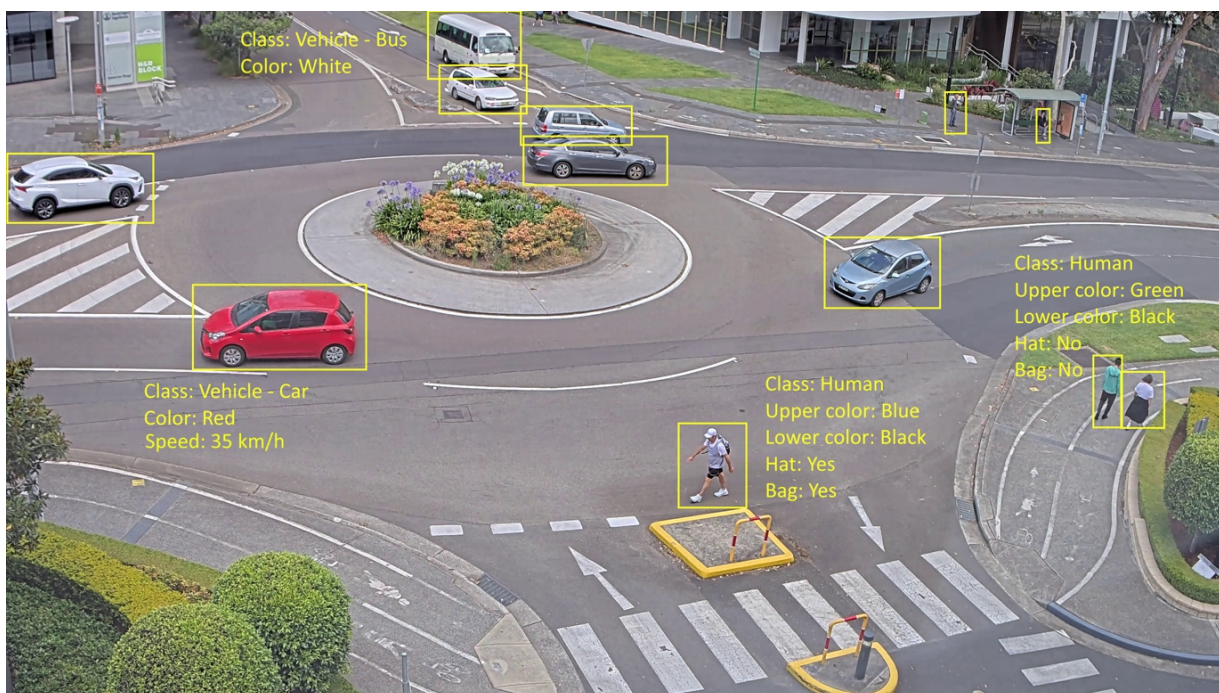


Figure 1. Przykład klatki wideo, w której interesujące obiekty są wykrywane i analizowane w celu utworzenia metadanych.

## 3 Generowanie metadanych w urządzeniach brzegowych

Dawniej wysokowydajne narzędzia do analiz wideo działały na serwerze, ponieważ ich potrzeby dotyczące mocy obliczeniowej przekraczały możliwości urządzeń brzegowych. Jednak w ostatnich latach rozwój algorytmów i wzrost mocy obliczeniowej urządzeń brzegowych umożliwiły prowadzenie zaawansowanych

analiz na brzegu sieci. Oznacza to, że metadane są generowane w urządzeniu i mogą być bezpośrednio w nim wykorzystywane przez inne narzędzia analityczne. Ponadto strumień wideo i strumień metadanych mogą być przekazywane do oprogramowania zarządzającego materiałem wizyjnym lub innej aplikacji w celu dalszego przetwarzania.

Brzegowe narzędzia analityczne mają dostęp do nieskompresowanego materiału wizyjnego z bardzo niewielkim opóźnieniem. Umożliwia to działanie w czasie rzeczywistym, a jednocześnie pozwala uniknąć dodatkowych kosztów i złożonych czynności związanych z przenoszeniem danych na potrzeby przetwarzania do innego miejsca w systemie. Ponadto analizy brzegowe wiążą się z niższymi kosztami sprzętu i wdrożenia, ponieważ system wymaga mniejszej ilości zasobów serwerowych.

Generowanie metadanych w urządzeniach brzegowych oznacza wydobywanie danych z materiału wizyjnego bez utraty informacji podczas kompresji czy przesyłania. Pozwala to pozyskiwać z materiału wizyjnego bardziej precyzyjne metadane i wykonywać na nim dokładniejsze analizy. Im lepsza jakość obrazu, tym dokładniejsze metadane.

## 4 Zastosowania

Metadane dostarczają nie tylko szczegółowych informacji na temat obiektów w danej scenie. Zapewniają również kontekst zdarzeń oraz umożliwiają szybkie sortowanie i przeszukiwanie dużej ilości materiału. Otwiera to drogę do licznych zastosowań, które można z grubsza podzielić na prace wyjaśniające prowadzone po zdarzeniu, wykorzystanie w czasie rzeczywistym oraz identyfikację trendów, wzorców i spostrzeżeń.

### 4.1 Wykorzystywanie w czasie rzeczywistym do natychmiastowych działań

Dzięki wykorzystaniu metadanych w czasie rzeczywistym operatorom łatwiej jest szybko reagować na zmieniające się warunki. Metadane to również cenny zasób, który wspiera proces decyzyjny lub umożliwia automatyczne wykonywanie działań. Działające w czasie rzeczywistym brzegowe narzędzia analityczne, które korzystają z wysokiej jakości metadanych, pomagają w dbaniu o bezpieczeństwo ludzi, zakładów i budynków oraz chronieniu ich przed zamierzonymi i przypadkowymi szkodami. Dzięki szybkiej detekcji, weryfikacji i ocenie zagrożeń można sprawnie na nie reagować.

### 4.2 Prace wyjaśniające

Metadane umożliwiają szybkie i skuteczne wyszukiwanie interesujących obiektów oraz zdarzeń. Dzięki temu śledczy mogą zaoszczędzić wiele godzin, zwłaszcza w przypadku przeszukiwania ogromnych ilości materiału wizyjnego pochodzącego z wielu źródeł. Możliwe jest wyszukiwanie takich obiektów jak ludzie czy pojazdy z wykorzystaniem różnych parametrów, które pozwalają zawęzić tę operację do ograniczonej liczby potencjalnych kandydatów. Parametry wyszukiwania mogą obejmować na przykład ruch, czas i cechy obiektu.

Tabela 4.1 Wyszukiwanie jest łatwiejsze dzięki różnym kategoriom metadanych.

Kategorie metadanych	Przedmiot wyszukiwania	Przykłady
Ruch	Jak obiekt się porusza	Kierunek, prędkość, inne zachowania
Czas	Kiedy obiekt się pojawia	Dzień tygodnia, godzina, czas zatrzymania

Tabela 4.1. Wyszukiwanie jest łatwiejsze dzięki różnym kategoriom metadanych. (Kontynuacja)

Kategorie metadanych	Przedmiot wyszukiwania	Przykłady
Lokalizacja	Gdzie obiekt się znajduje	Miejsce, pole widzenia kamery
Klasyfikacja obiektu	Jakiego rodzaju jest to obiekt	Człowiek, pojazd (samochód osobowy, autobus, samochód ciężarowy, rower/motocykl)
Atrybuty obiektu	Jakie cechy ma obiekt	Odzież, dodatki takie jak nakrycie głowy lub torba, cechy fizyczne takie jak kolor odzieży

Nawet jeśli dostępna jest tylko jedna kategoria metadanych, na przykład czas, może się ona okazać kluczowa dla znalezienia potrzebnych wyników.

Metadane dotyczące ruchu umożliwiają wyszukiwanie obiektów na podstawie ich prędkości i kierunku przemieszczania się. Dane dotyczące klasyfikacji obiektów umożliwiają wyszukiwanie uwzględniające szerszy zakres szczegółów. Kamery wyposażone w procesory głębokiego uczenia zazwyczaj dostarczają wzbogacone metadane z bardziej szczegółową klasyfikacją obiektów, umożliwiając na przykład wyszukanie zielonej ciężarówki lub osoby w niebieskim płaszczu.

### 4.3 Identyfikowanie trendów i wzorców w celu pozyskiwania wiedzy

Platformy z zakresu IoT i efektywności operacyjnej na potrzeby raportowania statystycznego mogą wykorzystywać metadane do celów zliczania gości, mierzenia prędkości, gromadzenia danych dotyczących ruchu i innych rodzajów automatycznego zbierania danych. Dane są analizowane w celu generowania wartościowej wiedzy praktycznej.

## 5 Gdzie wykorzystuje się metadane?

Istnieje wiele korzyści wynikających z zastosowania metadanych w celu poznania cech i zawartości sceny. Głównych odbiorców metadanych można podzielić na poniższe kategorie.

**Aplikacje brzegowe.** Działające w kamerze narzędzia analityczne mogą stosować filtry logiczne i reguły do informacji na temat określonego obiektu obecnego w scenie. Dzięki temu aplikacja analityczna może wyzwać działania na podstawie zdefiniowanych wartości progowych lub określonych zachowań, np. sterować kamerą PTZ na podstawie detekcji i ruchu osoby występującej w scenie.

**Systemy zarządzania materiałem wizyjnym (VMS).** W kontekście dozoru wizyjnego metadanych często używa się w systemach VMS do wyświetlania nakładek wizualnych wokół potencjalnych interesujących obiektów, które są obecnie w scenie. Dzięki rozwojowi bardziej zaawansowanych algorytmów detekcji i klasyfikacji obiektów operator może obecnie lokalizować interesujące obiekty na podstawie określonych cech, np. koloru odzieży. Możliwość wyszukiwania z wykorzystaniem tych punktów danych znacznie ogranicza potrzebę ręcznego przeglądania ogromnej ilości materiału.

**Platformy IoT.** Metadane mogą być zbierane i prezentowane na platformach business intelligence, zapewniając wartościową wiedzę praktyczną dzięki analizie aktualnych i historycznych trendów. Analizy statystyczne oparte na przepływie klientów lub ich wrażeniach umożliwiają podejmowanie decyzji opartych na danych w celu usprawniania działalności operacyjnej.

**Druga warstwa analiz.** W niektórych aplikacjach do wykonywania bardziej zaawansowanych analiz wymagane jest połączenie przetwarzania brzegowego i serwerowego. Przetwarzanie wstępne wykonuje kamera, a za dalsze czynności odpowiada serwer. Taki hybrydowy system umożliwi ekonomiczne skalowanie analiz, ponieważ na serwer przesyłany jest tylko niezbędny materiał wizyjny i niezbędne metadane.

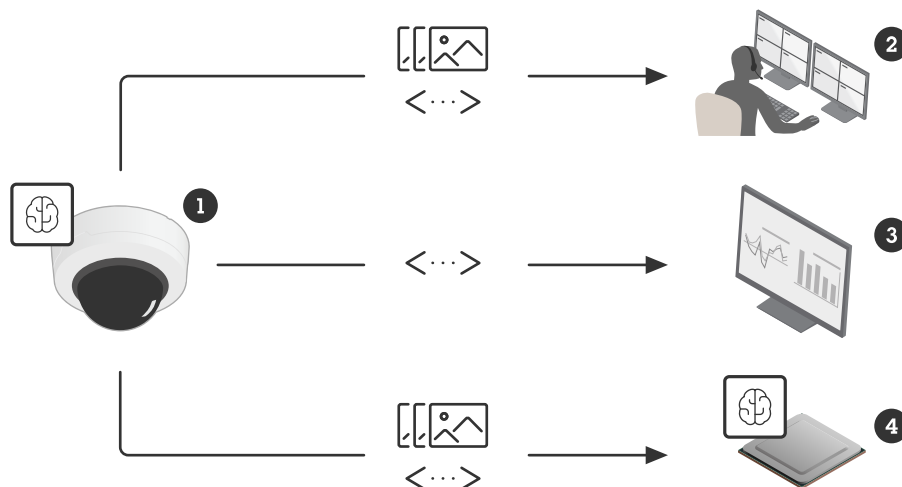


Figure 2. Odbiorcy metadanych

- 1 Aplikacje brzegowe
- 2 VMS
- 3 Platformy IoT
- 4 Druga warstwa analiz

## 6 Jak są dostarczane metadane?

Wygenerowane metadane mogą być dostarczane w różny sposób zależnie od zamierzonych zastosowań. W przypadku aplikacji działających w czasie rzeczywistym metadane muszą być stale przesyłane do odbiorcy na żądanie, ponieważ jest to niezbędne do zapewnienia odpowiedniej reakcji i świadomości sytuacyjnej. W przypadku mniej istotnych aplikacji, które nie wymagają działania w czasie rzeczywistym, przed dostarczeniem metadanych do odbiorcy możliwa jest ich dalsza konsolidacja, na przykład na

podstawie ścieżki każdego z obiektów w scenie. Pozwala to zmniejszyć ogólną ilość danych wymagających przechowywania i przetwarzania.

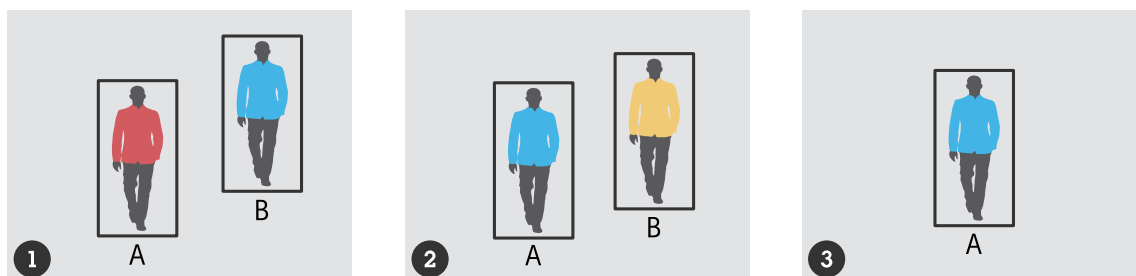


Figure 3. Ta ilustracja przedstawia strumieniowe przesyłanie metadanych, gdzie ciągłe klatki metadanych przekazywane z kamery dostarczają informacje o scenie w czasie rzeczywistym. Każda klatka zawiera obraz sceny z określonej chwili, niezależnie od wcześniejszych zdarzeń.

- 1 W klatce 1 następuje wykrycie obiektów A i B oraz sklasyfikowanie obiektu A jako człowieka w czerwonym ubraniu, a obiektu B jako człowieka w niebieskim ubraniu.
- 2 W klatce 2 kamera aktualizuje klasyfikację, ustalając, że obiekt A w rzeczywistości ma na sobie niebieskie ubranie, a obiekt B – ubranie żółte. Choć obiekty są takie same jak w klatce 1, zmieniają się atrybuty ich koloru, co znajduje odzwierciedlenie w metadanych.
- 3 W klatce 3 nie ma już obiektu B, a kamera śledzi tylko obiekt A, nadal sklasyfikowany jako człowiek w niebieskim ubraniu.

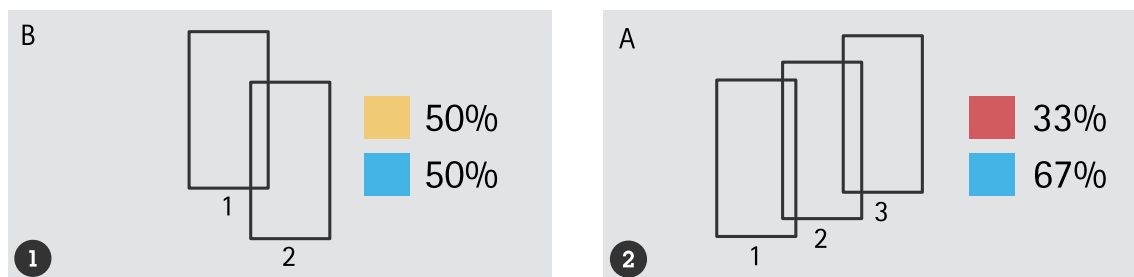


Figure 4. Ta ilustracja przedstawia dostarczanie skonsolidowanych metadanych, gdzie kamera przekazuje informacje w ujednoliconym formacie na podstawie wykrytej ścieżki obiektów w scenie. Klatki dotyczące każdego obiektu zawierają wszystkie znane szczegóły z całego okresu istnienia ścieżki obiektu.

- 1 W pierwszej klatce prezentowane są szczegóły dotyczące obiektu B, które obejmują pierwszą i ostatnią detekcję, podsumowanie trajektorii oraz atrybuty wykryte podczas ścieżki. Według systemu istnieje 50% prawdopodobieństwa, że obiekt B ma na sobie żółte ubranie, oraz 50% prawdopodobieństwa, że ma ubranie niebieskie.
- 2 W drugiej klatce ten sam format został zastosowany do obiektu A, w przypadku którego system wskazał 33% prawdopodobieństwa, że obiekt ma na sobie czerwone ubranie, oraz 67% prawdopodobieństwa, że ma ubranie niebieskie.



Zaletą metody skonsolidowanej polega na tym, że kamera znacznie zmniejsza ilość danych wysyłanych do odbiorcy, ponieważ dostarcza metadane tylko wtedy, gdy scena zawiera obiekty, a w takim przypadku podsumowanie (konsolidacja) metadanych ułatwia interpretację. W przypadku metody strumieniowej w każdej klatce przesyłany jest pełny opis sceny, nawet jeśli nie występuje w niej żadna aktywność ani obiekty, a odbiorca musi interpretować otrzymywane dane odpowiednio do swoich potrzeb. Jak już wspomniano, metoda strumieniowa przydaje się w zastosowaniach opartych na czasie rzeczywistym, natomiast metoda skonsolidowana jest optymalna w przypadku późniejszego przetwarzania, gdy użytkownik nie musi podejmować natychmiastowych działań.

Zrozumienie atutów i ograniczeń każdej metody jest niezbędnym warunkiem projektowania architektury systemu. Jeśli przykładowo klient korzysta z platformy IoT generującej spostrzeżenia na podstawie metadanych, dobrym rozwiązaniem będzie przesyłanie podsumowania na temat obiektów w scenie po wystąpieniu zdarzenia, ponieważ tego typu platformy zazwyczaj dysponują ograniczoną przepustowością i pamięcią masową.

Ponadto metadane mogą być dostarczane przy użyciu różnych protokołów komunikacyjnych i formatów plików odpowiednio do specyficznych potrzeb i preferencji docelowego użytkownika.

## **7 Metadane dotyczące dźwięku**

Niektóre kamery potrafią dekodować dźwięk i pozyskiwać metadane na jego temat. Aplikacja analityczna do rozpoznawania audio może wykrywać wzorce dźwiękowe oraz wyróżniać interesujące odgłosy w przekazywanym na żywo i nagrany materiał audio. W ten sposób system rozpoznawania audio połączony z urządzeniami do dozoru może na bieżąco ostrzegać operatora o potencjalnych incydentach, wskazując widok z odpowiedniej kamery. System może na przykład identyfikować agresję werbalną, aby zapobiegać eskalacji i napaściom fizycznym, wykrywać tłuczenie szkła, aby zapobiegać włamaniom, oraz wysyłać wczesne ostrzeżenia o pacjentach potrzebujących pomocy. Pozwalając operatorowi nie tylko widzieć, ale i słyszeć, co dzieje się w monitorowanej scenie, system rozpoznawania dźwięków umożliwi wczesną detekcję i szybką interwencję, a w wielu przypadkach pomaga też zapobiegać eskalacji zdarzeń. System rozpoznawania dźwięków może również służyć jako dodatkowy sposób weryfikacji.

Narzędzia analityczne przeszkolone do rozpoznawania wzorców dźwiękowych zazwyczaj wychwytyują określone połączenia cech, od stopnia natężenia dźwięku po poziom energii na różnych częstotliwościach w funkcji czasu. Określone wzorce dźwiękowe mogą być wykrywane i oznaczane w podobny sposób jak klasy obiektów w materiale wizyjnym.

## **8 Łączenie metadanych z wielu źródeł**

Prawdziwy potencjał metadanych ujawnia się w przypadku ich zastosowania do wielu rodzajów danych wejściowych, takich jak wizyjne, dźwiękowe oraz związane z aktywnością i procesami. W zarządzaniu każdym obiektem cenną pomocą są tagi RFID, współrzędne GPS, alarmy sabotażowe, odczyty liczników (np. poziomy temperatury lub substancji chemicznych), dane z systemu detekcji hałasu czy dane transakcyjne z punktów sprzedaży. Dane pochodzące ze wszystkich tych źródeł można harmonizować na podstawie znaczników czasowych.

Połączenie metadanych z różnych źródeł pozwala pozyskać więcej wiedzy, niż kiedykolwiek byłoby to możliwe z każdego źródła osobno. Do płynnej integracji metadanych niezbędne są otwarte protokoły i standardy.

# O firmie Axis Communications

Axis umożliwia tworzenie mądrzejszego i bezpieczniejszego świata, tworząc rozwiązania zwiększające bezpieczeństwo i wydajność biznesową. Jako firma z branży technologicznej będąca liderem na rynku, Axis oferuje systemy dozoru wizyjnego, kontroli dostępu, domofonowe i rozwiązania audio. Rozwiązania te są wzbogacone o inteligentne aplikacje analityczne i wysokiej jakości szkolenia

Firma Axis zatrudnia około 4000 zaangażowanych pracowników w ponad 50 krajach i współpracuje z partnerami z sektora technologii oraz integracji systemów na całym świecie, aby dostarczać rozwiązania dla klientów. Firma Axis powstała w 1984 roku, a jej siedziba znajduje się w Lund w Szwecji