BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

Using AI in Real-Time Social Media Analytics: How OSINT Can Be Used to Detect Drug Trafficking Activities

Rishita Choudhary, Prof. Geeta Kodabagi

Department of Artificial Intelligence and Data Science Engineering
Ajeenkya DY Patil School of Engineering, Pune, Maharashtra, India
Academic Year 2026–27

Abstract

More and more, drug trafficking groups have been communicating via public social media sites under layers of coded language, hashtags, emojis, and images. Monitoring such a huge variety of online content manually is a slow and laborious task, especially when phrases and patterns change. This research presents Artificial Intelligence (AI) and Open-Source Intelligence (OSINT) as means of helping the examination of published material from social media in an effort to detect content potentially needing further inspection. The methods covered by the study are the application of Natural Language Processing (NLP) for analysing written language and evaluating what a person is feeling or planning, the use of Computer Vision for the exploration of pictures that are publicly available, and machine-learning for the labelling of a suspicious content or a piece of material. A conceptual AI–OSINT workflow diagram is illustrated, starting with data acquisition, going through preprocessing, then multimodal analysis, followed by risk scoring, ending with human reviewing. According to the research papers and seminar content, the challenge of fusing both image and text analyses, responding to the development of secret language, having large enough databases, and producing clear explanations are still major issues. As the internet material can sometimes be incomplete and an automatic alarm system cannot be used to confirm a crime has been committed, AI-supported OSINT should be seen as a decision aid that is backed by privacy safeguards, a fair assessment, and the human double-check.

Keywords : Artificial Intelligence, Open-Source Intelligence, OSINT drug trafficking recognition, analysis of social networks, NLP, computer graphics, information security

A. Motivation

Publically shared information on social media may serve purposes in identifying and solving cases of cybercrime and other organized criminal acts. However, considering the huge volume of posts, informal speech or secret code used, and the presence of images all of these make the manual process of going through such a collection of material rather challenging or not a realistic possibility if a short time is given. AI may be a solution that allows analysts not only to process the flow of public information smoothly but also to detect elements in it that, for instance, require being looked at more closely. Open-Access Intelligence (OSINT) is an essential tool that helps in collecting and interpreting information available to the general public. One of the main benefits of using OSINT is that it allows to very quickly

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

select and prioritise pieces of evidence for the investigators who are in the know; at the same time, the results depend on the accuracy of the data, the presence of the context, the reliability of the model of interpretation and, most importantly, the ethical use of information. The goal of this paper is to conduct research and present a conceptual workflow rather than to claim that a working system for detection purposes has been implemented.

B. Objectives

This paper aims to:

1. Give an insight into how an analyst can employ both AI and OSINT to look for patterns in data on social media which could indicate the existence of criminal or other undesirable activities.
2. Show the way of using NLP techniques for analysing text, hashtags, and secret codes.
3. Describe Computer Vision's function in the identification of images posted on social networks.
4. Browse the list of references given in the seminar slides and note the difficulties pointed out alongside the areas for future research.
5. Design an AI assisted analysis concept workflow with also the criteria that will be used as a basis for the evaluation and ethics protection measures that relate to AI assisted analysis.

Seminar presentation highlights an array of sources covering drug trafficking, human-centred social media analysis, open sources intelligence (OSINT), annotations at the level of human beings, textual context and author's intention, and AI risk management. Using those sources one can investigate AI assisted analysis potential and constraints. Still, the slide does not feature the authors full names, the datasets and the comparable numerical results for each listed reference. Therefore, the following discussion is based only on a qualitative point and doesn't offer any rankings which are not supported with evidence.

A. OSINT and Social-Media Intelligence

Open Source Intelligence depends heavily on publicly available material to be used for collecting information and conducting subsequent analysis. Europol's Internet Organised Crime Threat Assessment (IOCTA, 2024) is a source in the list of references for the seminar and is relevant to the topic of online organised crime. World Drug Reports by UNODC for 2024 and 2025 provide a broader understanding of drug markets. Social-network research, as in an Australian poly-drug trafficking network paper that has been mentioned, illustrates that relationships and patterns should be considered instead of looking at each post separately. Nevertheless, public sources can give us only a partial understanding of events because, on their own, they cannot reveal private or encrypted communications.

B. NLP, Computer Vision and Context Analysis

We discussed natural language processing, context analysis, and computer vision at the meeting. The presentation features a paper on developing a comprehensive platform integrating social media content analysis and context and intent recognition (2024), detecting drug trafficking on social forums via knowledge-informed ChatGPT (2024), and natural-language processing for social-media intelligence (Yang et al., 2021). These subjects focus on uncovering the signals of the context that extend beyond

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

individual words. The paper cited under Computer Vision (Alzahrani et al., 2023) deals with online video content analysis. Despite the fact that textual and visual content may together represent a context, standalone processing systems have limited capability for joint caption-image conversation comprehension as they are unable to trace their relationships through such different representations. Language, such as slang and emojis, multiple-language posts, and deliberately altered dictionary choices may significantly affect the accuracy.

C. Research Gap

Research issues identified in the seminar PPT are: (i) text and image analyses are not joined; (ii) the change in time of expressions like slang, emojis, and encrypted codes may affect the outcome of the

work; (iii) bigger and more diverse data collection is necessary; (iv) public social media posts usually have minimal explanations for what the authors meant and therefore human verification is required; and (v) explainable AI is essential so that an investigator can see reasons behind content selection. The reason for the paper's refusal to declare a method as quantitatively superior lies in the fact that the listed references don't contain a common benchmark or directly comparable performance measures for all mentioned different methods.

The proposed architecture is a high-level overview drawing on the seminar topic. It is an illustration of a possible way of conducting a project without any real implementation involved or the result of an experiment. The general idea is to go through the following steps: get the social media public content, do some preparation, make NLP and image work, conduct a contextual risk assessment, review it with a human. It is only legally obtained materials that should be considered at each step.

A. Data Collection

In the first step, the tool is looking for posts that are public from the approved sites only together with the metadata available on them. The way of collecting data is a combination of adhering to the platform terms, the local laws, and institutional policy. It must stay clear from breaching security settings, or trying to recover the private messages. Every item should contain only the information necessary for the purposes in question, such as text, time, a publicly accessible post or a file identifier, etc. In addition to the public data, the tool must capture and store the provenance, the source information, and the point in time at which a particular record was retrieved. It will allow the users to assess the authenticity and to discover the limitations if not possible to verify the facts in the given item or its source.

B. Preprocessing and Feature Extraction

The first stage involves the identification by means of the system of posts and of metadata which are publicly available and which are related, that is, they are all about the same event, situation, etc., from the approved sources. The collection must adhere to the terms of the platform, to applicable law, and to the institution's guidelines. The system is not allowed to bypass access control mechanisms or to attempt to access private messages. Only those fields which are essential for the analysis should appear in each record. Example: public text; the time; a public link or identifier; an image where

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

appropriate; and a publicly available image. Recording of data source, provenance and data collection time is an essential step, as only with that information the analysts can understand where the source comes from and the items' limitations.

C. Model Inference and Human Review

Through machine learning model or classifier training risk rating may be given to either a single public post or to a batch of public linked items. When the items are validated based on the scores exceeding the established threshold level, these should be prioritized for review, the machine model at the same time revealing the signals which have resulted in that risk score However a trained human will have to assess if there is some other explanation or a possible mistake in the context for investigative decision. If the automated scorer has found drug trafficking possibility that alone is not enough to carry out some punitive action.

Artificial Intelligence methodologies enable us to tackle challenging issues in which algorithms can be taught from previous experiences. The techniques include machine learning, the application of deep learning, and neural networks. In machine learning, the system finds the features in the data which can be used for further steps such as the making of decisions or the provision of predictions. On a deeper technical level, deep learning refers to multilayer networks that use neural networks in a way that human

22

beings understand things. These networks process through various stages, from input to output. By applying mathematical rules we can formalize the AI processes to some degree, and therefore the mathematical model is used for describing all the steps and rules that the machine learning system undergoes during the learning process. Using this model, you can anticipate the AI's performance, discover weaknesses, and optimize the system. The cooperation between the AI mechanisms and the mathematical theories opens ways to construct systems which are intelligent, can learn by experience, and adapt in real situations. The discussed methodologies consist of classifiers for supervised machine learning, models from natural language processing for text and contexts, image processing techniques for visuals, and a few that are combining multiple senses. The presentation slides only mention these technologies and don't present any result of a new model that is trained or an experiment; therefore, the model described is merely an illustration or a method of scoring and it is not an actual result of measurements. This approach has one great limitation: it is sensitive to the use of vague terminology, which can lead to different interpretations, and also to language style being changed over time that could make the same term mean different things.

A. Candidate Techniques

On the other hand, the advantage of natural language processing models is that they can interpret the content as a whole and thus are better at detecting sarcasm or subtle meaning. Visual processing techniques can look for clues that are visually conveyed; multimodal models can be trained to take advantage both of word-based features and of picture-based ones. The study of interactions through social network may help identify the relations between accounts or the posts which are allowed for the public observation. It is important therefore, when comparing the proposed techniques with the others, that the same data splits, evaluation methods, and criteria of human review be used for assessing them.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

It should also be pointed out that in general more advanced models do not mean that they will deliver the better or even the most suitable results.

B. Risk Score

Let T stand for the text analysis signal, I the image analysis one, and C be a public contextual signal. So, the conceptual score, R, would be $R = w_tT + w_iI + w\square C$, the weights being selected and tested via an appropriate dataset. If the score R is above a given level, τ, the item will be reviewed by a human first. We also want to make it clear that this formula is merely an example of possible combination of the signals; the paper does not make any claims at all about the precise values of the weights, the threshold, or the results. If we set the τ lower, a greater quantity of items will be reviewed but there will also be more of false positives. However, with a higher set of τ we may overlook relevant ones.

C. Evaluation Measures

Precision, recall, F1-score, false-positive rate, and processing latency are all metrics to think about while rating. Precision counts the number of flagged items against the number of correct answers provided in labels; recall counts the number of actual positive cases that were caught. Precision and recall values are incorporated into the F1-score. Genuine cases may not be found frequently and also there may be uncertainties in the labels, so it is not fair to rely solely on accuracy. It is essential to have representative and carefully labelled data, to be mindful that there will be no leakage between the training and test sets, to evaluate the system with different languages and varying terminology, and to involve human review of the false positives and negatives.

A. Data Processing and Deployment Considerations

Depending upon the model size, data flow and volume of data, the latency requirements, and security policy, the processing may occur on a desktop machine, behind the firewall of a managed server, or even on the cloud. This paper does not recommend any special system of deployment or that the real-time aspect has been shown. In any case, deployment would involve controlling access, secure storage, recording and monitoring and documenting a retainable policy.

B. Operational Applications

Public posts can be filtered or grouped for analyst review and new content from the public can be brought onto the screen or system for the use of public keywords, recurring or public image patterns or other ways of detection by law enforcers or digital forensics investigators, for instance. Also included would be the identification of relations visible from publicly accessible posts. While the system could be of great help as a source for organizing information to trained investigators, no individual should be labeled as guilty merely by virtue of his profile being flagged by the system. Analysts need the whole story with different possible explanations besides the source reliability to make any conclusion.

C. Advantages

Possible benefits comprise the screening of large volumes of public material, saving time on repetitive manual searching, the blending of textual and visual cues, and aiding an individual who is a professional in their work area to focus on the most valuable material to look into. Outputs that can

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

be explained may facilitate the auditability of the processes involved. These are the potential positive aspects of the suggested method and they need to be tested via controlled experiments; presenting them as results from a deployed system would be incorrect.

A. Data and Model Limitations

Public posts may present information in an incomplete, misleading, sarcastic or disconnected manner from their original setting. Coded language updates rapidly, which can be region or community specific. Dataset imbalance, poor labelling or unrepresentativeness may lead to false alarms and overlooked cases. Images may be not only blurry but also not matching with the accompanying text. Models may also drift as language and platform behavior change, therefore validation and documentation of system performance should be an ongoing concern.

B. Privacy and Cybersecurity

Limit analysis only to publicly accessible material gathered for a specific and lawful reason. The system can access public material only with compliance with the platform policy and local legislation laws which means that it should not access private content without any proper authorisation and that it should minimize the quantity of personally identifiable information collected. To keep such information, the data should have a limited period of existence and be kept confidential by authorized personnel only. Before the start of operational use, security mechanisms, retention limits, audit logs, and risk assessment should be set up, together with NIST AI Risk Management Framework (AI RMF 1.0, 2023), listed in the seminar PPT, as structured AI risk management is one of the areas, the reference is very relevant.

C. Human Review and Accountability

AI-generated content is only a first-level alert for further investigation, and therefore does not confirm that a person has committed a crime. Human review is the next step where one needs to understand the situation, verify the source, record the basis for taking a decision, and change erroneous flagging records. This paper reference to having humans keep in the loop of decisions (ICWSM, 2025) points out that human-centered annotation and oversight are crucial. Explanation of AI actions, responsibility for its use, and error correction methods are vital for the ethically sound application of AI. To put it briefly, AI-based OSINT can help to compile publicly available social-media data for manual review at a later stage however, its effectiveness must be verified through open evaluation of lawful and representative data. The proposed workflow and risk score that were described in the main text of the

33

paper, are merely a concept for the research discussion - there is no attempt of claiming real results or the development of a system is involved.

REFERENCES

1. World drugs report, 2025 by United Nations Office on Drugs and Crime (UNODC).
2. Keeping Humans in the Loop: Human Centred Automated Annotation with Generative AI. Proceedings of ICWSM, 2025.

BODHIVRUKSHA JOURNAL OF DIVERSE DISCIPLINE
E-ISSN : 3139-1486
An Open Access, Peer-Reviewed, Multidisciplinary Scholarly Journal
Bi-monthly | Multilingual (English, Hindi, Marathi) | Academic Research
@ Volume 2, Issue 5 (September-October 2026) (Special Issue 1) ISSN: 3139-1486 (Online)

3. Analyzing Textual Context and Intent in Social Media: A Unified Framework. ACM Transactions on Intelligent Systems and Technology, 2024.
4. Knowledge Prompted ChatGPT: Enhancing Drug Trafficking Detection on Social Media. Information & Management, 2024.
5. 2024 World drugs report by United Nations Office on Drugs and Crime (UNODC).
6. Europol. IOCTA (Organised Crime Internet Threat Assessment), 2024.
7. NIST publishes the AI RMF v.1.0, 2023.
8. Alzahrani, M. and others. Computer vision technology for content analysis through the internet in 2023.
9. Yang, X., et al. NLP for the intelligence of social media media in 2021.
10. Alsaedi, N. 与其他的 Machine learning techniques for social media mining and analysis, 2017
11. Drug Supply Networks: A Systematic Review of the Organization Structure of Illicit Drug Trade. Crime Science, 2017.
12. Social Network Analysis of Australian Poly-Drug Trafficking Network. Social Networks, Vol. 51, 2017.
13. Risk, Price and Positions. International Journal of Drug Policy, Vol. 25, 2014.
14. Ferrara, E., et al. Social Bot Increase. Communications of the ACM, 2016.
15. [15], Crime, Law and Social Change 2012.