

IA e social media contro l'uso di sostanze

di Paolo Berretta

Opportunità emergenti per la prevenzione, tra sorveglianza digitale, personalizzazione dei messaggi e tutela dei dati

| | |
|-------------------------|--|
| Meta descrizione | L'IA può supportare la prevenzione dell'uso di sostanze, ma servono evidenze più robuste su esiti reali e tutela dei dati. |
| Tag | intelligenza artificiale; social media; prevenzione; uso di sostanze; adolescenti; salute pubblica; dipendenze; machine learning; privacy; sanità digitale |

Indice

- IA e social media: strumenti emergenti, non soluzioni autonome
- Social media: ambiente di esposizione e canale preventivo
- Che cosa può fare l'IA
- La revisione esplorativa: pochi studi, molti limiti
- Sorveglianza digitale e sistemi di allerta precoce
- Predizione del rischio: utile, ma non diagnostica
- Personalizzazione dei messaggi preventivi
- Valutazione delle campagne
- Il nodo dello slang, delle emoji e dei linguaggi locali
- Privacy, consenso e governance
- Data donation e screen tracking: possibilità e cautela
- Un modello operativo realistico
- Conclusione
- Take home messages
- Tabella delle fonti aggiunte
- Link cliccabili di approfondimento

IA e social media: strumenti emergenti, non soluzioni autonome

L'intelligenza artificiale sta entrando nel campo della prevenzione dell'uso di sostanze attraverso un canale ormai centrale nella vita quotidiana di adolescenti e giovani adulti: i social media. Le piattaforme digitali sono oggi luoghi di relazione, esposizione, normalizzazione dei comportamenti, costruzione dell'identità e accesso a informazioni più o meno affidabili. Per questo motivo possono contribuire sia alla diffusione di contenuti favorevoli all'uso di sostanze, sia alla costruzione di messaggi preventivi più tempestivi, credibili e aderenti ai linguaggi dei destinatari.

La tesi scientificamente più solida, tuttavia, è prudente: l'IA non previene da sola l'uso di sostanze. Può supportare la prevenzione attraverso sorveglianza digitale, analisi dei trend, personalizzazione dei messaggi e valutazione delle campagne, ma l'evidenza sugli esiti preventivi reali è ancora preliminare e richiede validazione attraverso studi longitudinali e interventistici. Questa distinzione è decisiva: monitorare conversazioni online, classificare contenuti o stimare un rischio non equivale a dimostrare una riduzione dell'iniziazione, del consumo o dei comportamenti problematici.

Il contesto rende il tema rilevante. I dati WHO Europe/HBSC mostrano che l'uso problematico dei social media tra gli adolescenti è aumentato dal 7% nel 2018 all'11% nel 2022; la stessa fonte segnala che oltre un terzo dei giovani riferisce contatti online costanti con gli amici e richiama la necessità di alfabetizzazione digitale, tutela della salute mentale e interventi educativi adeguati all'età. In parallelo, il rapporto ESPAD 2024, condotto su 113.882 studenti di 15-16 anni in 37 Paesi europei, conferma che l'uso di sostanze tra gli adolescenti resta un tema di salute pubblica: il 13% degli studenti riferisce almeno un uso nella vita di sostanze illecite, mentre il 44% ha provato sigarette elettroniche almeno una volta.

Social media: ambiente di esposizione e canale preventivo

I social media hanno una doppia natura. Da un lato espongono adolescenti e giovani a contenuti che possono normalizzare alcol, cannabis, nicotina, vaping o altre sostanze. Dall'altro offrono canali di comunicazione preventiva difficilmente raggiungibili con strumenti tradizionali.

Uno studio qualitativo pubblicato su Health Education Research ha intervistato adolescenti tra 14 e 18 anni sulla loro esposizione a contenuti relativi alle sostanze e ai messaggi di prevenzione. La maggior parte dei partecipanti aveva visto post di pari che rappresentavano l'uso di sostanze in modo positivo, mentre i messaggi formali di prevenzione erano spesso percepiti come meno credibili o troppo distanti dall'esperienza osservata online.

Il problema non è solo la quantità di esposizione, ma la qualità narrativa del contenuto. I post dei pari tendono a mostrare socialità, appartenenza, divertimento e identità. I messaggi preventivi istituzionali, invece, spesso insistono solo su rischio, danno e divieto. Quando le due rappresentazioni divergono troppo, il messaggio preventivo perde credibilità. La prevenzione sui social deve quindi essere scientificamente corretta, ma anche comunicativamente realistica: deve riconoscere il linguaggio dei giovani, evitare paternalismo e costruire fiducia.

Che cosa può fare l'IA

L'intelligenza artificiale applicata ai social media può operare su grandi quantità di dati non strutturati: testi, immagini, video, commenti, interazioni, hashtag, emoji e pattern di engagement. Le principali famiglie di strumenti sono tre: natural language processing, machine learning tradizionale e computer vision. La presentazione collegata alla revisione esplorativa ISSUP/Hanoi Medical University/We Care Social Enterprise indica che gli studi inclusi utilizzano modelli NLP come BERT, RoBERTa, Word2Vec e LDA; modelli di machine learning come SVM, regressione logistica, random forest e gradient boosting; e modelli di computer vision come ResNet18, AlexNet e CLIP.

Il natural language processing consente di analizzare testi, commenti, post e conversazioni, individuando termini ricorrenti, sentiment, percezioni del rischio e possibili segnali emergenti. Il machine learning tradizionale può classificare contenuti, stimare pattern di rischio o associare determinati segnali digitali a comportamenti osservati. La computer vision permette invece di riconoscere immagini o oggetti collegati a specifici comportamenti, come dispositivi per vaping, strumenti per il fumo o contesti visivi ricorrenti.

Queste tecnologie non devono però essere confuse con una diagnosi clinica o con una valutazione individuale certa. L'IA può individuare probabilità, pattern e correlazioni. Non può stabilire, da sola, se una persona abbia un disturbo da uso di sostanze, né distinguere con sicurezza tra uso episodico, uso rischioso, uso problematico e dipendenza sulla base di un singolo contenuto social.

La revisione esplorativa: pochi studi, molti limiti

La fonte scientifica centrale è la revisione esplorativa di Nguyen e colleghi, pubblicata su Public Health nel 2026. La revisione esamina le applicazioni di IA e machine learning sui social media per la prevenzione dell'uso di sostanze e conclude che il campo è ancora dominato da attività descrittive, come monitoraggio dei trend e predizione del rischio, con forte dipendenza dal natural language processing. Gli autori segnalano barriere importanti: ambiguità linguistica, squilibrio dei dati, limiti etici e necessità di validazione umana.

La presentazione metodologica della review mostra un processo selettivo rigoroso: 1.418 record identificati, 1.359 dopo rimozione dei duplicati, 19 articoli valutati full-text e 15 studi inclusi nella revisione finale. I database considerati comprendevano PubMed, Sage Journal, Embase, PsycINFO, ACM Library, IEEE Xplore e Scopus.

Un primo limite riguarda la rappresentatività. I dati analizzati sono prevalentemente testuali, in inglese e provenienti da piattaforme occidentali come Twitter/X, Facebook, Reddit e Instagram. La presentazione segnala oltre 133 milioni di dati testuali, circa 476 mila dati immagine e circa 254 mila dati relativi a like o page-like; tuttavia, il peso dei dati testuali è dominante. Questo significa che molte aree linguistiche, culturali e geografiche restano sottorappresentate.

Un secondo limite riguarda gli obiettivi applicativi. Nei 15 studi inclusi, sette lavori riguardavano sorveglianza e monitoraggio epidemiologico, sei analisi comportamentale e predizione, e solo due valutazione di campagne comunitarie. Questo conferma che la letteratura attuale è più forte nel descrivere o classificare fenomeni digitali che nel dimostrare effetti preventivi concreti.

Un terzo limite riguarda la valutazione degli esiti. La review e la presentazione indicano performance positive in alcuni studi, approssimativamente nell'ordine del 70-80% per specifici compiti di classificazione, comprensione di atteggiamenti, analisi comportamentale o predizione del rischio. Questo dato non va però interpretato come "l'IA previene l'uso di sostanze con un'affidabilità del 70-80%". Le metriche sono eterogenee, dipendono dal dataset, dal modello, dalla sostanza analizzata e dal compito specifico.

Sorveglianza digitale e sistemi di allerta precoce

Una delle applicazioni più concrete dell'IA riguarda la sorveglianza digitale. L'analisi automatizzata dei social media può aiutare a identificare conversazioni emergenti su nuove sostanze, prodotti a base di nicotina, farmaci usati impropriamente, pattern territoriali o mutamenti del linguaggio. In questo senso, l'IA può rafforzare i sistemi di early warning, non sostituirli.

Uno studio pubblicato sul Journal of Medical Internet Research ha esaminato come l'IA possa contribuire a identificare pattern di rischio e danno droga-correlati monitorando attività social di gruppi pubblici e istituzionali. Lo studio ha raccolto oltre 40 mila tweet da 145 account rilevanti in tre province canadesi, con l'obiettivo di supportare l'identificazione di trend emergenti nelle comunità.

Questo tipo di applicazione è utile sul piano della salute pubblica perché riduce il ritardo tra comparsa di un segnale e risposta istituzionale. Tuttavia, il passaggio dal segnale digitale all'intervento preventivo richiede verifica umana, contestualizzazione epidemiologica e integrazione con fonti tradizionali: servizi territoriali,

IA e social media nella prevenzione dell'uso di sostanze
pronto soccorso, tossicologia forense, osservatori epidemiologici, dati scolastici e sistemi di sorveglianza nazionali.

Predizione del rischio: utile, ma non diagnostica

Un'altra area promettente è la stima del rischio. Modelli di deep learning possono analizzare profili social e contenuti multimodali per identificare pattern associati a uso di alcol, tabacco o altre sostanze. Uno studio pubblicato su *Neuropsychopharmacology* ha sviluppato un metodo basato su immagini e testi Instagram per classificare il rischio di uso di alcol, tabacco e droghe in 2.287 utenti attivi, utilizzando reti neurali convoluzionali per le immagini e modelli LSTM per il testo.

Il valore di questi studi è metodologico: mostrano che i dati social possono contenere segnali utili per la ricerca e la prevenzione di popolazione. Il limite è altrettanto chiaro: il rischio stimato da un modello non coincide con una diagnosi individuale. Inoltre, i dati social sono selettivi, performativi e incompleti. Non tutti gli adolescenti usano le stesse piattaforme, non tutti pubblicano contenuti rappresentativi dei propri comportamenti reali, e molti segnali possono essere ironici, imitativi, identitari o contestuali.

Per questo motivo, l'uso di modelli predittivi in prevenzione deve restare orientato alla popolazione e alla programmazione degli interventi, non alla classificazione automatica dei singoli individui. Una tecnologia che etichetta un giovane come "a rischio" senza verifica umana e senza garanzie procedurali può produrre stigma, sfiducia e danni relazionali.

Personalizzazione dei messaggi preventivi

L'IA può offrire un contributo importante nella costruzione di messaggi preventivi più pertinenti. Può analizzare quali temi generano maggiore attenzione, quali linguaggi risultano più comprensibili, quali immagini funzionano meglio per specifici target, quali contenuti sono ignorati o respinti. Può inoltre aiutare i professionisti a differenziare messaggi per età, piattaforma, contesto culturale e livello di rischio.

La personalizzazione non deve però diventare manipolazione. In salute pubblica, adattare un messaggio significa renderlo più chiaro, rilevante e accessibile, non sfruttare vulnerabilità cognitive o psicologiche. Questo è particolarmente importante quando il target è costituito da minori o gruppi vulnerabili. L'uso dell'IA generativa per produrre contenuti preventivi deve quindi prevedere revisione esperta, controllo delle fonti, verifica del linguaggio e valutazione del rischio di messaggi fuorvianti.

Un'esperienza interessante, anche se ancora pilota, è il programma "Living the Example", valutato su JMIR Mental Health. Il programma prevedeva la formazione di giovani nella creazione di messaggi preventivi da

IA e social media nella prevenzione dell'uso di sostanze
condividere sui social media, secondo una logica peer-to-peer. Questo modello suggerisce una direzione operativa robusta: l'IA dovrebbe supportare campagne co-progettate con i giovani, non sostituire la relazione educativa o la credibilità del pari.

Valutazione delle campagne

La valutazione delle campagne preventive è uno degli ambiti in cui l'IA può essere più immediatamente utile. Attraverso analisi del sentiment, engagement, commenti, condivisioni e pattern di interazione, è possibile comprendere se un messaggio viene visto, discusso, ignorato, contestato o frainteso.

Questa valutazione non deve però ridursi alla misurazione di like e visualizzazioni. L'engagement è un indicatore intermedio, non un esito preventivo. Una campagna può essere molto vista ma inefficace; può generare interazioni elevate perché controversa; può produrre commenti ironici o oppositivi. Gli indicatori più rilevanti dovrebbero includere conoscenze, percezione del rischio, norme sociali, intenzione di evitare o ridurre l'uso, richiesta di aiuto, accesso ai servizi e, quando possibile, cambiamenti comportamentali misurati nel tempo.

La revisione esplorativa indica che solo due degli studi inclusi riguardavano la valutazione di campagne comunitarie, confermando che questo settore è ancora poco sviluppato rispetto al monitoraggio e alla predizione. Per passare da sperimentazione tecnologica a prevenzione validata servono disegni longitudinali, gruppi di confronto, misure pre-post e outcome comportamentali.

Il nodo dello slang, delle emoji e dei linguaggi locali

Uno dei limiti tecnici più rilevanti riguarda l'evoluzione rapida dei linguaggi digitali. Slang, emoji, meme, abbreviazioni, ironia e termini in codice cambiano rapidamente e possono variare per età, comunità, lingua, piattaforma e contesto locale. La presentazione della review segnala che le applicazioni attuali sono ancora limitate da difficoltà di adattamento ai social media languages, inclusi slang, emoji e coded terms.

Questo limite è cruciale per la prevenzione. Un modello addestrato su dati in inglese e su piattaforme occidentali può funzionare male in contesti culturali differenti. Localizzare non significa tradurre parole da una lingua all'altra; significa comprendere norme sociali, codici comunicativi, piattaforme realmente usate, riferimenti culturali e sottogruppi digitali. Per questo motivo, i modelli dovrebbero essere sviluppati con approcci partecipativi, coinvolgendo giovani, operatori, esperti di salute pubblica, linguisti computazionali e comunità locali.

Privacy, consenso e governance

L'uso dei dati social in prevenzione apre questioni etiche rilevanti. Molti contenuti online sono tecnicamente accessibili, ma questo non significa che possano essere usati senza limiti per finalità di ricerca o salute pubblica. La raccolta automatizzata dei dati, soprattutto se riferita a minori o gruppi vulnerabili, richiede basi giuridiche solide, minimizzazione, anonimizzazione o pseudonimizzazione, protezione dei dati e valutazione del rischio.

Il GDPR protegge i dati personali indipendentemente dalla tecnologia utilizzata e considera trattamento qualsiasi operazione come raccolta, registrazione, organizzazione, conservazione, consultazione, uso o diffusione di dati personali. I principi richiamati dalla Commissione europea includono liceità, correttezza, trasparenza, limitazione delle finalità, minimizzazione, accuratezza, integrità, riservatezza e accountability.

Anche il quadro europeo sull'IA è rilevante. L'AI Act introduce un approccio basato sul rischio e mira a promuovere sistemi di IA affidabili, sicuri, rispettosi dei diritti fondamentali e centrati sull'essere umano. Nel settore della salute pubblica, la guida dell'OMS sull'etica e la governance dell'IA richiama la necessità di mettere etica e diritti umani al centro della progettazione, dello sviluppo e dell'uso dei sistemi intelligenti.

Questi principi hanno implicazioni pratiche: nessun sistema di IA dovrebbe produrre decisioni automatizzate stigmatizzanti; l'uso dei dati dei minori deve essere sottoposto a criteri rafforzati; gli output algoritmici devono essere spiegabili e verificabili; l'intervento umano deve restare essenziale; ogni progetto dovrebbe prevedere una valutazione d'impatto etico e giuridico.

Data donation e screen tracking: possibilità e cautela

La data donation e lo screen tracking sono spesso proposti come metodi per superare i limiti dei dati social pubblici. La data donation prevede che i partecipanti condividano volontariamente alcuni dati digitali con i ricercatori, secondo protocolli di consenso informato. Lo screen tracking può invece monitorare nel tempo l'uso di dispositivi o applicazioni, offrendo dati più longitudinali rispetto alla semplice raccolta via API.

Questi strumenti sono promettenti, ma non semplici. La stessa presentazione della review indica che i dati disponibili sono per lo più trasversali, raccolti tramite API tradizionali, e che data donation e screen tracking richiedono molte più risorse e tempo. Inoltre, pongono problemi elevati di privacy, proporzionalità e sicurezza. In particolare, non è accettabile parlare genericamente di "donare tutti i propri dati social": un protocollo corretto deve prevedere selezione dei dati necessari, consenso granulare, possibilità di revoca, protezione dei contenuti sensibili e controllo indipendente.

Un modello operativo realistico

Una strategia di prevenzione che integri IA e social media dovrebbe essere costruita su cinque livelli.

Primo: ascolto digitale strutturato. Le istituzioni possono monitorare trend pubblici, linguaggi emergenti, nuove sostanze, prodotti di nicotina, pattern territoriali e percezioni del rischio, sempre con supervisione umana e nel rispetto della normativa.

Secondo: co-progettazione con i destinatari. I messaggi devono essere costruiti con adolescenti, giovani adulti, educatori e operatori, evitando comunicazioni calate dall'alto. La credibilità è una variabile preventiva.

Terzo: personalizzazione responsabile. L'IA può aiutare a differenziare contenuti per piattaforma, età, linguaggio e contesto, ma ogni messaggio deve essere validato da esperti e non deve sfruttare vulnerabilità individuali.

Quarto: valutazione multilivello. Le campagne devono essere misurate non solo in termini di visualizzazioni o engagement, ma anche rispetto a conoscenze, percezioni, intenzioni, accesso ai servizi e possibili cambiamenti comportamentali.

Quinto: governance etica e giuridica. Ogni progetto deve prevedere protezione dei dati, trasparenza, documentazione delle decisioni, verifica dei bias, supervisione umana e tutela specifica dei minori.

Conclusione

L'IA e i social media possono aprire nuove opportunità per la prevenzione dell'uso di sostanze, ma solo se inseriti in una strategia di salute pubblica rigorosa. Il loro valore non sta nella promessa di sostituire educatori, clinici, operatori o comunità, ma nella capacità di rafforzare alcune funzioni: ascoltare segnali deboli, leggere trend emergenti, adattare i messaggi, valutare campagne e orientare interventi più tempestivi.

L'evidenza disponibile resta però preliminare. La letteratura mostra buone potenzialità nella sorveglianza digitale e nella predizione di pattern di rischio, ma non dimostra ancora in modo robusto che l'IA riduca l'iniziazione all'uso di sostanze o i comportamenti problematici. Per questo servono studi longitudinali, interventi controllati, outcome comportamentali e protocolli etici solidi.

La prevenzione digitale del futuro non sarà semplicemente più tecnologica. Dovrà essere più precisa, più partecipativa, più trasparente e più responsabile. L'IA può diventare una componente utile di questa evoluzione, ma solo se resta al servizio della competenza umana, della tutela dei minori e della salute pubblica.

Tabella delle fonti aggiunte

| Autore/Anno | Titolo | Rivista/Fonte | Impact Factor | Contributo al testo |
|--|---|--------------------------------------|--|---|
| Nguyen et al., 2026 | Artificial intelligence innovations in substance use prevention on social media: A scoping review | Public Health | Non verificato in JCR in questa sessione | Fonte scientifica principale; definisce stato dell'arte, applicazioni, limiti e prospettive. |
| ISSUP/Hanoi Medical University/We Care, 2026 | AI & Substance Use: Research Sharing | Presentazione scientifica webinar | Non applicabile | Dati metodologici: 1.418 record, 15 studi inclusi, piattaforme, modelli, limiti tecnici ed etici. |
| WHO Europe/HBSC, 2024 | Teens, screens and mental health | Fonte istituzionale WHO | Non applicabile | Contesto epidemiologico su uso problematico dei social media negli adolescenti. |
| ESPAD/EUDA, 2025 | ESPAD Report 2024 | Fonte istituzionale europea | Non applicabile | Dati europei su adolescenti, sostanze, e-cigarette, benessere mentale e prevenzione. |
| Hashemi & Vogel, 2024 | Adolescents' perceptions of substance use messaging in the age of social media | Health Education Research | Non verificato in JCR in questa sessione | Mostra la dissonanza tra contenuti dei pari e messaggi preventivi formali. |
| Evans et al., 2017 | The Living the Example Social Media Substance Use Prevention Program | JMIR Mental Health | Non verificato in JCR in questa sessione | Esempio di programma peer-to-peer di prevenzione tramite social media. |
| Hassanpour et al., 2019 | Identifying substance use risk based on deep neural networks and Instagram social media data | Neuropsychopharmacology | Non verificato in JCR in questa sessione | Esempio di deep learning applicato a dati Instagram per stima del rischio. |
| Fisher et al., 2023 | Automating Detection of Drug-Related Harms on Social Media | Journal of Medical Internet Research | Non verificato in JCR in questa sessione | Esempio di IA per early warning e monitoraggio dei danni droga-correlati. |
| WHO, 2021 | Ethics and governance of artificial intelligence for health | Fonte istituzionale WHO | Non applicabile | Cornice etica: diritti umani, governance, responsabilità e supervisione. |
| Commissione europea, 2024-2026 | AI Act e GDPR | Fonte istituzionale UE | Non applicabile | Cornice normativa europea su dati personali, IA affidabile e approccio basato sul rischio. |

Link cliccabili di approfondimento

- Articolo principale su IA, social media e prevenzione dell'uso di sostanze: <https://www.issup.net/knowledge-share/resources/2026-03/artificial-intelligence-innovations-substance-use-prevention>
- Presentazione PDF del webinar ISSUP: https://www.issup.net/files/2026-05/AI%26SU_SR_Researching%20Sharing.pdf
- WHO Europe su adolescenti, schermi e salute mentale: <https://www.who.int/europe/news/item/25-09-2024-teens--screens-and-mental-health>
- ESPAD Report 2024 / EUDA: https://www.euda.europa.eu/publications/joint-publications/espas-report-2024_en
- Dati chiave ESPAD 2024 su sostanze ed e-cigarette: https://www.euda.europa.eu/publications/data-factsheets/espas-2024-key-findings_en
- Studio qualitativo su adolescenti e messaggi preventivi nei social media: <https://academic.oup.com/her/article-abstract/39/1/1/7512621>
- Programma peer-to-peer "Living the Example": <https://mental.imir.org/2017/2/e24/>
- Studio su Instagram e deep learning per rischio di uso di sostanze: <https://www.nature.com/articles/s41386-018-0247-x>
- IA e early warning sui danni droga-correlati: <https://www.imir.org/2023/1/e43630/>
- AI Act europeo: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
- GDPR e protezione dei dati personali: https://commission.europa.eu/law/law-topic/data-protection/data-protection-explained_en
- OMS: etica e governance dell'IA in salute: <https://www.who.int/publications/i/item/9789240029200>