

SHARING RESEARCH FINDINGS

SCOPING REVIEW OF AI/ML APPLICATIONS IN SUBSTANCE USE PREVENTION ON SOCIAL MEDIA



RESEARCH TEAM

Researchers



LÊ PHAN KHÁNH HUY



NGUYỄN THANH VÂN

Scientific Supervisor



NGUYỄN THU TRANG, PhD

PART 1

CONTEXT, OBJECTIVES,
& METHODOLOGY

CONTEXT

Substance use remains a major public health challenge globally, with profound impacts on individuals, families and society.

Individuals	Families	Society
In 2022, over 64 million people globally experienced substance use disorders, but only 1 in 11 received treatment.	Parental substance use can lead to family separation and severe negative outcomes.	Substance use creates a great economic burden.
WHO reports 600,000 deaths annually due to drug-related neurotoxicity.	Children with substance-using parents face elevated risks of health, psychological, and social issues, with potential for substance use replication.	In Canada, substance-related costs exceed 49 billion CAD/year, with 13.1 billion for healthcare.
Of these, opioids cause more than 75% of deaths, with amphetamines implicated in nearly 50,000 cases each year.	Between 2009 and 2014, approximately 8.7 million children under 18 lived with a parent with a substance use disorder.	In South Korea (2015), estimated \$2.65 billion USD spent on opioid abuse.

CONTEXT

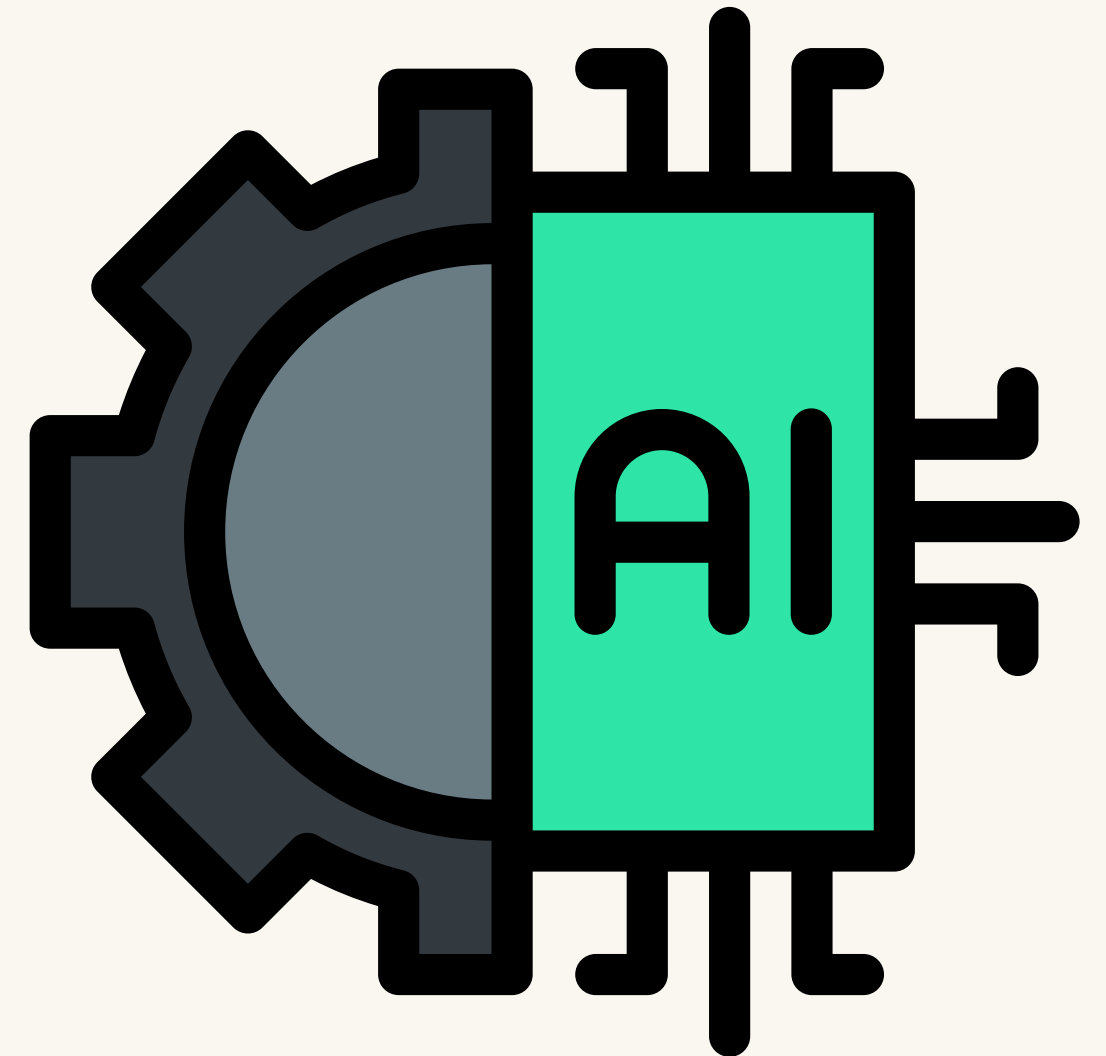


- Social media offers broad reach and interactivity, particularly for youth.
- As of 2024, about 1/3 of U.S. teens use social media daily; nearly 50% are online constantly.
- Social media is both a risk factor and a prevention tool in shaping young people's attitudes and behaviors toward substance use.
 1. Pros: Provides information about community resources and important health issues.
 2. Cons: Peer pressure increases as teens are exposed to pro-substance use content from celebrities and social media influencers.

CONTEXT

- Artificial intelligence (AI) and machine learning (ML) technologies enhance the potential of social media by analyzing large amounts of unstructured data, allowing systems to learn patterns and make predictions without being explicitly programmed.
- AI helps track trends, predict risks, and detect risky behavior early in real time.
- Study (Prochaska et al., 2020) shows that personalized AI increases craving control and reduces substance use.

=> AI has the potential to deliver data-driven, personalized, and scalable solutions to preventative programs efficiently.



OBJECTIVES

- 01** Describe the types of AI/ML applications in substance use prevention on social media currently available in literature.
- 02** Explore common challenges and potential applications of AI/ML in substance use prevention on social media.

METHODOLOGY

Scoping Review

07 databases: PubMed, Sage Journal, Embase, PsycINFO, ACM Library, IEEE Xplore, Scopus

**STUDY
DESIGN**

**INCLUSION
CRITERIA**

DATABASES

**RESEARCH
TIMELINE**

Must be an article published in a **peer-reviewed journal**.

Features a **clear discussion** on AI/ML application in addiction prevention on social media

Total research duration:
11/2024 - 06/2025

Data collection duration:
12/2024 - 02/2025

METHODOLOGY

SEARCH TERMS ON DATABASES

Search Term 1	Search Term 2	Search Term 3
Artificial intelligence [MeSH]	Social media [MeSH]	Substance-Related Disorders [MeSH]
Machine Learning [MeSH]	Twitter (X)	Substance Use
Deep Learning [MeSH]	Facebook	Tobacco
Chatbot	Instagram	Alcohol
AI assistant	YouTube	Marijuana
ChatGPT	Reddit	Amphetamine/ATS

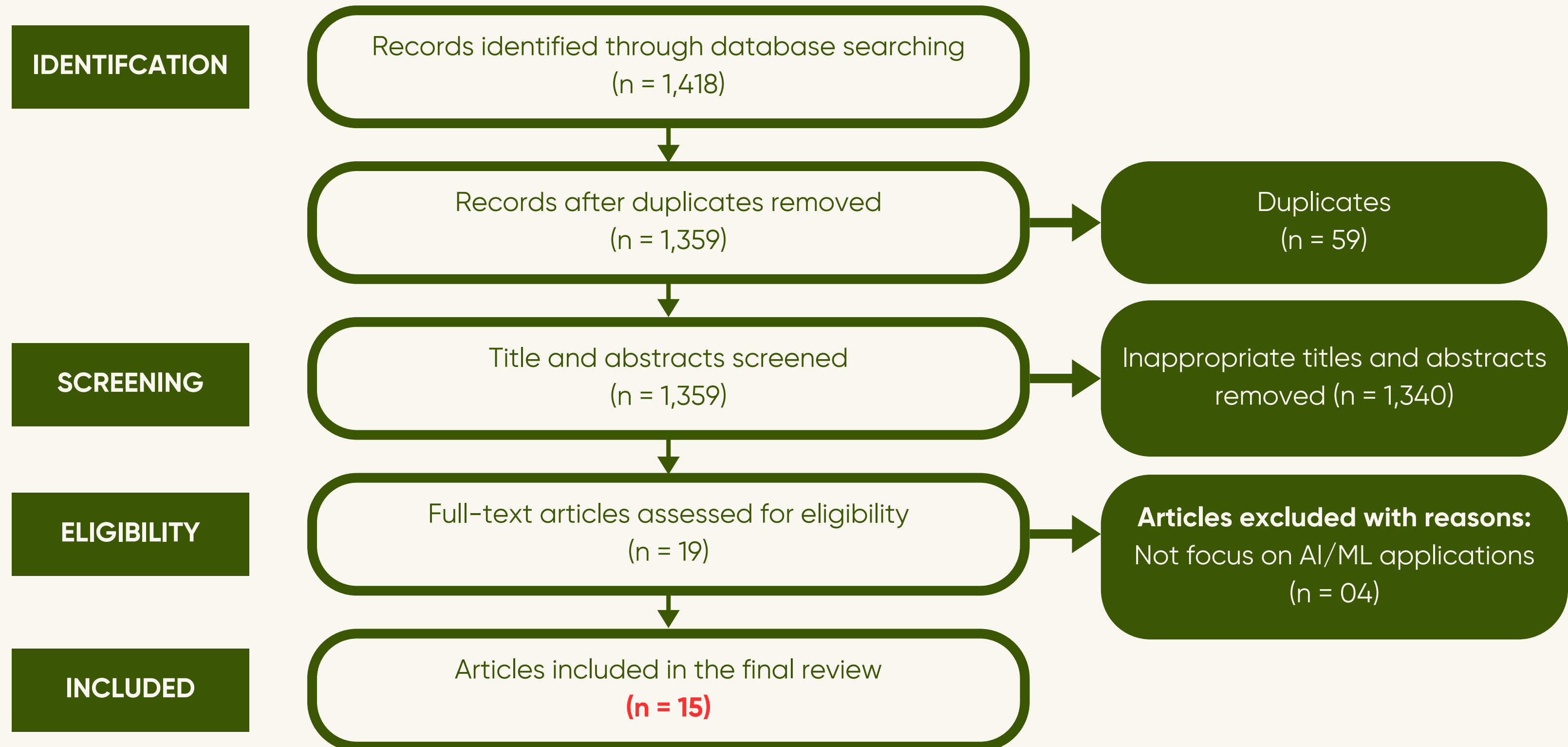
How search terms were used:

- Use OR for phrases/synonyms within the same Term.
- Use AND to join different Terms together.

PART 2

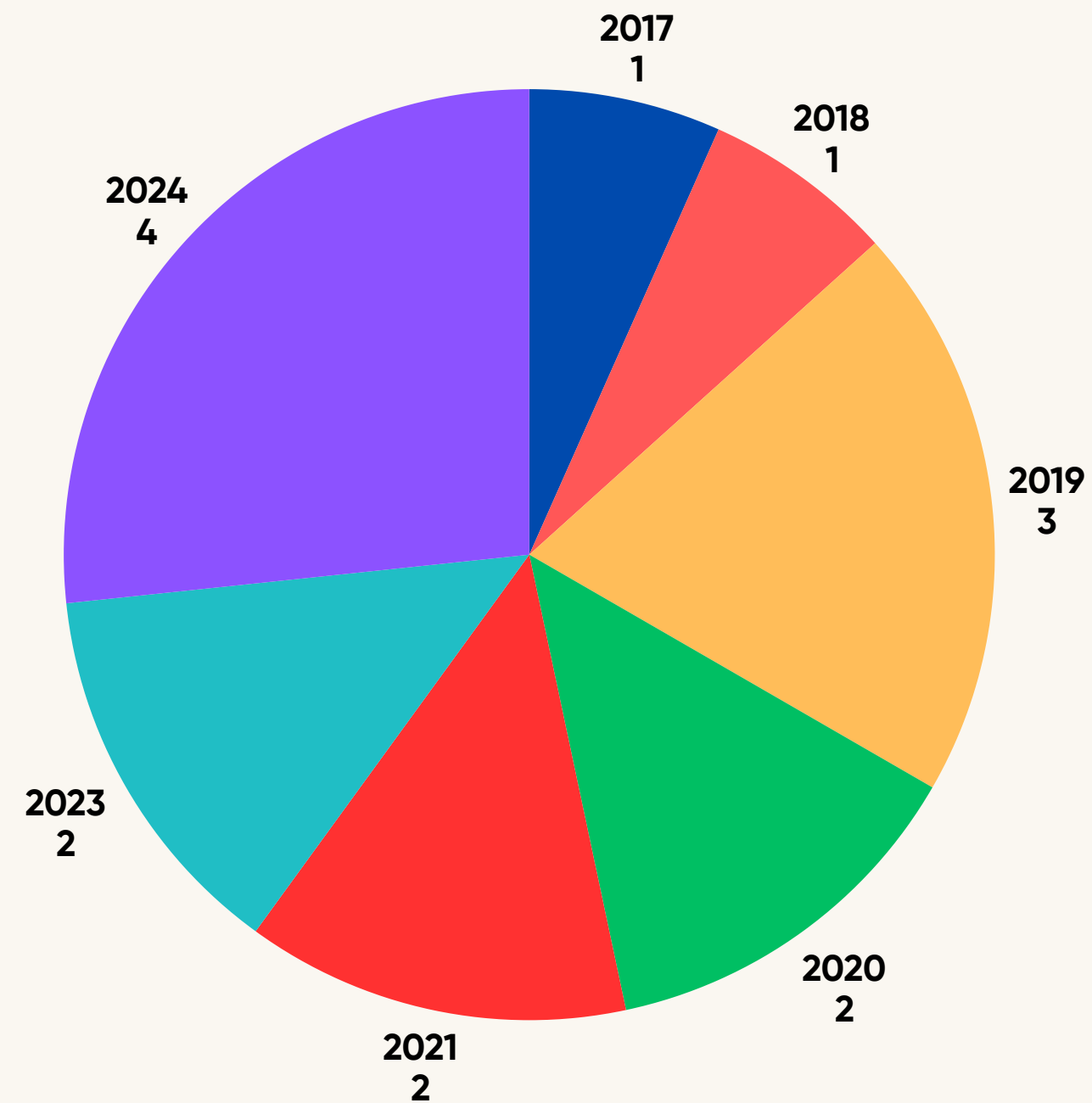
RESEARCH FINDINGS

PRISMA FLOWCHART

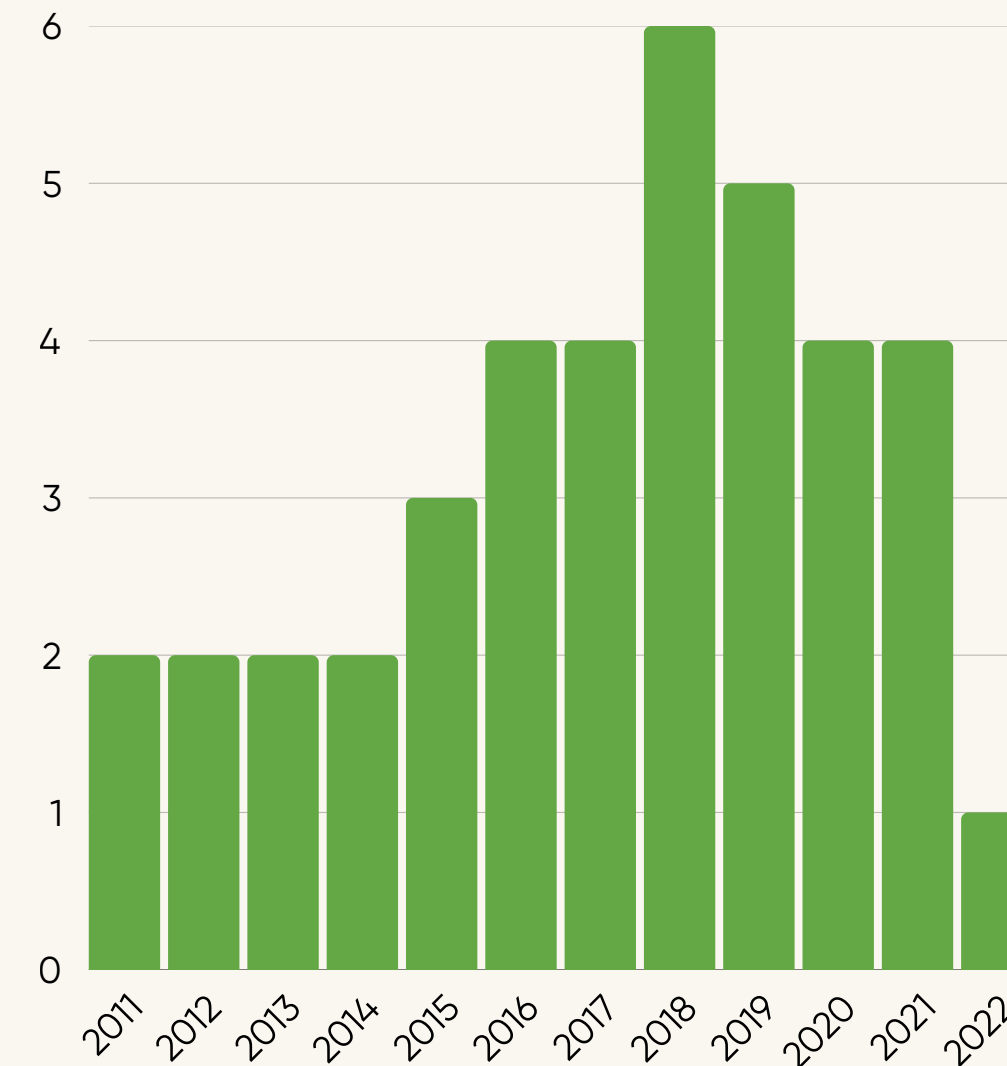


FINDINGS

Year of publication of articles



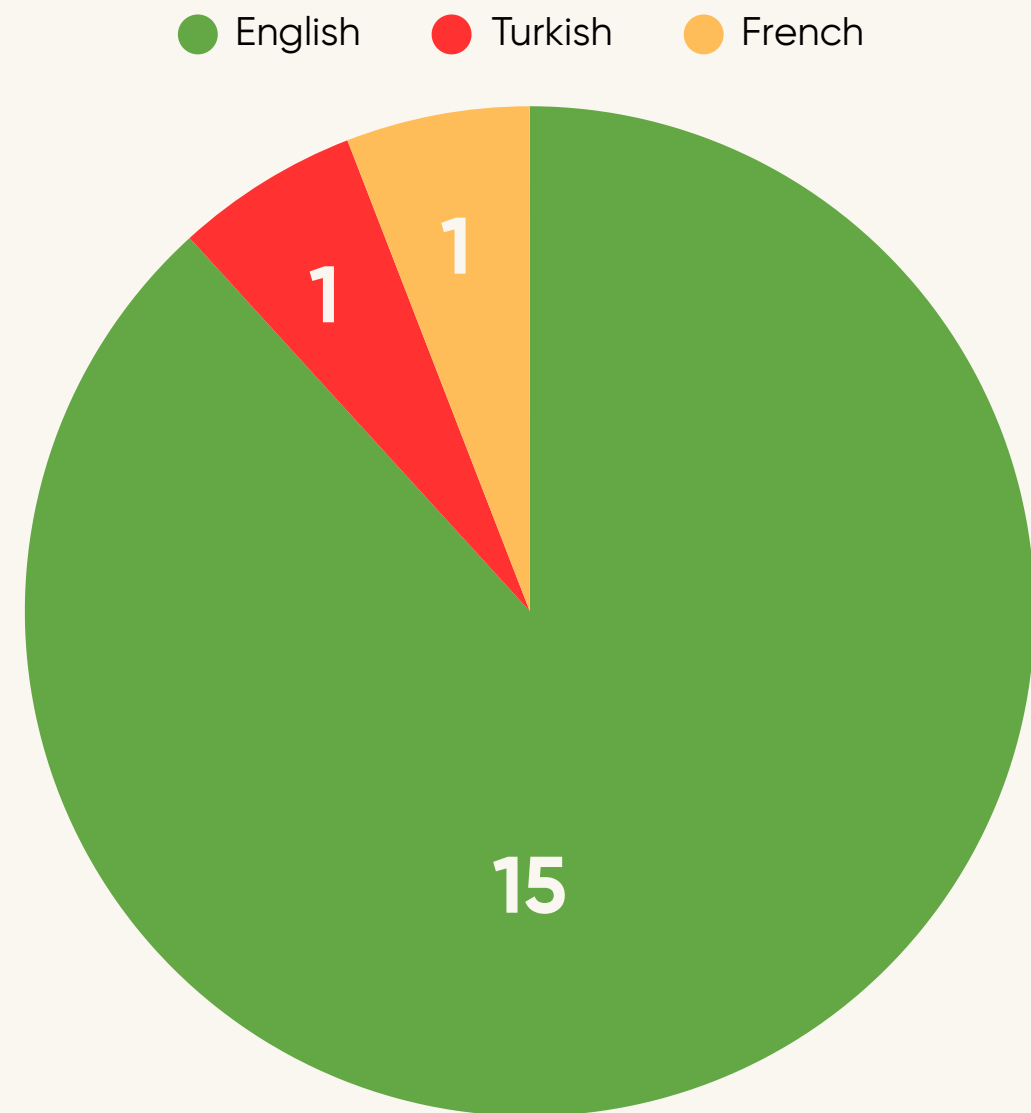
Data collection years



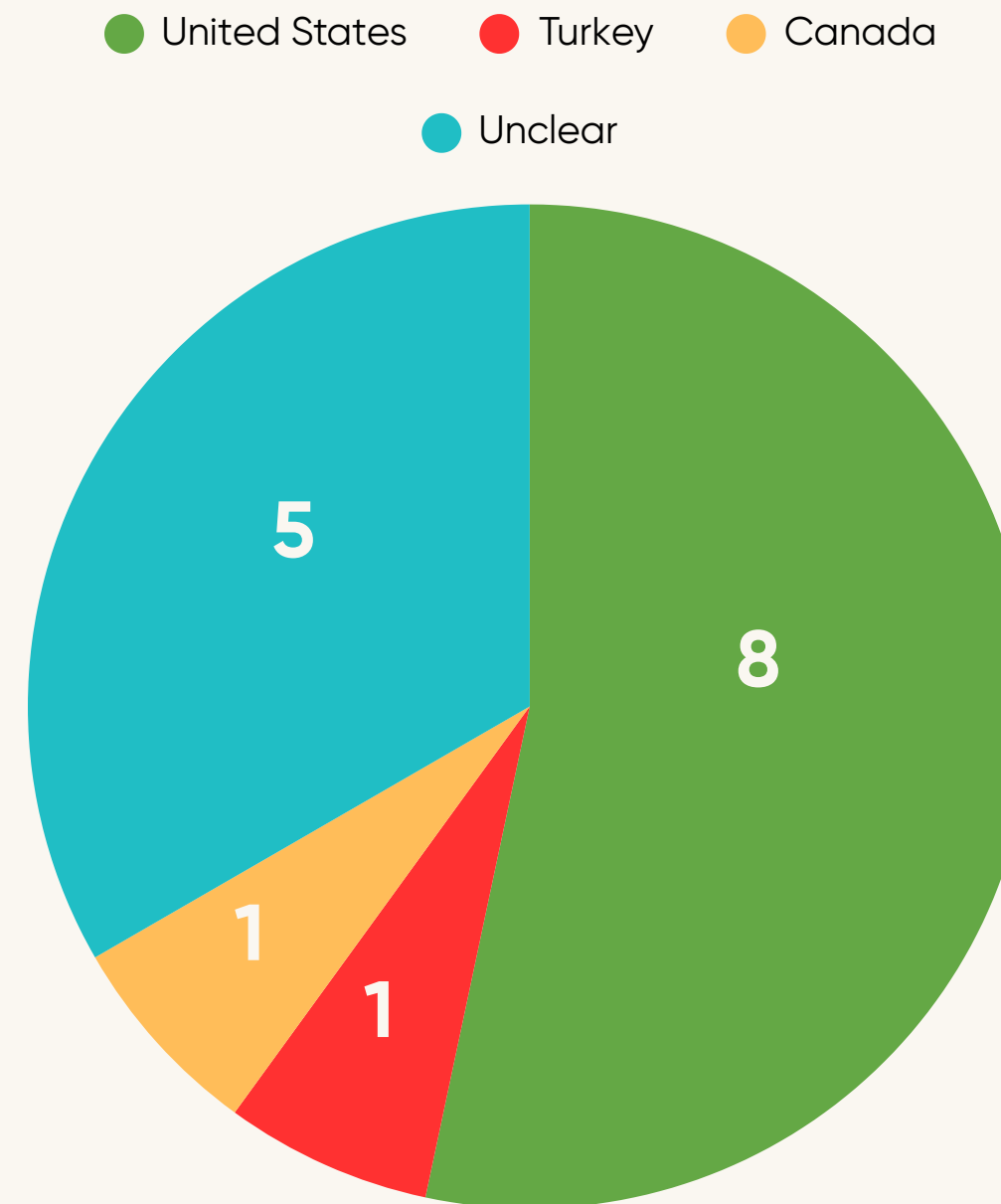
***Note:** $n = 1$ article (Ding et al., 2020) did not clearly state the year of data collection for their study.

FINDINGS

Languages of collected data



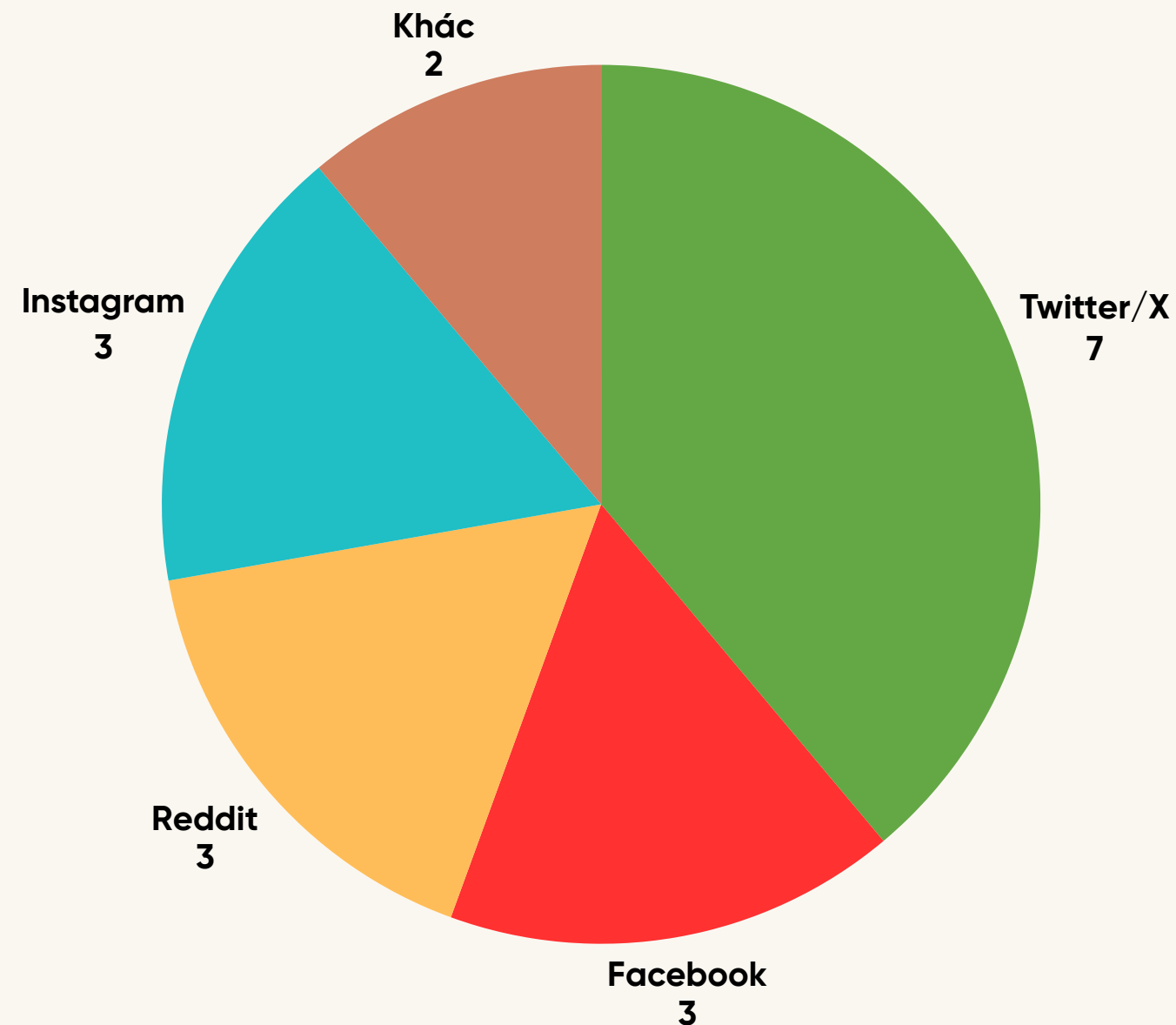
Countries in which data were collected



***Note:** *n* = 2 articles (Fisher et al., 2023; Küçükali & Erdoğan, 2024) used multiple languages (English and a secondary language).

FINDINGS

Data sources



**Note: n = 1 article (Ovalle et al., 2021) collected data from multiple sources (Facebook, Twitter and 2 websites)*

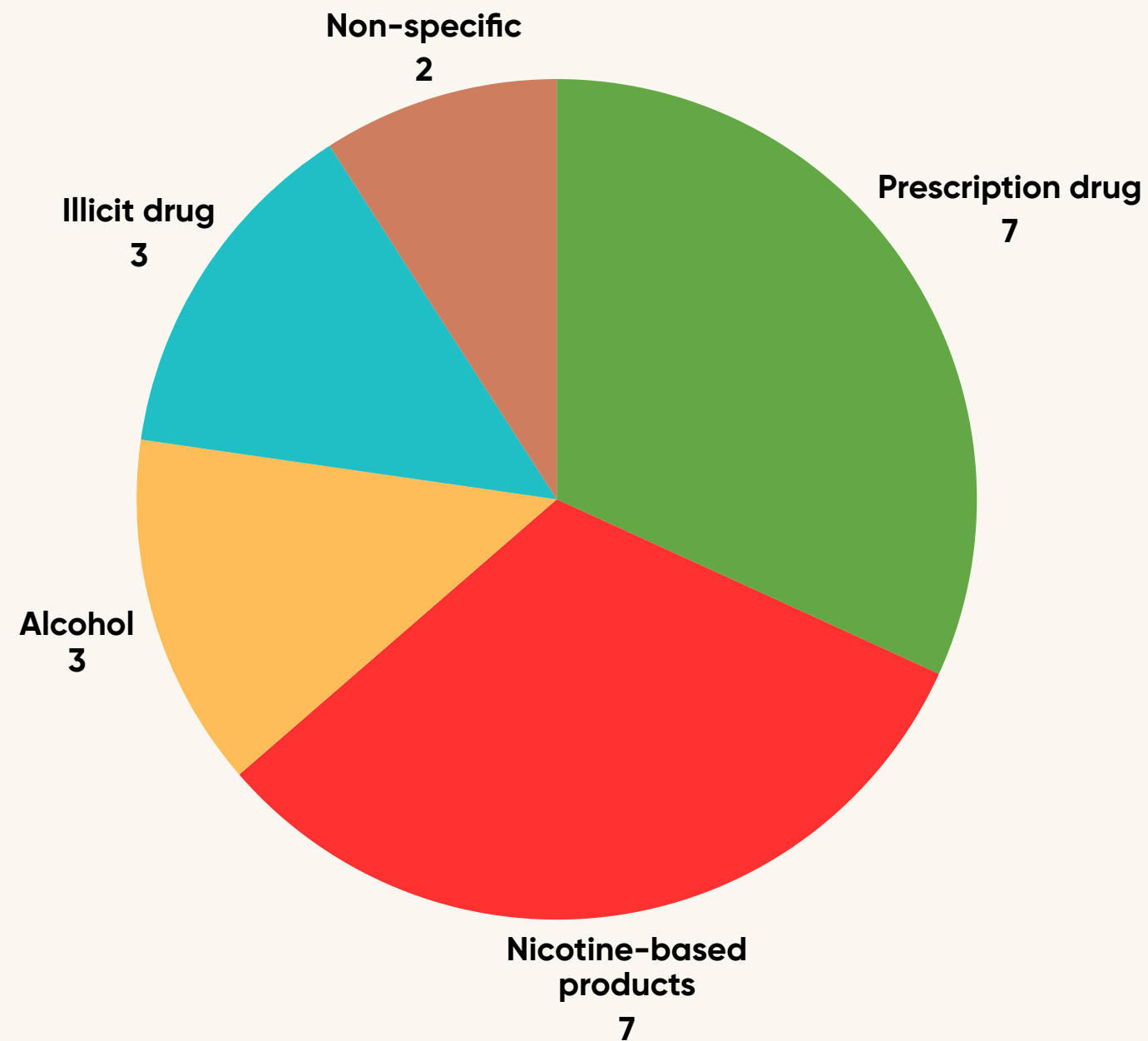
DATA SUMMARY:

The vast majority of data used is **text-based**. Specifically, the number of data types is as follows:

- Text-based data: **133,314,480** (including posts, and comments)
- Image-based data: **476,019** (currently only applied to studies on Instagram)
- Others: **253,980** (data on Likes - Page Likes - of social pages on Facebook)

FINDINGS

Substances of Interest



DATA SUMMARY:

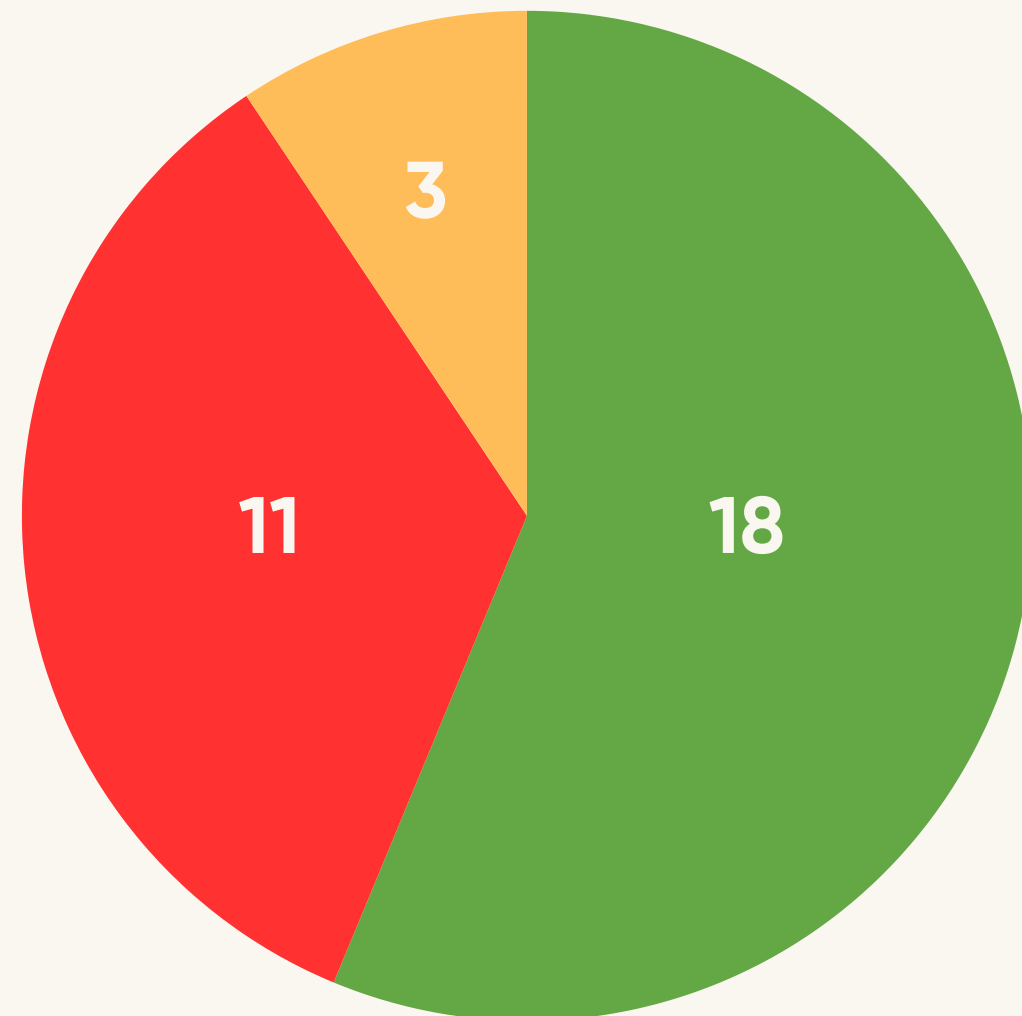
Prescription drugs (which include fentanyl and methylphenidate) and **nicotine-based products** (tobacco and vaping) are frequently targetted in the included studies, both mentioned on 7 articles.

Meanwhile, **alcohol** and **illicit drugs** (non-specific) followed the ranking with mentions in 3 articles each.

FINDINGS

AI/ML Models in application

- Natural Language Processing (NLP)
- Traditional Machine Learning (ML)
- Computer Vision (CV)



NOTES AND SUMMARY:

Natural Language Processing (NLP) models include models that help AI understand and process text-based data. Models used include: BERT, RoBERTa, Zero-shot Classifier, Word2Vec, GloVe, LDA, BTM and some new/unknown models.

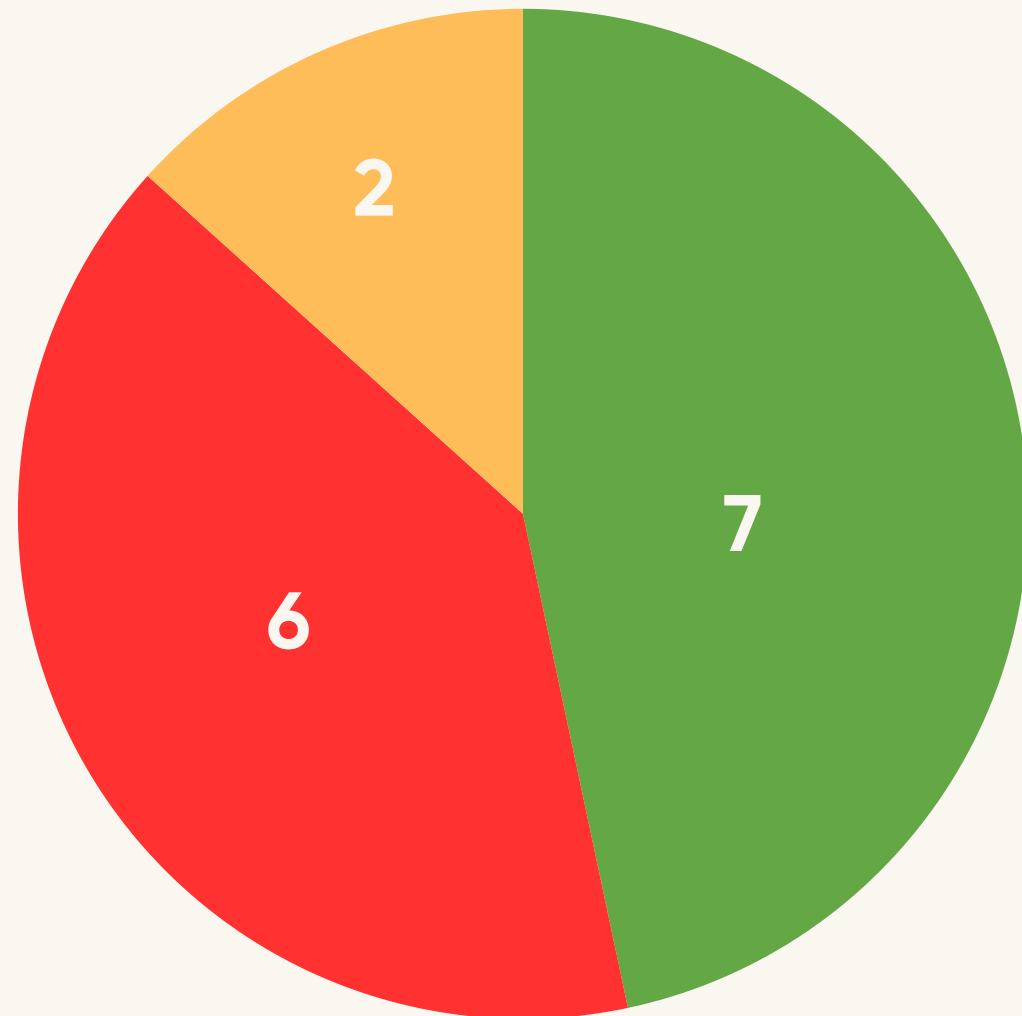
Traditional Machine Learning (ML) models include models that apply mathematical equations to make predictions. Models used include: SVM, Logistic Regression, Naive Bayes, Random Forest and GBT.

Computer Vision (CV) models include models that help AI recognize content in images and videos. Models used include: ResNet18, AlexNet, and CLIP.

FINDINGS

AI/ML application goal

- Epidemiological surveillance and monitoring
- Behavioral analysis and prediction
- Community campaign evaluation



NOTES AND SUMMARY:

Epidemiological surveillance and monitoring: 07 articles applied AI/ML models with the aim of screening, monitoring trends (in terms of language and user perspective) and discovering risk factors for substance addiction.

Behavioral analysis & prediction: 06 articles applied AI/ML models with the aim of analyzing behavior and predicting the risk of using/relapsing addictive substances.

Community campaign evaluation: In particular, 02 articles applied AI/ML models to support the evaluation of the effectiveness of community communication campaigns on social media and analyzing success factors for the campaigns.

FINDINGS

Common challenges:

- **Technical challenges:** Currently, the vast majority of articles pointed out that the biggest weaknesses of their studies were in the technical aspect, including: limited amount of data, bias towards some groups is unavoidable, unreliable performance, and difficulty in accessing/extracting data sources.
- **Ethical and legal challenges:** Due to the nature of social media included in the studies, the articles all raise concerns about security issues as well as the lack of legal corridors to regulate the use of personal data of social media users.

Potential and future directions:

- **Collaboration and resource mobilization for research and development:** Technical issues related to AI/ML models can all be solved through increasing resources for research and development of general and specific models for locality.
- **Applications in public health and substance use prevention:** The diverse applications paint a potential future in applying AI/ML in public health projects of government agencies, research institutes and community organizations interested in substance use prevention, especially among young people.

PART 3

DISCUSSION

DISCUSSION

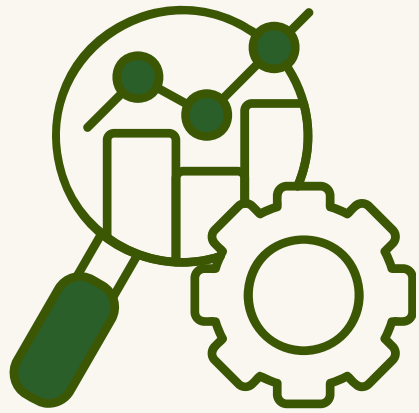
From the results of the scoping review, the following conclusions can be drawn:

- The data analyzed is still largely **text-based, in English, and extracted from popular Western social media outlets** such as Twitter, Facebook, Reddit, and Instagram. The biggest challenge at present is **adjusting these methods to local language (with greater complexity for AI/ML models), cultural norms and popular local social media sites.**
- The outputs of the studies show positive results (approximately 70–80% reliability) when applying AI/ML in understanding the status, attitudes and behaviors as well as making risk predictions. However, due to **anonymity nature of social media data**, existing studies still **lack inclusivity and targeting of vulnerable groups** such as youths, LGBTQIA+, ethnic minorities, etc.
- Data was collected “cross-sectionally” using traditional API (application programming interface) method and **lack behavioral tracking over-time.** Though alternative methods of data donation and screen tracking are promising, they require much **more resources and time.**

DISCUSSION

- Current AI/ML applications are effective but only **limited to controlled settings** due to **limitations in adapting to fast-evolving “social media languages”** (e.g. slang, emojis, coded terms, etc.). It requires future models to be more dynamic and learn to capture nuanced in daily communications.
- Current practices raised questions on **data privacy safeguarding and responsible AI deployment**. It highlights a need for government regulations on AI use in regards to private data and platforms need to adopt a collaborative ruleset to allow for researchers to leverage data safely.

RECOMMENDATIONS



For researchers: More research and development efforts are needed from public health experts collaborating in multiple disciplines, especially computer scientists and AI developers, on the topic of “AI & Substance Use”, not only expanding data sources to perfect models but also needing local projects, focusing on adapting models and tools to Global South.

For government agencies: The utilization of social media data for research purposes is still a difficult problem, with no clear regulations. Therefore, legislators can consider discussing and issuing regulations to protect the security and privacy of users’ data while serving the effective development of knowledge.



For community organizations: Enhance knowledge and capacity to leverage and keep up with new 4.0 technologies such as AI/ML and social media to enhance the effectiveness of community campaigns not only in terms of communication but also in community interventions and research.

***IF YOU HAVE ANY QUESTIONS OR CONCERNS ABOUT
THE STUDY, PLEASE CONTACT THE RESEARCH TEAM
AT khanhhuy@wecare.org.vn***