

MODULO S5.2 • AGOSTO 2026

# Multimedia Forensics e Contenuti Sintetici

La perizia su immagini, audio e video  
nell'era dei generatori

`exiftool -a -G1 -s file.mp4 • ffprobe -show_streams • c2patool file.jpg`

Corso di Laurea Magistrale in Sicurezza Informatica

# Al termine del modulo lo studente sarà in grado di

## 01 • ATTRIBUIRE

Un contenuto a un dispositivo o a una catena di elaborazione, dichiarando i limiti di ogni tecnica impiegata

## 02 • ACCERTARE

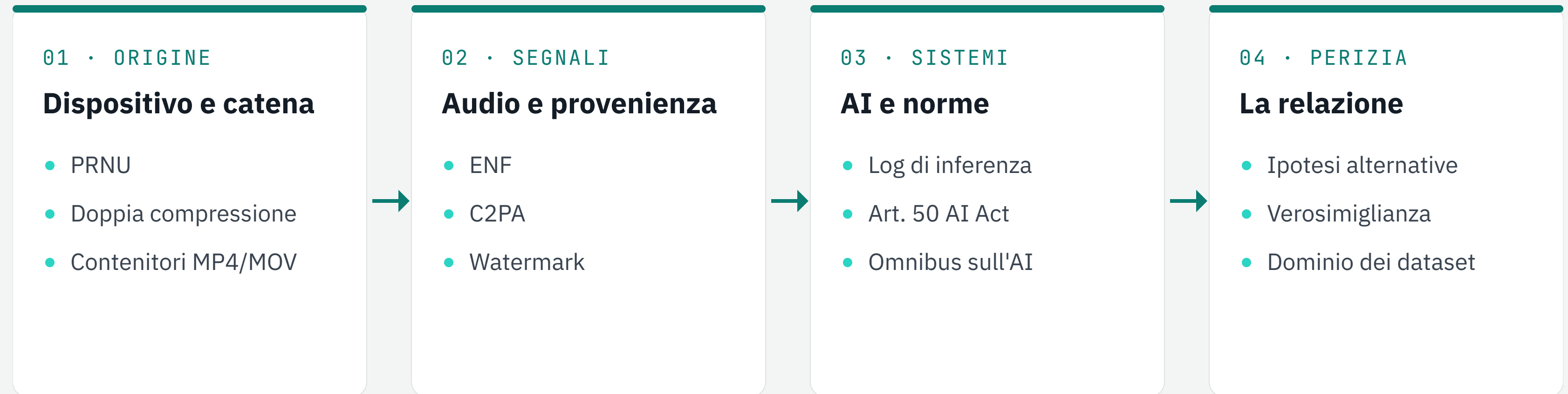
Manipolazioni e generazione sintetica, combinando metodi indipendenti su immagine, audio e video

## 03 • REDIGERE

Una perizia per ipotesi alternative, coerente con il quadro normativo vigente sui contenuti sintetici

**Prerequisiti.** Il modulo presuppone M10 (forensics dell'AI) e i moduli core M7 e M8 su metriche, soglie e validazione. La valutazione di ammissibilità dei risultati viene ripresa in M11.

# Dal file alla perizia per ipotesi alternative



Le quattro aree si susseguono nelle slide del modulo nello stesso ordine dello schema: ciascuna presuppone la precedente.

# Il rumore del sensore come impronta della fotocamera

Ogni sensore d'immagine presenta minime disuniformità nella risposta dei pixel alla luce. Il **PRNU** (photo-response non-uniformity) lascia in ogni fotografia un rumore debole e stabile, caratteristico del singolo esemplare di fotocamera e indipendente dalla scena ripresa.

Il metodo stima un'**impronta di riferimento** da decine di immagini note dello stesso dispositivo e la confronta con il residuo di rumore dell'immagine contestata. Il risultato si esprime come compatibilità, insieme al tasso di falsi abbinamenti misurato su altre fotocamere dello stesso modello.

Ricompressione, ridimensionamento, ritaglio e stabilizzazione indeboliscono l'impronta, come la fusione di più scatti negli smartphone.

## ESEMPIO • LA FOTO DEL SOPRALLUOGO

Una fotografia prodotta in giudizio dovrebbe provenire dalla fotocamera di un tecnico. Il perito ottiene trenta scatti di riferimento dallo stesso apparecchio.

La correlazione con l'immagine contestata supera nettamente i valori misurati su dieci fotocamere dello stesso modello. La relazione esprime un forte sostegno all'ipotesi di provenienza e precisa che il file esaminato non era stato ricompresso.

# Le impronte dell'elaborazione

## Doppia compressione JPEG

Un'immagine salvata due volte con tabelle di quantizzazione diverse lascia periodicità nei coefficienti DCT. Se l'effetto riguarda una sola regione, suggerisce un **innesto** da un'altra immagine.

## Contenitori MP4 e MOV

Il video è una sequenza di atomi: **ftyp** dichiara il formato, **moov** contiene metadati e indici, **mdat** i dati. Ordine degli atomi ed encoder rivelano il software che ha scritto il file.

## Catena di condivisione

Le piattaforme ricomprimono con parametri propri e rimuovono gran parte dei metadati. La ricompressione identifica il **servizio attraversato** e cancella molti indizi di manipolazione.

**ESEMPIO** • Un video presentato come ripresa originale ha atomi ordinati diversamente dai file dello stesso modello di telefono, encoder di un'app di montaggio e risoluzione tipica di una piattaforma di messaggistica. Il file risulta rielaborato e ricondiviso; la falsità del contenuto richiede altre analisi.

# Continuità, spettro e frequenza di rete

Su una registrazione contestata la prima verifica riguarda la **continuità**: tagli e inserimenti lasciano discontinuità nel rumore di fondo, nella fase e nello spettro. L'analisi spettrale rende visibili questi punti e permette di confrontare il canale acustico con quello dichiarato.

La **ENF** (electric network frequency) sfrutta un riferimento esterno: la rete elettrica europea lavora a 50 Hz nominali con microvariazioni continue, che una registrazione vicina a impianti alimentati può catturare. Il confronto con archivi di riferimento aiuta a datare il file e a rilevare tagli.

La traccia ENF può mancare nei dispositivi a batteria: l'assenza va dichiarata come limite.

## ESEMPIO • LA TELEFONATA REGISTRATA

Una registrazione di quaranta minuti documenterebbe una minaccia. La difesa sostiene che la frase decisiva sia stata spostata dalla fine all'inizio.

L'analisi rileva due discontinuità del rumore di fondo attorno alla frase e un salto della traccia ENF nello stesso punto rispetto all'archivio di riferimento. Gli indizi convergono e la relazione esprime un forte sostegno all'ipotesi di montaggio.

# Credenziali e watermark alla prova

## Leggere il manifest

**c2patool** legge i manifest C2PA incorporati: firmatario, certificato, modifiche dichiarate. L'esito si salva integralmente, con versione dello strumento e hash del file.

## Firma e contenuto

Una firma valida attesta l'**integrità del manifest** e il titolare del certificato. La veridicità della scena resta da accertare: anche un falso può essere firmato da chi lo produce.

## Watermark e assenze

I watermark si verificano con gli strumenti del fornitore, se esistono. Credenziali rimosse e modelli che non marcano rendono l'**assenza di segnali** compatibile con file autentici e sintetici.

**ESEMPIO** • La foto di un incidente porta credenziali C2PA valide di un software di fotoritocco, con un riempimento generativo dichiarato su una regione. Il perito descrive l'area e valuta se incide sui fatti del quesito; una copia della stessa foto scaricata da un social non ha credenziali.

# Quando l'oggetto della perizia è il sistema di AI

In un incidente che coinvolge un'applicazione basata su modelli, le fonti di prova cambiano: **log di inferenza**, prompt e output conservati, configurazione, versione del modello in uso alla data dei fatti, documenti recuperati da eventuali pipeline RAG. Si acquisiscono con la disciplina di ogni reperto: hash, orari, catena di custodia.

La **provenienza del modello** è un tema emergente: pesi, schede tecniche e tecniche di impronta aiutano a stabilire quale modello abbia prodotto un output, o se un modello distribuito derivi da un altro. Il NIST AI 600-1 (2024) offre un riferimento per classificare rischi e incidenti dell'AI generativa.

Senza log conservati all'epoca dei fatti la ricostruzione resta parziale: le politiche di logging vanno chieste e verbalizzate per prime.

## ESEMPIO • L'ASSISTENTE CHE DIVULGA DATI

Il chatbot di assistenza di un'azienda mostra a un cliente i dati di un altro.

Il perito acquisisce i log delle conversazioni, la configurazione e l'indice dei documenti usato per il recupero. Ne emerge che un documento interno era stato indicizzato per errore: la causa risiede nella pipeline di recupero. La relazione tiene distinte le cause tecniche dalle valutazioni giuridiche.

# Art. 50 AI Act: obblighi e date

## Le date

Obblighi applicabili dal **2 agosto 2026**. L'omnibus sull'AI (in vigore dal 27 luglio 2026) ne lascia invariata la sostanza; per i sistemi generativi già sul mercato la marcatura slitta al 2 dicembre 2026.

## Chi deve fare che cosa

I **fornitori** marcano gli output in formato leggibile dalla macchina. I **deployer** che diffondono deepfake ne dichiarano l'origine artificiale. Linee guida della Commissione: luglio 2026.

## Per il perito

Un contenuto senza marcatura può venire da un sistema non conforme, da un **modello a pesi aperti** o da un file ricompresso. Per questo l'origine del file si accerta con l'analisi tecnica.

**ESEMPIO** • Un video diffamatorio diffuso a settembre 2026 è privo di marcature. Il perito individua i generatori compatibili con le tracce, verifica quali fossero soggetti all'obbligo a quella data, tenendo conto della proroga, e chiede ai fornitori gli strumenti di verifica.

# Scrivere la perizia per ipotesi alternative

La valutazione di un contenuto contestato si costruisce su almeno due ipotesi esplicite, per esempio «ripresa originale» e «contenuto generato o manipolato». Per ogni indizio il perito stima quanto sia atteso sotto ciascuna ipotesi: il **rapporto di verosimiglianza** esprime il peso dell'indizio senza pronunciarsi sulla probabilità che un'ipotesi sia vera.

La **ENFSI Guideline for Evaluative Reporting in Forensic Science** (2015) suggerisce di comunicare quel peso con una scala verbale ancorata a intervalli numerici, da sostegno debole a sostegno molto forte. La probabilità delle ipotesi resta materia del giudice, che dispone anche delle altre prove.

Ogni relazione dichiara metodi, strumenti con versione, dominio di validazione e domande a cui l'analisi non sa rispondere.

## ESEMPIO • DUE FORMULAZIONI A CONFRONTO

Da evitare: «il video è falso con probabilità del 95%».

Corretta: «le anomalie osservate sono molto più probabili se il video è sintetico che se è una ripresa originale; i risultati forniscono un sostegno forte all'ipotesi di generazione sintetica, nei limiti del dominio di validazione dichiarato».

# Dataset e limiti dei detector

## Dataset video

**FaceForensics++** raccoglie video con volti sostituiti o animati, a più livelli di compressione. Il **DFDC** offre un corpus ampio girato con attori consenzienti. Entrambi hanno condizioni d'uso specifiche.

## Dataset audio

Le edizioni di **ASVspoof** forniscono parlato autentico e contraffatto (sintesi, conversione della voce, riproduzione) con protocolli di valutazione standard per confrontare i detector.

## Generalizzazione

Fuori dal dominio di addestramento un detector perde accuratezza: generatori nuovi, compressioni, condizioni di ripresa. In perizia si dichiara su quali dati lo strumento è stato **validato**.

**ESEMPIO** • Un detector con ottime prestazioni su FaceForensics++ perde accuratezza in modo marcato su video ricompressi da una piattaforma di messaggistica e generati con tecniche recenti. La relazione riporta entrambe le misure e usa lo strumento come un indizio fra gli altri.

# Una perizia completa su un video di prova

## Laboratorio

- Acquisizione documentata del video
- Contenitore, compressione, coerenza
- Ipotesi alternative e peso degli indizi

DELIVERABLE: relazione peritale

## Lecture obbligatorie

- Farid, *Photo Forensics*
- Rathgeb et al., *Handbook of Digital Face Manipulation*
- ENFSI, *Evaluative Reporting* (2015)

## Verifica

Ogni conclusione rimanda a un'analisi ripetibile.

**Prossimo modulo:** M11 Profili giuridici ed etici (traccia core), oppure il modulo base della seconda specializzazione scelta.

La perizia regge quando ogni conclusione ha un'analisi ripetibile e un limite dichiarato.



# Grazie per l'attenzione

**Vincenzo Calabrò**

Domande, materiali del corso e contatti  
sul sito del docente



[vincenzocalabro.it](https://vincenzocalabro.it)