

ホワイトペーパー

AXIS Camera Station Proの フリーテキスト検索

1月 2025

概要

AXIS Camera Station Proには、いくつかのビデオフォレンジック検索ツールがプリインストールされています。これには、タイムラインスクラビング機能、データ検索、事前に分類されている物体に基づくスマート検索、フリーテキスト検索が含まれます。

フリーテキスト検索を活用することで、独自の単語を用いて検索用語を入力し、移動物体を検索することができます。広範な記述属性を使用して詳細な検索フィルタを自由に作成できるため、関連性の高い映像をより迅速に見つけることが可能となります。

フリーテキスト検索機能は、事前にトレーニングされたオープンソースの基盤モデルによるテキストと画像のマッチングに基づいています。この基盤モデルは、監視用途向けにAxisで最適化されています。検索は1台のカメラに適用することも、複数のカメラに同時に適用することもできます。

フリーテキストクエリの数値表現と検知された移動物体の画像の数値表現の比較が行われます。結果は、カメラ名、時間、日付が含まれているサムネイルとして表示され、検索クエリとの関連性の高さに基づいて並べ替えられます。

AIを活用するフリーテキスト検索により、フォレンジック検索ソリューションの精度と効率が高まります。これが、最終的には人の意思決定の強化につながります。法的基準と倫理基準に準拠することを目的として、検索機能にはAxisが開発した別のモデレーション機能が含まれています。この機能により、検索クエリにおける不適切な単語や表現の使用が抑制されます。すべての検索はログに記録されます。管理者はそのログを表示できるため、不正な使用が発生した場合には是正措置を講じることが可能となります。

目次

1	はじめに	4
2	背景：AXIS Camera Station Proのフォレンジック検索	4
3	フリーテキスト検索の機能の仕組み	4
4	検索クエリの作成	5
	4.1 プロンプトのガイドライン	6
	4.2 クエリのモデレーション	6
	4.3 検索クエリのログ記録	6
5	テキストと画像のマッチング	7
	5.1 監視向けに最適化された基盤モデル	7
	5.2 特徴ベクトル	7
6	プロセスの概説	7
7	責任あるAIの使用	9

1 はじめに

AXIS Camera Station Proのフリーテキスト検索ツールを活用することで、既定のフィルタではなく、独自の単語を使用して録画を検索できるようになります。

本ホワイトペーパーには、検索方法の仕組みの概要、およびその使用方法に関するガイドラインが記載されています。また、法的基準と倫理基準への確実な準拠を目的として導入されているモデレーション機能とクエリログに関する説明も含まれています。

2 背景：AXIS Camera Station Proのフォレンジック検索

AXIS Camera Station Proには、タイムラインスクラビング機能、データ検索、事前に分類されている物体に基づくスマート検索、フリーテキスト検索など、いくつかのビデオフォレンジック検索ツールがプリインストールされています。

スマート検索機能では、Axisデバイスで生成されたシーンメタデータが使用されます。メタデータには、移動物体のタイプ（人物、車両タイプ、不明な物体）と共に、衣服や車両の色、ナンバープレート、速度、場所、タイムスタンプといった属性が含まれます。

分析機能に限界のあるデバイスの場合は、デバイスの動体検知とAXIS Camera Station Proサーバーで実行される物体分類の組み合わせに基づいて検索が実行されます。AXIS Camera Station Proのフォレンジック検索は、エッジデバイスの機能を可能な限り活かしながら、必要に応じてサーバーからのデータで補完するハイブリッドソリューションとなります。

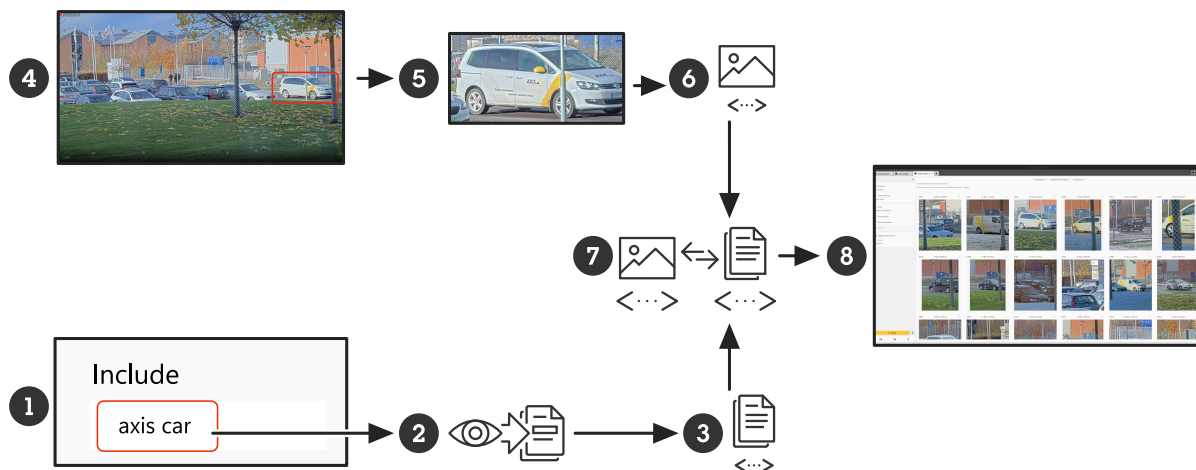
従来的には、シーンメタデータを使用する検索の場合は、既定の検索フィルタを使用する必要がありました。この場合は、物体のタイプ（例：「車両」）、車両タイプ（例：該当する場合は「車」）、色（例：「青」）など、リストから固定物体記述子を選択します。これとは異なり、新たなフリーテキスト検索方法では、独自の検索フィルタを作成することができます。

事前に分類されたデータに基づく検索の場合は、高精度の結果が得られるものの、既定されていない新たな物体のタイプを検出することができません。フリーテキスト検索では、独自の単語を使用して自由かつ柔軟に検索を実行できるため、こうした制限を克服することができます。自然言語と関連付けを用いて、移動物体をより詳細に説明し、検索を微調整できることから、検索の成果が高まります。

3 フリーテキスト検索の機能の仕組み

フリーテキストクエリの数値表現と検知された移動物体の画像の数値表現の比較が行われます。テキストと画像が一致した結果が、最も一致度の高いものから順に並べ替えられて

表示されます。結果は、カメラ名、時間、日付が含まれているサムネイルとして表示され、検索クエリとの関連性の高さに基づいて並べ替えられます。



フリーテキスト検索プロセスの簡易概要。手順4～6は、検索時以外も、検知されたすべての移動物体の特徴ベクトルを作成するために継続的に実行されるプロセスです。

- 1 ユーザーがフリーテキスト検索クエリを入力する。
- 2 モデレーションモジュールにより、有害で非倫理的な言葉の使用が防止される。
- 3 基盤モデルにより、検索クエリの数値表現（特徴ベクトル）が作成される。
- 4 カメラにより、シーン内の動きが検知される。
- 5 カメラにより、移動物体を表すトリミング画像（切り取られた画像）1つが選択される。
- 6 基盤モデルにより、物体の形状、パターン、色などが分析された後に、物体の特徴ベクトルが作成される。
- 7 2つの特徴ベクトルの比較が行われる。
- 8 比較の結果が、最も一致度の高いものから順に並べ替えられて、サムネイルとして表示される。

フリーテキスト検索は1台のカメラに適用することも、複数のカメラに同時に適用することもできます。

類似検索や時間ベースの検索など、他のスマート検索機能と組み合わせ、複数の検索の種類を使用することで、フリーテキスト検索の範囲を絞り込むことができます。

4 検索クエリの作成

あらゆる移動物体やあらゆる種類の車両を検索することができます。最良の結果が得られるように、ガイドラインに従ってください。

検索する対象は、移動物体のみとなることに注意してください。多くの場合、静止物体の検索は良好に機能しません。

非倫理的な検索行為の防止を目的として、検索の文章はモデレーションの対象となり、記録されます。

4.1 プロンプトのガイドライン

- 画像を描写するように状況を記述します。モデルには静止画像が取り込まれるため、動作（転倒、走行、窃盗など）を検索することは困難になる可能性があります。これは、動作にはより多くのコンテキストが必要になるためです。
- 「a person in a red sweater（赤いセーターを着た人）」や「a yellow pickup truck（黄色のピックアップトラック）」など、いくつかの重要な記述子を使用して物体を説明します。他のマルチモーダルモデルと同様に、フリーテキスト検索モデルの場合も、物体や色といった記述子を使えば良好に機能しますが、数（例：「three persons（3人）」）、俗語、感情的な記述（例：「angry-looking man（怒っているように見える男性）」）は適切ではありません。物体の説明に、主観的な記述や曖昧すぎる言葉を使用しないように、また具体的な詳細を多く含めすぎないように注意する必要があります。
- 「and（および）」を使用して、複数の物体の属性を組み合わせます。例として、「person with red hat and backpack（赤い帽子とリュックを持った人物）」といった文章が挙げられます。
- テキスト、テキストのロゴ、ブランド名を記述します。例として、「van with text Axis（Axisという文字が付いているバン）」といった文章が挙げられます。
- 環境の説明に焦点を当てないでください。物体のトリミング画像（切り取られた画像）を用いて処理が行われるため、モデルで物体の周囲や背景が認識されない可能性があります。そのため、広範なシーンや環境を説明する記述子（例：「city（都市）」、「urban（都会）」、「park（公園）」、「garden（庭）」、「lake（湖）」、「beach（ビーチ）」）では、適切な結果が得られない可能性があります。
- 結果が思わしくない場合は、別の表現や言い回しを試してください。
- フリーテキストプロンプトは英語でのみサポートされています。

4.2 クエリのモデレーション

効果的なフィルタリングが実現するように、検索機能には一般的な慣行に基づくクエリモデレーションが実装されています。モデレーションモデルは、自然言語処理モデルです。これにより、クエリがチェックされ、不適切な単語や表現が抑制されます。テキスト文字列全体のチェックが行われ、有害、不適切、有毒なコンテンツが存在しないかが確認されます。さらに、禁止検索カテゴリや単語のカスタムリストなど、当社独自の対策によってこの機能が強化されています。このリストに含まれている単語や句がクエリに存在すると検索が拒否されるため、安全な検索環境が維持され、倫理的な結果が確保されます。ユーザーは匿名でフィードバックをAxisに提供し、特定の単語がブロックされていることに異議を唱えること、またブロックすべき単語を提案することができます。

4.3 検索クエリのログ記録

AXIS Camera Station Proでは、ユーザー操作の監査証跡が保持されます。監査証跡では、特定の操作とユーザーIDが記録されるだけでなく、操作で使用されたデータも保持されます。つまり、検索プロンプトを含め、すべてのユーザー検索が記録されるということです。管理者はログを使用することで、ユーザーによる不適切な検索操作を特定し、非倫理的な検索プロンプトにフラグを立て、そして是正措置を講じることができます。

ビデオデータがAxisと共有されることはありません。データはサーバーに残ります。

5 テキストと画像のマッチング

フリーテキストクエリを使用すればビデオメタデータを検索できることから、検索機能を著しく拡張することが可能となります。つまり、これまで既定の属性リストが基準となっていた検索をほぼ無制限の検索基準まで広げることができるということです。AXIS Camera Station Proに搭載されているこの機能は、オープンソースの基盤モデルに基づいて構築されており、数十億の画像とテキストのペアでトレーニングされているだけでなく、監視用途向けにAxisで微調整されていることから、より優れた性能を発揮します。

5.1 監視向けに最適化された基盤モデル

基盤モデルは、テキストと画像のペアが含まれている大規模なデータセットを用いてトレーニングされたText-to-Image（テキストから画像を生成する）モデルです。これは、テキストと関連性の高い画像を一致させるゼロショットモデルです。ゼロショットモデルとは、特定の事前トレーニングデータがなくても、物体や概念を認識して分類できるAI（人工知能）モデルの一種です。言い換えれば、これは未見のタスクの場合も、そのタスクを実行できる能力を備えたモデルです。自然言語と画像を一致させる際に最適な性能を実現する上で、この能力は不可欠となります。

このモデルは大量のテキストと画像の組み合わせでトレーニングされており、25億超のパラメーターが投入されたニューラルネットワークで動作します。Axisは一般的な監視カメラにおける表示と物体の画像の解釈能力を向上させることを目的として、独自のトレーニング材料を使用してこのモデルをさらに調整しています。つまり、このモデルは監視用途向けに最適化されているということです。

5.2 特徴ベクトル

フリーテキスト検索を行う際に、基盤モデルによって検索クエリの特徴ベクトルが作成されます。

また、基盤モデルでは、シーンで追跡の対象となったすべての物体の説明が継続的に生成され、それぞれを表す特徴ベクトルが作成されます。各物体は1つの特徴ベクトルのみで表されます。これが、データベースに保存されます。特徴ベクトルがあらかじめ計算されており、直ちにデータベースで利用できるようになっていることから、高速に検索を行うことができます。

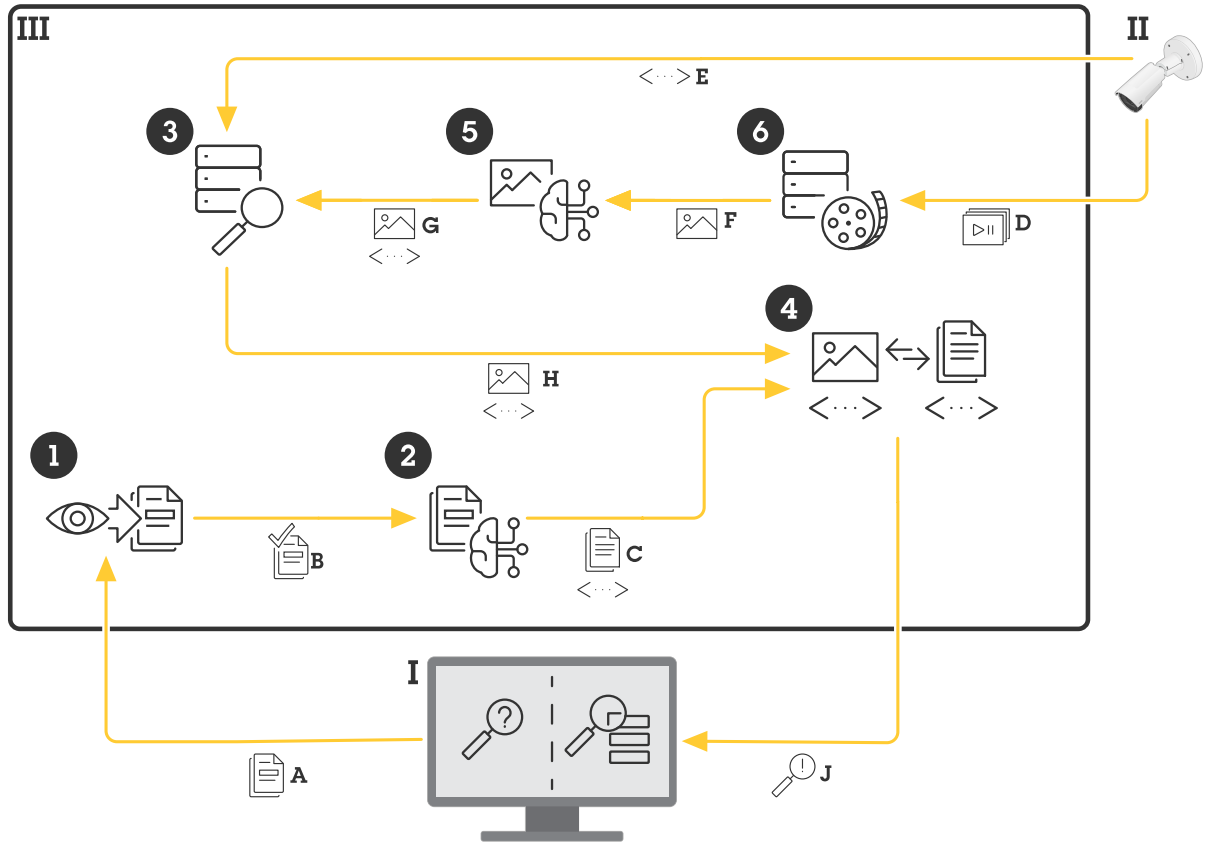
両方のタイプの特徴ベクトルがベクトル比較エンジンに入力され、検知された物体を表すすべての特徴ベクトルと検索クエリとの間の類似距離が決定されます。

特徴ベクトルとは、テキストや画像を数値的に表現したベクトルのことです。そのため、人物や物体の特徴ベクトルは、人物または物体の外観の抽象的な表現にすぎません。頭髪や衣服の色というものは、特定の人物に明示的にマッピングされる特徴またはその識別に使用できる特徴ですが、特徴ベクトルには人間が解釈できる情報は含まれていません。特徴ベクトルは、他の特徴ベクトルとの比較にのみ使用できる要素です。

6 プロセスの概説

プロセスの概説図に、各段階が実行される場所および各段階で生成されるデータの種類など、主要プロセスの段階が示されています。

カメラ (II)、録画ストレージ (6)、基盤モデル (5)、検索データベース (3) などが含まれている図上部の枠内には、検索時のみでなく、検知されたすべての移動物体の特徴ベクトルを作成するために継続的に実行されるプロセスが示されています。



フリーテキスト検索プロセスのメインの場所 (I-III)

- I AXIS Camera Station Proクライアント
- II カメラ
- III AXIS Camera Station Proサーバー

メインのプロセス段階 (1~6)

- 1 検索クエリのモデレーション
- 2 基盤モデル (テキスト)
- 3 検索データベース
- 4 ベクトルの比較
- 5 基盤モデル (画像)
- 6 録画ストレージ

データのタイプまたは結果 (A-J)

- A 文字列
- B 文字列
- C 特徴ベクトル (テキスト)
- D ビデオ
- E メタデータ
- F 画像
- G 特徴ベクトル (画像)
- H 特徴ベクトル (画像)

J 検索結果

(I) **AXIS Camera Station Proクライアント**：ここで、ユーザーが検索クエリを入力し、並べ替えられた検索結果を表示します。

(II) **カメラ**：フリーテキスト検索は、AXIS OS 5.51以降を搭載したAxisカメラで実行することができますが、カメラの性能が高いほど、取得できる結果の質も向上します。古いデバイスの場合は、動体検知のみに基づいた粗粒度のメタデータしか得られません。より新しいデバイスでは、AXIS Scene Metadataを活用することができます。これには、物体分類が含まれます。カメラの移動物体検知と追跡機能を用いて、検知された各物体の代表的な画像が1つ特定されます。これにより、サーバーで分析しなければならない画像の数が削減されます。

(III) **AXIS Camera Station Proサーバー**：ここで、カメラで生成されたすべてのメタデータとビデオデータの処理と保存が行われます。フリーテキスト検索を実行できるようになる前に、まずサーバーで（検知された移動物体ごとに）ビデオがデコードされ、検知された物体の画像が抽出されます。デコードと抽出が完了すると、次に基盤モデルでその画像が処理されて特徴ベクトルが作成されます。処理能力という観点から、こうした操作には非常にコストがかかるため、性能を向上させるために、特徴ベクトルはデータベースに保存されます。これにより、今後の検索を迅速に行うことができますようになります。サーバーに余裕がある場合は、最も重要なカメラで生成されたビデオのバックグラウンド処理を有効化することが強く勧められます。これにより、検索が大幅に高速化されます。

(1) **検索クエリのもデレーション**：モデレーションモデルにより、クエリがチェックされて、不適切なコンテンツが制限されます。

(2) **基盤モデル（テキスト）**：基盤モデルにより、モデレートされた検索クエリテキスト文字列の数値表現（特徴ベクトル）が作成されます。

(3) **検索データベース**：属性、時間、位置、特徴ベクトルによる物体分類データなど、AXIS Scene Metadataまたはサーバーで生成された完全なメタデータが検索データベースに保存されます。

(4) **ベクトルの比較**：検索クエリテキスト文字列の特徴ベクトル表現とビデオで検知された物体画像の特徴ベクトル表現の比較が行われます。

(5) **基盤モデル（画像）**：基盤モデルにより、追跡されてビデオに録画された各物体の数値表現（特徴ベクトル）が作成されます。これは、検索時以外でも継続的に実行されるプロセスです。

(6) **録画ストレージ**：カメラで生成されたビデオがこのストレージに保存されます。基盤モデルにより、ここから画像が取得されます。

7 責任あるAIの使用

AIを活用するフリーテキスト検索により、フォレンジック検索ソリューションの精度と効率が高まります。これが、最終的には人の意思決定の強化につながります。

AIに関するAxisのアプローチでは、責任と説明責任がその基盤となっています。これには、倫理原則を遵守したAIシステム、法律に準拠したAIシステム、そしてリスクを効果的に管理できるAIシステムを構築するということが含まれます。Axisは、顧客が安心して業務の完全性を維持できるツールを提供しています。AXIS Camera Station Proのフリーテキスト検索機能には、適切なテキストプロンプトモデレーションが実現するように微調整されたテキスト分類モデルが含まれています。検索クエリをモデレートすることで、不正使用を防止できるように、当社はこのモデルを開発しました。

新たな規制や要件に継続的に準拠するためには、AIモデルを更新していく必要があります。その更新の必要性の有無を確認するため、週に1回フリーテキスト検索機能からAxisクラウドサービスへの接続が行われます。この接続に失敗すると、接続が再確立されるまでフリーテキスト検索を利用することができなくなります。

AIの応用において、関連する法的基準と倫理基準をより確実に遵守するため、当社の製品には、ユーザー認証資格情報とアクセス権限に基づくアクセスコントロールが備わっています。これにより、当社の顧客はユーザーの運用ポリシーへの遵守を強制的に確保することができます。

Axis Communicationsについて

Axisは、セキュリティとビジネスパフォーマンスを向上させるソリューションを生み出すことで、よりスマートでより安全な世界の実現を目指しています。ネットワークテクノロジー企業として、また業界をけん引するリーダーとして、Axisは映像監視、アクセスコントロール、インターコム、音声システムなどに関連するソリューションを提供しています。これらのソリューションは、インテリジェントアプリケーションによって強化され、質の高いトレーニングによってサポートされています。

Axisは50ヶ国以上に4,000人を超える熱意にあふれた従業員を擁し、世界中のテクノロジーパートナーやシステムインテグレーションパートナーと連携することで、カスタマーソリューションをお届けしています。Axisは1984年に創業し、本社はスウェーデン・ルンドにあります。