Come usare l'intelligenza artificiale nelle indagini OSINT: guida pratica, workflow e compliance

Maria Cattini | 02/02/2026 | Intelligenza Artificiale

Hai mai passato ore a setacciare centinaia di PDF, thread infiniti su forum oscuri e fughe di dati incomprensibili? Se fai indagini OSINT, il problema non è trovare informazioni, ma capire quali contano davvero e farlo abbastanza in fretta da essere utile.

Questa guida pratica mostra come integrare l'intelligenza artificiale nel flusso di lavoro OSINT: strumenti concreti (anche a basso costo), template di prompt riutilizzabili, un workflow operativo riproducibile e le regole di compliance per mantenere la difendibilità legale degli output. Nessuna teoria inutile: istruzioni testate sul campo e risorse per implementare subito.

Cosa fa davvero l'IA per un investigatore OSINT

Prima di buttarti su ChatGPT aspettandoti magie, capiamo cosa l'IA sa fare bene e cosa no.

• Elaborazione massiva di testo. Un LLM può leggere e sintetizzare 50 pagine di documenti in secondi. Tu impiegheresti ore, se non giorni.
• Riconoscimento di pattern. L'IA individua contraddizioni, ripetizioni sospette e anomalie che a occhio umano passerebbero inosservate. Nei dataset complessi, questo vale oro.
• Estrazione dati strutturati. Dare all'IA un elenco disordinato e ricevere una tabella pulita con nomi, date, luoghi e ruoli? Fattibile, se sai come chiederlo.
• Traduzione contestuale. Non si limita a tradurre parola per parola: cerca di mantenere il senso del testo originale. Certo, per lingue complesse serve sempre una verifica umana.
• Generazione di ipotesi. Bloccato su un caso? L'IA può suggerirti angolazioni che non avevi considerato, partendo dai dati che già possiedi.

Tuttavia, l'IA non sostituisce gli strumenti OSINT tradizionali né scopre dati dal nulla: non accede a banche dati riservate e non crea prove vere. L'IA è un assistente che aiuta a dare senso ai dati raccolti con Maltego, Google Dorking, OSINT Framework e simili.

Prompting: il linguaggio dell'IA e come usarlo bene

Un prompt è l'insieme di istruzioni che dai all'IA. Se le istruzioni sono confuse, il risultato sarà confuso.

Anatomia di un buon prompt per OSINT

1. Il ruolo. Dici all'IA come comportarsi. Esempio: "Agisci come un analista OSINT esperto."
2. Il compito. Cosa deve fare. Esempio: "Estrai tutti i nomi di persona, le date e le località menzionate in questo documento."
3. Le regole. Cosa può e non può fare. Esempio: "Cita sempre la fonte. Se un'informazione non è chiara, scrivi 'sconosciuto' invece di indovinare."
4. Il formato. Come vuoi i risultati. Esempio: "Presenta i dati in una tabella con tre colonne: Nome, Data, Località."

Prompting avanzato e prompt-chaining per indagini OSINT

Di seguito tecniche e template avanzati pronti all'uso, con esempi di input/output attesi e indicazioni su come misurare la qualità degli output.

1) Estrazione entità con citazione fonti (template)

Prompt: "Agisci come analista OSINT. Estrai tutte le entità PERSONA, ORGANIZZAZIONE, DATA e LUOGO dal testo seguente. Per ogni entità fornisci: entità, tipo, fonte (URL o documento), timestamp della fonte, e un breve commento di confidenza (alta/media/bassa). Se non trovi fonte verificabile, scrivi 'SCONOSCIUTO' nella colonna fonte."

2) Timeline generator

Prompt: "Ricevi in input questo set di post, email e documenti. Crea una timeline cronologica con eventi numerati, indicandone la fonte (URL/file), la data (standard ISO) e la prova in una riga. Ordina per data e segnala eventuali gap temporali o incongruenze."

3) Verifica incrociata

Prompt: "Data l'affermazione X, cerca e riporta fino a 3 fonti primarie che la confermano o la smentiscono. Per ogni fonte indica: tipo (registro ufficiale, articolo, archive.org), URL, data di accesso e un giudizio di affidabilità (1-5). Se non trovi fonti primarie, rispondi 'SCONOSCIUTO' e suggerisci i termini di ricerca da usare."

4) Analisi stilometrica rapida

Prompt: "Confronta il testo A e il testo B. Identifica differenze stilistiche significative (lessico, lunghezza frasi, uso di parole chiave) e segnala le probabilità che provengano dalla stessa fonte autore usando indicatori qualitativi. Fornisci esempi di frasi che evidenziano la differenza."

5) Prompt di fallback per ridurre allucinazioni

Prompt: "Se durante l'analisi non trovi una fonte verificabile per un'affermazione, rispondi esclusivamente 'SCONOSCIUTO' e fornisci 3 termini di ricerca mirati per ulteriori verifiche manuali."

Concetto di prompt chaining (esempio pratico)

Esempio: Step A - estrai entità da 1000 documenti; Step B - normalizza nomi/varianti; Step C - costruisci grafo delle relazioni; Step D - calcola centralità e segnala i nodi con score alto. Ogni step è una chiamata separata al modello con output salvato e hash per tracciabilità.

Metriche di qualità suggerite: valutare precision e recall su un campione verificato di 100 estrazioni; target operativo: precision ≥ 90% per entità critiche, recall target ≥ 75% con revisione umana per gli elementi mancanti.

Gli strumenti IA che funzionano per l'OSINT

Non esiste un unico strumento che faccia tutto. Qui sotto trovi i tool principali, i loro punti di forza, limiti e quando usarli.

ChatGPT e Claude: analisti generici

Questi modelli linguistici sono il cervello extra della tua indagine. Lavorano con testo e documenti, e se ben istruiti diventano co-investigatori efficaci.

• Cosa fanno meglio: Riassumere documenti di 100+ pagine; estrarre dati strutturati; individuare

incongruenze; generare report formattati.
• Limiti: non sempre accesso realtime e possono allucinare se non guidati; preferire DPA e soluzioni on-prem per dati sensibili.

Maltego e visualizzatori OSINT

Questi strumenti mappano le relazioni tra dati e visualizzano reti di persone, aziende, domini.

Hunchly: archivista automatico

Hunchly registra ogni pagina web che visiti durante un'indagine e conserva URL, timestamp e screenshot in modo ricercabile.

Elastic AI e Haystack: processori di massa

Quando hai migliaia di documenti o leak giganteschi servono processori su larga scala per indicizzazione e clustering.

Tabella comparativa strumenti AI per OSINT: criteri e raccomandazioni

Strumento	Tipologia	Costo indicativo	Privacy	Accesso realtime	Migliore per	Limiti principali	Raccomandazione pratica
ChatGPT / GPT-4	LLM generalista	Low-High (piani)	Cloud	Limitato	Riassunti, estrazione entità	Tendenza ad allucinare; privacy	Usare per pre-analisi testuale; non inviare dati sensibili senza DPA
Claude	LLM generalista	Medium-High	Cloud	Limitato	Documenti lunghi, analisi contestuale	Costi e privacy	Usare con DPA e logging delle richieste
Maltego	Visualizzazione grafica	Licenza	On-prem / Cloud	No	Network mapping	Curva di apprendimento	Usare per mapping relazioni e validazione grafica
Hunchly	Archivio forense	Licenza	On-prem / Cloud	No	Logging e catena di custodia	Non analizza contenuti	Indispensabile per evidenze forensi
Elastic AI / Haystack	Indexing & Retrieval	Open / Enterprise	On-prem possibile	Sì (configurato)	Processing di grandi corpora	Setup complesso	Usare per leak massivi; preferire on-prem
TinEye / Google Lens	Reverse image	Freemium	Cloud	Sì	Verifica immagini	Dipendenza da database	Combinare con analisi metadati locali
BuiltWith	Profiling web	Freemium	Cloud	No	Mappare stack tecnologico	Non fornisce TTP	Usare per superficie d'attacco e profilo sito

Consiglio per il budget: per team individuali: LLM + Hunchly; per team enterprise: Elastic AI on-prem + Maltego e Hunchly.

Workflow operativo passo-passo: dall'acquisizione alla verifica con IA

Workflow riproducibile con checkpoint forensi e template pronti:

1. Definizione obiettivo e limiti legali (15-30 min) • Definisci scopo, scope e vincoli legali (es. dati personali, limiti territoriali).• Checklist: autorizzazioni, NDA, ruolo investigatore.
2. Raccolta dati iniziale (30-180 min) • Strumenti: Google Dorking, Maltego, query booleane, scraping controllato.• Log sessione: salva URL visitati, query usate, timestamp (uso Hunchly).• Pre-processing per l'IA (30-120 min) Pulisci, normalizza formati, anonimizza dati sensibili prima dell'invio a modelli cloud.• Esempio comando: estrai testo da PDF → rimuovi PII → CSV con colonna origine.• Analisi LLM divisa in micro-task (variabile) Micro-task tipici: estrazione entità, normalizzazione, clustering, timeline generation.• Esempio prompt per estrazione entità: vedi template nella sezione Prompting avanzato.• Checkpoint: salva output con hash e timestamp; non procedere senza verificare entità critiche.• Cross-check automatizzato e manuale (30-240 min) Attività: reverse image search, domain age, verifica registri ufficiali (es. registri di imprese).• Regola: per ogni affermazione rilevante richiedi almeno una fonte primaria verificabile.• Visualizzazione relazioni e validazione Tool: Maltego, neo4j. Valida collegamenti sospetti con evidenze multiple.• Conservazione e catena di custodia Registra hash di input/output, timestamp, screenshot (Hunchly); archiviare in repository cifrato.• Redazione report verificabile Template obbligatorio: sintesi esecutiva; metodologia; fonti; output AI con timestamp e hash; valutazione confidenza; appendici tecniche.• Sintesi esecutiva (2-4 paragrafi)• Obiettivo e scope• Metodologia e strumenti (includere versioni e endpoint dei modelli)• Fonti primarie e secondarie (URL, data di accesso, screenshot/hash)• Output AI salvati (file, timestamp, hash)• Valutazione confidenza e rischi• Raccomandazioni operative• Appendici tecniche (prompt usati, comandi, log)• Base giuridica: definire la base per il trattamento (es. interesse legittimo, consenso, obbligo legale). Documentare decisione e valutazione interessi.• Minimizzazione dei dati: raccogli solo i dati necessari per lo scopo investigativo; pseudonimizza o anonimizza prima di inviare a modelli cloud.• Limitazioni al profiling: evitare decisioni automatizzate che producono effetti legali senza revisione umana.• Conservazione: definire retention policy minima e sicura; cancellazione sicura quando scadono i tempi. Esempi pratici e regole operative • Anonymize: rimuovi PII (codici fiscali, numeri identificativi, UBO) prima di inviare dataset a modelli cloud.• Processare localmente: dati sensibili e UBO, numeri identificativi, dossier legali. Se necessario usare modelli self-hosted o on-prem.• NDA e autorizzazioni: firmare NDA e DPA con fornitori cloud quando si processano dati non pubblici. Responsabilità professionale e catena di custodia Documentare ogni passaggio per preservare la catena di custodia digitale: • Registrare hash e timestamp di input e output.• Salvare prompt completi, versione del modello e endpoint usato.• Certificare che decisioni investigative rilevanti siano supportate da fonti primarie verificate e da revisione umana. Checklist di conformità operativa (breve) Non inviare dati sensibili a modelli cloud senza DPA.
3. Loggare tutte le richieste AI e conservare gli output con hash.
4. Valutare accordi DPA e responsabilità contrattuali con fornitori.
5. Applicare retention policy minima e cancellazione sicura.

Risorse e riferimenti

Documenti utili: Regolamento (UE) 2016/679 (GDPR), linee guida ENISA su digital forensics e best practice DPA dei fornitori cloud. Formati di logging consigliati: JSONL con campi (timestamp, user-id, prompt, model-version, input-hash, output-hash, source-ids).

Esempio di modulo interno per condivisione file con servizi esterni

Include: descrizione file, tipologia dati (PII/Sensibile/Non sensibile), motivo condivisione, provider esterno, DPA firmato (Sì/No), data invio, responsabile autorizzazione.

Caso pratico: smascherare un profilo falso con l'IA

Ricevi un messaggio LinkedIn da una persona che si presenta come “Astra Velorin, Ambasciatrice del Braccio Spirale Esterno”. Ti offre un lavoro come “Assistente ai Rapimenti”. Alieno vero o truffatore creativo? Usiamo l’IA per scoprirlo.

1. Analizzare la bio — copi la biografia di Astra e la incolli in ChatGPT con questo prompt...
2. Generare pivot investigativi — Chiedi a ChatGPT: “Mostrami tre controlli di follow-up...”
3. Analizzare documenti allegati — Astra ti invia un PDF di 40 pagine chiamato “Proposta di Primo Contatto”...
4. Ricerca inversa delle immagini — Usi Google Lens o TinEye...
5. Report finale — Con Hunchly hai documentato ogni passaggio...

Nel report finale includi: URL/screenshot con hash, prompt usati (versionati), risultati LLM con timestamp, valutazione confidenza e raccomandazioni operative. Di conseguenza la conclusione è documentata e riproducibile.

Metriche e KPI per misurare l'efficacia dell'IA in OSINT

KPI concreti, formule e frequenze di misura:

• Tempo medio per analisi (T): T = tempo umano medio / tempo con IA. Target: riduzione percentuale > 70% per task ripetitivi.
• Precision of entity extraction (P): P = entità corrette / entità estratte. Soglia di allerta: P • Recall su fonti verificate (R): R = fonti rilevanti trovate / fonti note esistenti. Target operativo: R ≥ 0.75 per corpus non triviale.
• Confidence-adjusted accuracy (CAA): percentuale di output con fonte verificata e confidenza alta.
• Reduction in manual hours per caso (H): ore risparmiate = ore umane pre-IA - ore con IA.

Esempio numerico (simulato): su 10.000 documenti, Elastic AI + LLM ha ridotto il triage da 120 a 8 ore; precision estrazione entità = 92%; recall fonti critiche = 78%.

Reportistica suggerita: dashboard Grafana con metriche di job (durata job, tassi di errore, percentuale di output verificati, latenza). Frequenza report: settimanale per operazioni continue; post-mortem dopo ogni indagine significativa.

Errori da evitare (e come risolverli)

• Fidarsi ciecamente dell’IA. L’IA può inventare fonti, nomi, date. Verifica sempre i risultati critici con fonti primarie.
• Prompt troppo vaghi. “Analizza questo documento” non è un prompt. Specifica cosa cerchi: nomi, date, contraddizioni, pattern.
• Non iterare. Il primo prompt raramente è perfetto. Testa, modifica, riprova. Salva i prompt che funzionano.
• Usare un solo strumento. ChatGPT non può fare tutto. Combina LLM, visualizzatori, archivisti e processori di massa.
• Dimenticare la catena di custodia. Se lavori su casi legali, ogni passaggio deve essere documentato. Usa strumenti come Hunchly per mantenere tracce verificabili.

Checklist forensica: catena di custodia e verifica quando si usa IA

1. Log completo delle sessioni: timestamp, URL, input testuali inviati all'IA e output ricevuti; salva hash di input e output.
2. Anonymizzazione preventiva: rimuovi o pseudonimizza dati sensibili prima di inviarli a modelli cloud; conserva copia originale cifrata separata.
3. Archiviazione forense: usa Hunchly o imaging forense per screenshot, HTML capture e metadati; associa ogni elemento a un ID investigativo.

4. Verifica delle fonti: per ogni affermazione critica prodotta dall'IA, richiedi almeno una fonte primaria verificabile e registra URL, data di accesso e screenshot.
5. Conservazione dei prompt: salva la versione completa dei prompt usati (istruzioni, parametri) per riproducibilità e auditing.
6. Controllo versioning modelli: registra quale modello/endpoint è stato usato (versione, provider) e note su aggiornamenti che potrebbero alterare output.
7. Escalation procedure: definisci quando un output richiede revisione umana o consulenza legale (es. potenziali implicazioni penali).

FAQ ottimizzate per featured snippets

L'IA può sostituire un investigatore OSINT?

No. L'IA accelera l'analisi e automatizza compiti ripetitivi, ma non sostituisce giudizio investigativo, verifica delle fonti e valutazione legale.

Posso inviare dati sensibili a ChatGPT per un'indagine?

No. Evita di inviare dati sensibili o personali a modelli cloud a meno che non ci sia un DPA e misure di anonimizzazione; preferisci soluzioni on-premise per dati riservati.

Come riduco le allucinazioni dell'IA?

Usa prompt che richiedono fonti, specifica "SCONOSCIUTO" se non trovi fonti, spezzetta il compito in più passaggi e verifica tutte le informazioni critiche con fonti primarie.

Quali strumenti usare per mantenere la catena di custodia?

Usa soluzioni di archiviazione forense (Hunchly, imaging forense) che registrano URL, screenshot e hash, affiancate da logging delle richieste AI e conservazione delle versioni dei prompt.

Qual è il miglior approccio per investigare immagini sospette?

Combina reverse image search (TinEye / Google Lens), analisi EXIF metadati, ricerca per similitudine e confronto con banche immagini; documenta ogni passaggio con screenshot e hash.

Quanto costa integrare IA in OSINT?

Dipende dal volume dati e dalla soluzione: piani consumer per LLM partono da decine di euro al mese; per progetti enterprise servono licenze, infrastruttura on-prem e personale specializzato.

Prossimi passi

L'intelligenza artificiale può ridurre significativamente il tempo speso in attività ripetitive e aumentare la qualità delle analisi OSINT, purché venga integrata con processi forensi, controllo umano e rispetto della normativa. Segui il workflow, applica la checklist forense e misura i KPI per dimostrare valore e conformità.

Vuoi iniziare subito? Prendi un caso OSINT su cui stai lavorando. Scegli uno strumento da questa lista. Scrivi un prompt seguendo la struttura che ti ho mostrato. Testa, itera, affina. In poche ore avrai integrato l’IA nel tuo flusso di lavoro.

Risorse e community:

- Newsletter: https://coondivido.substack.com/
- Telegram: https://t.me/osintaipertutti
- Telegram: https://t.me/osintprojectgroup

Ulteriori riferimenti

Per approfondire compliance e pratiche forensi consultare GDPR, le linee guida ENISA e la documentazione ufficiale dei tool citati; per template di logging usare JSONL con i campi consigliati nella sezione 'Quadro normativo'.

Hai mai passato ore a setacciare centinaia di PDF, thread infiniti su forum oscuri e fughe di dati incomprensibili? Se fai indagini OSINT, il problema non è trovare informazioni, ma capire quali contano davvero e farlo abbastanza in fretta da essere utile.

Questa guida pratica mostra come integrare l'intelligenza artificiale nel flusso di lavoro OSINT: strumenti concreti (anche a basso costo), template di prompt riutilizzabili, un workflow operativo riproducibile e le regole di compliance per mantenere la difendibilità legale degli output. Nessuna teoria inutile: istruzioni testate sul campo e risorse per implementare subito.

Cosa fa davvero l'IA per un investigatore OSINT

Prima di buttarti su ChatGPT aspettandoti magie, capiamo cosa l’IA sa fare bene e cosa no.

• Elaborazione massiva di testo. Un LLM può leggere e sintetizzare 50 pagine di documenti in secondi. Tu impiegheresti ore, se non giorni.
• Riconoscimento di pattern. L’IA individua contraddizioni, ripetizioni sospette e anomalie che a occhio umano passerebbero inosservate. Nei dataset complessi, questo vale oro.
• Estrazione dati strutturati. Dare all’IA un elenco disordinato e ricevere una tabella pulita con nomi, date, luoghi e ruoli? Fattibile, se sai come chiederlo.
• Traduzione contestuale. Non si limita a tradurre parola per parola: cerca di mantenere il senso del testo originale. Certo, per lingue complesse serve sempre una verifica umana.
• Generazione di ipotesi. Bloccato su un caso? L’IA può suggerirti angolazioni che non avevi considerato, partendo dai dati che già possiedi.

Tuttavia, l'IA non sostituisce gli strumenti OSINT tradizionali né scopre dati dal nulla: non accede a banche dati riservate e non crea prove vere. L'IA è un assistente che aiuta a dare senso ai dati raccolti con Maltego, Google Dorking, OSINT Framework e simili.

Prompting: il linguaggio dell'IA e come usarlo bene

Un prompt è l’insieme di istruzioni che dai all’IA. Se le istruzioni sono confuse, il risultato sarà confuso.

Anatomia di un buon prompt per OSINT

1. Il ruolo. Dici all’IA come comportarsi. Esempio: “Agisci come un analista OSINT esperto.”
2. Il compito. Cosa deve fare. Esempio: “Estrai tutti i nomi di persona, le date e le località menzionate in questo documento.”
3. Le regole. Cosa può e non può fare. Esempio: “Cita sempre la fonte. Se un’informazione non è chiara, scrivi ‘sconosciuto’ invece di indovinare.”
4. Il formato. Come vuoi i risultati. Esempio: “Presenta i dati in una tabella con tre colonne: Nome, Data, Località.”

Prompting avanzato e prompt-chaining per indagini OSINT

Di seguito tecniche e template avanzati pronti all'uso, con esempi di input/output attesi e indicazioni su come misurare la qualità degli output.

1) Estrazione entità con citazione fonti (template)

Prompt: "Agisci come analista OSINT. Estrai tutte le entità PERSONA, ORGANIZZAZIONE, DATA e LUOGO dal testo seguente. Per ogni entità fornisci: entità, tipo, fonte (URL o documento), timestamp della fonte, e un breve commento di confidenza (alta/media/bassa). Se non trovi fonte verificabile, scrivi 'SCONOSCIUTO' nella colonna fonte."

2) Timeline generator

Prompt: "Ricevi in input questo set di post, email e documenti. Crea una timeline cronologica con eventi numerati, indicandone la fonte (URL/file), la data (standard ISO) e la prova in una riga. Ordina per data e segnala eventuali gap temporali o incongruenze."

3) Verifica incrociata

Prompt: "Data l'affermazione X, cerca e riporta fino a 3 fonti primarie che la confermano o la smentiscono. Per ogni fonte indica: tipo (registro ufficiale, articolo, archive.org), URL, data di accesso e un giudizio di affidabilità (1-5). Se non trovi fonti primarie, rispondi 'SCONOSCIUTO' e suggerisci i termini di ricerca da usare."

4) Analisi stilometrica rapida

Prompt: "Confronta il testo A e il testo B. Identifica differenze stilistiche significative (lessico, lunghezza frasi, uso di parole chiave) e segnala le probabilità che provengano dalla stessa fonte autore usando indicatori qualitativi. Fornisci esempi di frasi che evidenziano la differenza."

5) Prompt di fallback per ridurre allucinazioni

Prompt: "Se durante l'analisi non trovi una fonte verificabile per un'affermazione, rispondi esclusivamente 'SCONOSCIUTO' e fornisci 3 termini di ricerca mirati per ulteriori verifiche manuali."

Concetto di prompt chaining (esempio pratico)

Esempio: Step A - estrai entità da 1000 documenti; Step B - normalizza nomi/varianti; Step C - costruisci grafo delle relazioni; Step D - calcola centralità e segnala i nodi con score alto. Ogni step è una chiamata separata al modello con output salvato e hash per tracciabilità.

Metriche di qualità suggerite: valutare precision e recall su un campione verificato di 100 estrazioni; target operativo: precision ≥ 90% per entità critiche, recall target ≥ 75% con revisione umana per gli elementi mancanti.

Gli strumenti IA che funzionano per l'OSINT

Non esiste un unico strumento che faccia tutto. Qui sotto trovi i tool principali, i loro punti di forza, limiti e quando usarli.

ChatGPT e Claude: analisti generici

Questi modelli linguistici sono il cervello extra della tua indagine. Lavorano con testo e documenti, e se ben istruiti diventano co-investigatori efficaci.

• Cosa fanno meglio: Riassumere documenti di 100+ pagine; estrarre dati strutturati; individuare incongruenze; generare report formattati.
• Limiti: non sempre accesso realtime e possono allucinare se non guidati; preferire DPA e soluzioni on-prem per dati sensibili.

Maltego e visualizzatori OSINT

Questi strumenti mappano le relazioni tra dati e visualizzano reti di persone, aziende, domini.

Hunchly: archivista automatico

Hunchly registra ogni pagina web che visiti durante un’indagine e conserva URL, timestamp e screenshot in modo ricercabile.

Elastic AI e Haystack: processori di massa

Quando hai migliaia di documenti o leak giganteschi servono processori su larga scala per indicizzazione e clustering.

Tabella comparativa strumenti AI per OSINT: criteri e raccomandazioni

Strumento	Tipologia	Costo indicativo	Privacy	Accesso realtime	Migliore per	Limiti principali	Raccomandazione pratica
ChatGPT / GPT-4	LLM generalista	Low-High (piani)	Cloud	Limitato	Riassunti, estrazione entità	Tendenza ad allucinare; privacy	Usare per pre-analisi testuale; non inviare dati sensibili senza DPA
Claude	LLM generalista	Medium-High	Cloud	Limitato	Documenti lunghi, analisi contestuale	Costi e privacy	Usare con DPA e logging delle richieste
Maltego	Visualizzazione grafica	Licenza	On-prem / Cloud	No	Network mapping	Curva di apprendimento	Usare per mapping relazioni e validazione grafica
Hunchly	Archivio forense	Licenza	On-prem / Cloud	No	Logging e catena di custodia	Non analizza contenuti	Indispensabile per evidenze forensi
Elastic AI / Haystack	Indexing & Retrieval	Open / Enterprise	On-prem possibile	Sì (configurato)	Processing di grandi corpora	Setup complesso	Usare per leak massivi; preferire on-prem
TinEye / Google Lens	Reverse image	Freemium	Cloud	Sì	Verifica immagini	Dipendenza da database	Combinare con analisi metadati locali
BuiltWith	Profiling web	Freemium	Cloud	No	Mappare stack tecnologico	Non fornisce TTP	Usare per superficie d'attacco e profilo sito

Consiglio per il budget: per team individuali: LLM + Hunchly; per team enterprise: Elastic AI on-prem + Maltego e Hunchly.

Workflow operativo passo-passo: dall'acquisizione alla verifica con IA

Workflow riproducibile con checkpoint forensi e template pronti:

1. Definizione obiettivo e limiti legali (15-30 min) • Definisci scopo, scope e vincoli legali (es. dati personali, limiti territoriali).• Checklist: autorizzazioni, NDA, ruolo investigatore.
2. Raccolta dati iniziale (30-180 min) • Strumenti: Google Dorking, Maltego, query booleane, scraping controllato.• Log sessione: salva URL visitati, query usate, timestamp (uso Hunchly).• Pre-processing per l'IA (30-120 min) Pulisci, normalizza formati, anonimizza dati sensibili prima dell'invio a modelli cloud.• Esempio comando: estrai testo da PDF → rimuovi PII → CSV con colonna origine.• Analisi LLM divisa in micro-task (variabile) Micro-task tipici: estrazione entità, normalizzazione, clustering, timeline generation.• Esempio prompt per estrazione entità: vedi template nella sezione Prompting avanzato.• Checkpoint: salva output con hash e timestamp; non procedere senza verificare entità critiche.• Cross-check automatizzato e manuale (30-240 min) Attività: reverse image search, domain age, verifica registri ufficiali (es. registri di imprese).• Regola: per ogni affermazione rilevante richiedi almeno una fonte primaria verificabile.• Visualizzazione relazioni e validazione Tool: Maltego, neo4j. Valida collegamenti sospetti con evidenze multiple.• Conservazione e catena di custodia Registra hash di input/output, timestamp, screenshot (Hunchly); archiviare in repository cifrato.• Redazione report verificabile Template obbligatorio: sintesi esecutiva; metodologia; fonti; output AI con timestamp e hash; valutazione confidenza; appendici tecniche.• Sintesi esecutiva (2-4 paragrafi)• Obiettivo e scope• Metodologia e strumenti (includere versioni e endpoint dei modelli)• Fonti primarie e secondarie (URL, data di accesso, screenshot/hash)• Output AI salvati (file, timestamp, hash)• Valutazione confidenza e rischi• Raccomandazioni operative• Appendici tecniche (prompt usati, comandi, log)• Base giuridica: definire la base per il trattamento (es. interesse legittimo, consenso, obbligo legale). Documentare decisione e valutazione interessi.• Minimizzazione dei dati: raccogli solo i dati necessari per lo scopo investigativo; pseudonimizza o anonimizza prima di inviare a modelli cloud.• Limitazioni al profiling: evitare decisioni automatizzate che producono effetti legali senza revisione umana.• Conservazione: definire retention policy minima e sicura; cancellazione sicura quando scadono i tempi. Esempi pratici e regole operative • Anonymize: rimuovi PII (codici fiscali, numeri identificativi, UBO) prima di inviare dataset a modelli cloud.• Processare localmente: dati sensibili e UBO, numeri identificativi, dossier legali. Se necessario usare modelli self-hosted o on-prem.• NDA e autorizzazioni: firmare NDA e DPA con fornitori cloud quando si processano dati non pubblici. Responsabilità professionale e catena di custodia Documentare ogni passaggio per preservare la catena di custodia digitale: • Registrare hash e timestamp di input e output.• Salvare prompt completi, versione del modello e endpoint usato.• Certificare che decisioni investigative rilevanti siano supportate da fonti primarie verificate e da revisione umana. Checklist di conformità operativa (breve) Non inviare dati sensibili a modelli cloud senza DPA.
3. Loggare tutte le richieste AI e conservare gli output con hash.
4. Valutare accordi DPA e responsabilità contrattuali con fornitori.
5. Applicare retention policy minima e cancellazione sicura.

Risorse e riferimenti

Documenti utili: Regolamento (UE) 2016/679 (GDPR), linee guida ENISA su digital forensics e best practice DPA dei fornitori cloud. Formati di logging consigliati: JSONL con campi (timestamp, user-id, prompt, model-version, input-hash, output-hash, source-ids).

Esempio di modulo interno per condivisione file con servizi esterni

Include: descrizione file, tipologia dati (PII/Sensibile/Non sensibile), motivo condivisione, provider esterno, DPA firmato (Sì/No), data invio, responsabile autorizzazione.

Caso pratico: smascherare un profilo falso con l'IA

Ricevi un messaggio LinkedIn da una persona che si presenta come “Astra Velorin, Ambasciatrice del Braccio Spirale Esterno”. Ti offre un lavoro come “Assistente ai Rapimenti”. Alieno vero o truffatore creativo? Usiamo l’IA per scoprirlo.

1. Analizzare la bio — copi la biografia di Astra e la incolli in ChatGPT con questo prompt...
2. Generare pivot investigativi — Chiedi a ChatGPT: “Mostrami tre controlli di follow-up...”
3. Analizzare documenti allegati — Astra ti invia un PDF di 40 pagine chiamato “Proposta di Primo

Contatto”...
4. Ricerca inversa delle immagini — Usi Google Lens o TinEye...
5. Report finale — Con Hunchly hai documentato ogni passaggio...

Nel report finale includi: URL/screenshot con hash, prompt usati (versionati), risultati LLM con timestamp, valutazione confidenza e raccomandazioni operative. Di conseguenza la conclusione è documentata e riproducibile.

Metriche e KPI per misurare l'efficacia dell'IA in OSINT

KPI concreti, formule e frequenze di misura:

• Tempo medio per analisi (T): T = tempo umano medio / tempo con IA. Target: riduzione percentuale > 70% per task ripetitivi.
• Precision of entity extraction (P): P = entità corrette / entità estratte. Soglia di allerta: P • Recall su fonti verificate (R): R = fonti rilevanti trovate / fonti note esistenti. Target operativo: R ≥ 0.75 per corpus non triviale.
• Confidence-adjusted accuracy (CAA): percentuale di output con fonte verificata e confidenza alta.
• Reduction in manual hours per caso (H): ore risparmiate = ore umane pre-IA - ore con IA.

Esempio numerico (simulato): su 10.000 documenti, Elastic AI + LLM ha ridotto il triage da 120 a 8 ore; precision estrazione entità = 92%; recall fonti critiche = 78%.

Reportistica suggerita: dashboard Grafana con metriche di job (durata job, tassi di errore, percentuale di output verificati, latenza). Frequenza report: settimanale per operazioni continue; post-mortem dopo ogni indagine significativa.

Errori da evitare (e come risolverli)

• Fidarsi ciecamente dell’IA. L’IA può inventare fonti, nomi, date. Verifica sempre i risultati critici con fonti primarie.
• Prompt troppo vaghi. “Analizza questo documento” non è un prompt. Specifica cosa cerchi: nomi, date, contraddizioni, pattern.
• Non iterare. Il primo prompt raramente è perfetto. Testa, modifica, riprova. Salva i prompt che funzionano.
• Usare un solo strumento. ChatGPT non può fare tutto. Combina LLM, visualizzatori, archivisti e processori di massa.
• Dimenticare la catena di custodia. Se lavori su casi legali, ogni passaggio deve essere documentato. Usa strumenti come Hunchly per mantenere tracce verificabili.

Checklist forensica: catena di custodia e verifica quando si usa IA

1. Log completo delle sessioni: timestamp, URL, input testuali inviati all'IA e output ricevuti; salva hash di input e output.
2. Anonymizzazione preventiva: rimuovi o pseudonimizza dati sensibili prima di inviarli a modelli cloud; conserva copia originale cifrata separata.
3. Archiviazione forense: usa Hunchly o imaging forense per screenshot, HTML capture e metadati; associa ogni elemento a un ID investigativo.
4. Verifica delle fonti: per ogni affermazione critica prodotta dall'IA, richiedi almeno una fonte primaria verificabile e registra URL, data di accesso e screenshot.
5. Conservazione dei prompt: salva la versione completa dei prompt usati (istruzioni, parametri) per riproducibilità e auditing.
6. Controllo versioning modelli: registra quale modello/endpoint è stato usato (versione, provider) e note su aggiornamenti che potrebbero alterare output.
7. Escalation procedure: definisci quando un output richiede revisione umana o consulenza legale

(es. potenziali implicazioni penali).

FAQ ottimizzate per featured snippets

L'IA può sostituire un investigatore OSINT?

No. L'IA accelera l'analisi e automatizza compiti ripetitivi, ma non sostituisce giudizio investigativo, verifica delle fonti e valutazione legale.

Posso inviare dati sensibili a ChatGPT per un'indagine?

No. Evita di inviare dati sensibili o personali a modelli cloud a meno che non ci sia un DPA e misure di anonimizzazione; preferisci soluzioni on-premise per dati riservati.

Come riduco le allucinazioni dell'IA?

Usa prompt che richiedono fonti, specifica "SCONOSCIUTO" se non trovi fonti, spezzetta il compito in più passaggi e verifica tutte le informazioni critiche con fonti primarie.

Quali strumenti usare per mantenere la catena di custodia?

Usa soluzioni di archiviazione forense (Hunchly, imaging forense) che registrano URL, screenshot e hash, affiancate da logging delle richieste AI e conservazione delle versioni dei prompt.

Qual è il miglior approccio per investigare immagini sospette?

Combina reverse image search (TinEye / Google Lens), analisi EXIF metadati, ricerca per similitudine e confronto con banche immagini; documenta ogni passaggio con screenshot e hash.

Quanto costa integrare IA in OSINT?

Dipende dal volume dati e dalla soluzione: piani consumer per LLM partono da decine di euro al mese; per progetti enterprise servono licenze, infrastruttura on-prem e personale specializzato.

Prossimi passi

L'intelligenza artificiale può ridurre significativamente il tempo speso in attività ripetitive e aumentare la qualità delle analisi OSINT, purché venga integrata con processi forensi, controllo umano e rispetto della normativa. Segui il workflow, applica la checklist forense e misura i KPI per dimostrare valore e conformità.

Vuoi iniziare subito? Prendi un caso OSINT su cui stai lavorando. Scegli uno strumento da questa lista. Scrivi un prompt seguendo la struttura che ti ho mostrato. Testa, itera, affina. In poche ore avrai integrato l’IA nel tuo flusso di lavoro.

Risorse e community:

• Newsletter: https://coondivido.substack.com/
• Telegram: https://t.me/osintaipertutti
• Telegram: https://t.me/osintprojectgroup

Ulteriori riferimenti

Per approfondire compliance e pratiche forensi consultare GDPR, le linee guida ENISA e la documentazione ufficiale dei tool citati; per template di logging usare JSONL con i campi consigliati nella sezione 'Quadro normativo'.