

ひょうごTECHイノベーションプロジェクト

SNS上に関連した犯罪から子どもたちを守りたい！
AIを活用したサイバーパトロール

最終報告報告

株式会社Spectee（スペクティィ）



実証の背景・目的と解決する課題

実証名：SNS上に関連した犯罪から子どもたちを守りたい！AIを活用したサイバーパトロール



実証の背景

SNS上に書き込まれた違法・有害情報が多くあり、サイバーパトロールを行っている



実証の目的

AI等を活用したシステムを活用することでより効率的なサイバーパトロールを行えるか検証する

想定する実証実験

AIを利用した検索ツールで、対象となる書き込みを効率よく発見できるかどうかを検証する



今後の展開等

実証の結果、有用と結果が出るようであれば本導入を進めていきたい



本実証実験の概要・課題の解決方法

キーワードによる検索、解析を繰り返し、県警本部として必要な情報が取得できるかどうかを実証実験を行います。

検索（キーワードによる抽出）

キーワード設定を行い情報を抽出



結果に対する情報の確認

結果に対する解析を行い
情報が取得できているか確認する。



除外キーワードを設定し再度検索

再度キーワード設定を行い
情報を検出



※上記サイクルを繰り返し、最終的に本実証で取り組みたい情報を取得できるかどうかを検証します。

モックアップ画面

利用者側で抽出された投稿についてメモや情報の必要性、また取得データの出力が行える画面のモックアップも作成します。

The screenshot shows a web application interface for Spectee. The main content area displays a list of posts. Each post has a header with a status icon (circle, triangle, or cross) and a text area. The posts are filtered based on these status icons. The interface includes a navigation bar at the top with the Spectee logo and a menu icon. On the right side, there are several buttons for navigation and filtering, such as '記事カテゴリ' (Article Category) with options like 'HOME', '国内', '海外', '話題', '感染症', 'お気に入り', '地図検索', 'リスク情報', and 'コロナ情報'. Below these are buttons for '台風情報' (Typhoon Information) and 'レポート' (Report). A '絞り込み' (Filter) section on the right allows users to set the number of columns (自動), classification (radio buttons for '関係がありそう', 'おそらく関係がありそう', '不必要そう'), keywords, and date range. A 'ダウンロード' (Download) button is available for the filtered results. The interface also includes a search bar and a 'ログアウト' (Logout) button.

投稿について○（必要）、△（要検討）、×（不要）と入力できます。

解析の結果、該当すると思われる投稿を表示します。

場所欄に位置情報や備考欄にメモを入力することができます

絞り込み
列数 自動
分類 ○(関係がありそう) △(おそらく関係がありそう) ×(不必要そう)
キーワード 絞り込むキーワードを入力してください
期間 from to
ファイル この条件でエクセルファイルをダウンロード

場所欄や備考欄のキーワード検索、返信内容の記号ごとの絞り込み、期間で絞り込み表示をしたり、結果をエクセルファイルでダウンロードできます。

実証の結果

本実証期間に2回のキーワード解析を行い、結果について検証した

1回目概要

- ・ 県警よりキーワードをいただく
- ※およそ90件のキーワード
- ・ 土曜日、日曜日の2日間のデータを抽出

2回目概要

- ・ 前回よりもキーワードを減らして実施
- ※普段県警で行っている内容で実施
- ・ 土曜日、日曜日の2日間のデータを抽出

- ① 毎回非常にノイズとなる投稿が多い※毎回数万件の投稿が収集されるが実際に必要と思われる投稿は数十件
- ② 普段利用しているキーワードのほうが精度が高い情報が出やすい
- ③ 継続して社内でデータ収集していたが、週末の投稿件数が非常に多い傾向がある

課題と精度向上、今後の展開に向けて

課題

- ・精度検証が十分でない
- ・様々な投稿のデータが必要
- ・精度向上のためのデータ量が足りない



今後に向けて

- ・必要なデータを自動的に取得し、表示をする。
- ・データ化して可視化させる
- ・他県警による注意喚起等を行っている投稿を表示・非表示機能の追加
- ・新しい隠語等自動対応できるようにする

他領域へ展開する場合の課題

SNS解析における情報収集料金（API）の費用、サーバー等の維持管理費が非常に高額化している

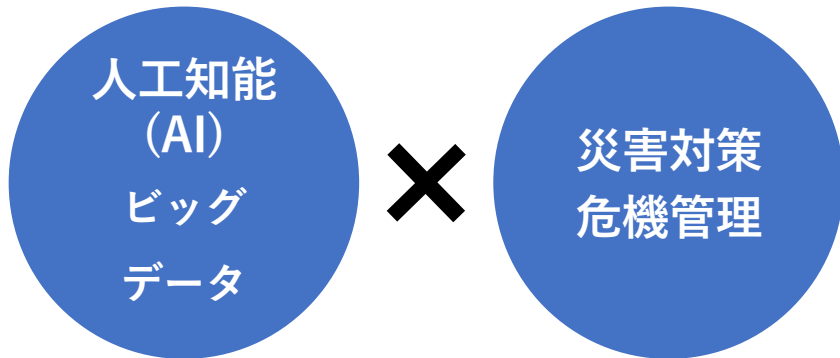
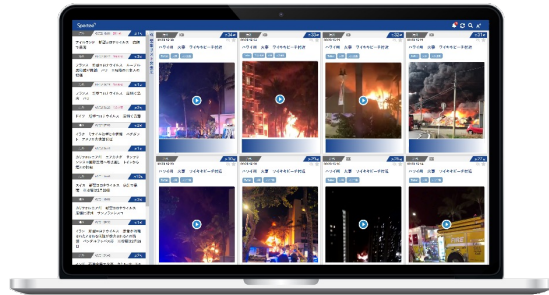
世界的な値上げによりサービス提供のサーバーやAI学習にかかる費用も高額化しているため、ある程度の顧客が見込める内容でないと非常に展開が難しい。

※展開ができないわけではないが、1ユーザーのみだと費用が高額化していくため、費用対効果が微妙になる。



参考資料

Specteeの紹介とメディア掲載について

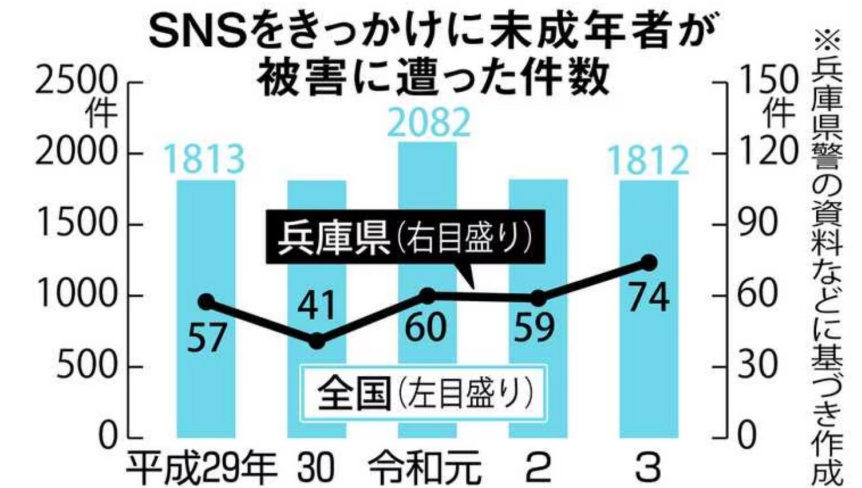


<独自>未成年者の性被害抑止にAIシステム導入へ 兵庫県警、SNSを自動検索

2023/1/27 12:04 倉持 亮

みんなの反応

社会 | 事件・疑惑 地方 | 近畿 兵庫
産経WEST | できごと



未成年者が性犯罪に巻き込まれるのを防ごうと、兵庫県警が、人工知能（AI）を活用して交流サイト（SNS）上の問題投稿を自動検出する新システムの導入を検討していることが27日、県警への取材で分かった。手作業で行っている現状のサイバーパトロールを効率化し、捜査や発信者への警告に人的資源を集中することで被害の抑止を目指す。

※2023年1月27日産経新聞より