

MODULO 15

Governance, etica e diritto

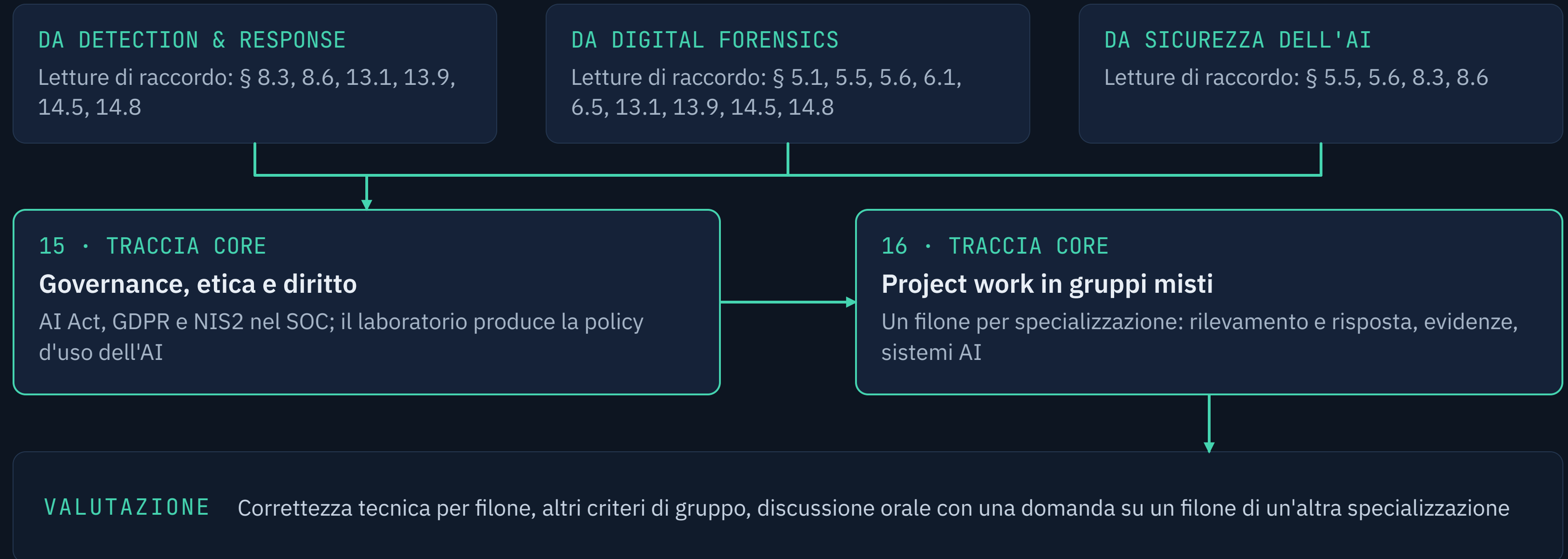
L'AI nella risposta agli incidenti dentro il quadro europeo: AI Act, GDPR, NIS2, responsabilità delle decisioni automatizzate, bias, spiegabilità e tutela dei dati di telemetria.

AI ACT

GDPR / NIS2

XAI

La chiusura comune a tutti i percorsi



Abstract e obiettivi formativi

Abstract

Il modulo colloca l'impiego dell'AI nella risposta agli incidenti nel quadro normativo europeo (AI Act, GDPR, NIS2) e ne discute le implicazioni etiche: responsabilità delle decisioni automatizzate, bias nei modelli di detection, sorveglianza dei dipendenti, spiegabilità come requisito organizzativo e probatorio, policy interna d'uso dell'AI nel SOC.

Obiettivi formativi (Bloom)

- **Applicare:** classificare un sistema AI di sicurezza secondo le categorie di rischio dell'AI Act
- **Analizzare:** valutare la base giuridica del trattamento dei dati di telemetria (legittimo interesse, DPIA)
- **Valutare:** attribuire le responsabilità in caso di danno da azione automatizzata

Traccia core

01 02 07

Detection & Response

03 04 05 06

Digital Forensics

08 09 10 11

Sicurezza dell'AI

06 12 13 14

Traccia core

15 16

Lecture di raccordo: § 5.1, 5.5, 5.6, 6.1, 6.5, 8.3, 8.6, 13.1, 13.9, 14.5, 14.8, salvo quelle già studiate

Vincoli giuridici ed etici sull'uso dell'AI



L'AI Act e gli strumenti di sicurezza

VIETATO

Art. 5: manipolazione, social scoring, inferenza delle emozioni da dati biometrici al lavoro e a scuola, salvo motivi medici o di sicurezza

ALTO RISCHIO

Monitoraggio e valutazione dei lavoratori o decisioni sul rapporto di lavoro (allegato III, punto 4): obblighi che si applicano dal 2 dicembre 2027

RISCHIO LIMITATO / MINIMO

Gran parte degli strumenti di detection e triage; trasparenza solo nei casi dell'art. 50

Implicazioni per il SOC

La classificazione si documenta sistema per sistema; se il sistema profila i dipendenti, la deroga dell'art. 6, par. 3, non si applica. Il deployer ha obblighi propri, distinti da quelli del fornitore. Secondo il considerando 55 i componenti usati solo per la cybersicurezza non dovrebbero essere considerati componenti di sicurezza delle infrastrutture critiche.

Obblighi del GDPR e della NIS2 per il SOC

Telemetria e GDPR

- Log con dati personali dei dipendenti: legittimo interesse come base giuridica, minimizzazione
- UEBA e monitoraggio esteso: DPIA quasi sempre necessaria; in Italia, anche lo Statuto dei lavoratori (art. 4)
- Trasferimenti verso servizi AI extra-UE: capo V del GDPR
- **Art. 22:** per un blocco automatico con effetti significativi, intervento umano e contestazione

NIS2 e responsabilità

- NIS2: gestione del rischio e degli incidenti come obbligo degli organi di vertice, sanzionabile
- Danno da azione automatizzata errata: risponde l'organizzazione che impiega il sistema; i contratti con i fornitori regolano i rapporti interni
- Con l'audit trail (Modulo 14) si ricostruisce e si motiva la decisione

Bias, proporzionalità, spiegabilità

BIAS

Un risk score che penalizza sistematicamente un gruppo, come i turnisti, un reparto o una sede, produce effetti discriminatori; va verificato con test per gruppo

PROPORZIONALITÀ

Il monitoraggio per finalità di sicurezza resta soggetto a necessità e proporzionalità: dati minimi, conservazione limitata, finalità dichiarate

SPIEGABILITÀ (XAI)

Requisito organizzativo e probatorio: ogni alert deve poter essere motivato verso l'analista, verso l'interessato e in giudizio

La policy d'uso dell'AI nel SOC

Documento, redatto in laboratorio, che raccoglie inventario dei sistemi AI e loro classificazione, dati ammessi verso ciascun servizio, punti di revisione umana obbligatoria, requisiti di registrazione e responsabilità nominative.

Le norme nella pratica del SOC



AI Act e strumenti di incident response

Il Regolamento (UE) 2024/1689 distingue pratiche vietate (art. 5), sistemi ad alto rischio (art. 6 e allegato III), sistemi soggetti a obblighi di trasparenza (art. 50) e sistemi a rischio minimo. Per un SOC l'esercizio consiste nel classificare ciascuno strumento in uso. Un filtro antispam ricade di norma nel rischio minimo. Un sistema che monitora e valuta il comportamento dei dipendenti, o che incide sulle decisioni relative al rapporto di lavoro, può rientrare nell'**alto rischio** (allegato III, punto 4), con obblighi di registrazione degli eventi, documentazione e sorveglianza umana che si applicano dal 2 dicembre 2027 (Reg. (UE) 2026/1744). Secondo il considerando 55, i componenti destinati solo alla cybersicurezza non dovrebbero essere considerati componenti di sicurezza delle infrastrutture critiche.

Scenario didattico

Si supponga che il censimento degli strumenti di un SOC individui quattordici sistemi con componenti AI. Dodici sono a rischio minimo o limitato; per l'UEBA che alimenta provvedimenti disciplinari e per il sistema antifrode interno si svolge l'analisi da alto rischio, con registro dei trattamenti aggiornato, log conservati e procedura di intervento umano scritta e verificata.

Nota normativa

L'organizzazione che utilizza il sistema, anche quando lo acquista da un fornitore, è deployer e ha obblighi propri (art. 26): uso conforme alle istruzioni, sorveglianza umana affidata a persone competenti, monitoraggio del funzionamento, conservazione dei log.

Decisioni automatizzate e art. 22 GDPR

L'art. 22 GDPR riconosce all'interessato il diritto di non essere sottoposto a una decisione **basata unicamente sul trattamento automatizzato** che produca effetti giuridici o incida in modo analogo significativamente sulla sua persona, salvo eccezioni (contratto, legge, consenso esplicito): per contratto e consenso servono almeno intervento umano, possibilità di esprimere la propria opinione e di contestare la decisione (art. 22, par. 3); per la legge, le tutele previste dalla legge stessa. Nella risposta agli incidenti un caso da valutare è la sospensione automatica dell'account di un dipendente segnalato dall'UEBA, con effetti sul rapporto di lavoro. L'intervento umano è effettivo se affidato a una persona con competenza e autorità per modificare la decisione; una ratifica sistematica, condizionata dall'automation bias (Modulo 5), non soddisfa il requisito. A monte servono DPIA (art. 35), informativa ai dipendenti e minimizzazione dei dati che alimentano i modelli.

Esempio

Nel playbook l'UEBA segnala, la piattaforma SOAR prepara la sospensione e l'esecuzione attende l'approvazione dell'analista, che esamina le evidenze e può dissentire; decisione e motivazione restano nel log. Il tempo di risposta resta nell'ordine dei minuti e la decisione non è basata unicamente sul trattamento automatizzato.

Nota normativa

Si applicano anche gli artt. 33 e 34: il titolare notifica senza ingiustificato ritardo e, ove possibile, entro 72 ore dalla conoscenza della violazione, salvo rischio improbabile, e informa gli interessati se il rischio è elevato. L'AI può accelerare triage e stima dell'impatto; il testo notificato richiede revisione umana.

Responsabilità per gli errori dei sistemi AI

Quando impiega un sistema automatico, l'organizzazione che lo utilizza resta responsabile insieme alle persone che ne decidono configurazione e uso. La catena delle responsabilità va definita in anticipo. Il **fornitore** risponde del sistema come progettato e documentato e il **deployer** dell'uso concreto, della configurazione, delle soglie e della scelta di automatizzare un'azione. Il singolo operatore risponde delle decisioni che gli competono. La documentazione interna che registra chi ha autorizzato quale livello di autonomia, e su quali valutazioni, deve esistere prima dei fatti, perché ricostruita a posteriori avrebbe scarso valore. Dopo un incidente la diligenza dell'organizzazione si valuta su questa documentazione e sul contratto con il fornitore, che fissa livelli di servizio, limiti dichiarati e aggiornamenti.

Scenario didattico

Un contenimento automatico isola per errore i sistemi di cassa in un fine settimana di forte attività, con danno economico. L'istruttoria interna trova una matrice tra azioni e livelli di autonomia approvata dal CISO, una procedura di rollback verificata e un errore entro il tasso dichiarato dal fornitore. Il processo è quindi difendibile e il caso entra nel golden set (Modulo 14).

Cautele

Oltre ai requisiti di legge restano questioni etiche: distribuzione delle segnalazioni tra gruppi di persone, proporzionalità del monitoraggio dei dipendenti, trasparenza verso chi è osservato.

NIS2 e obblighi di risposta agli incidenti

La Direttiva (UE) 2022/2555 (NIS2), recepita in Italia con il D.Lgs. 4 settembre 2024, n. 138, rende la capacità di risposta un obbligo per i soggetti essenziali e importanti, sotto la vigilanza dell'Agenzia per la cybersicurezza nazionale (ACN). Le misure di gestione del rischio comprendono la gestione degli incidenti. Per gli incidenti significativi è prevista una **notifica in più fasi** al CSIRT Italia: pre-notifica entro 24 ore e notifica entro 72 ore dalla conoscenza, relazione finale entro un mese dalla notifica. Gli organi di amministrazione e direttivi approvano le misure, ne sovrintendono l'attuazione e possono essere chiamati a risponderne. Poiché i termini decorrono dal momento in cui il soggetto viene a conoscenza dell'incidente significativo, la rapidità di qualificazione (Modulo 1) incide sul rispetto di termini presidiati da sanzioni.

Scenario didattico

Alle 22:00 il SOC viene a conoscenza di un incidente significativo: da quel momento decorre il termine di 24 ore per la pre-notifica. Con triage rapido e canale di notifica già predisposto il termine si rispetta, mentre un SOC che deve ancora stabilire destinatari e modalità rischia di superarlo. L'AI che accelera qualificazione e stima d'impatto (Modulo 5) contribuisce anche alla conformità.

Metodo

Per rispettare i termini si predispongono in anticipo modelli di notifica, criteri scritti per stabilire la significatività, contatti dell'autorità nel playbook e ruoli assegnati. La notifica è un atto con rilievo giuridico e va preparata con la funzione legale.

Bias e monitoraggio dei lavoratori

Molti sistemi difensivi, come UEBA, DLP e rilevazione delle minacce interne, osservano comportamenti umani e sono soggetti anche alle tutele delle persone. Il **bias** nasce dai dati storici: se alcuni ruoli, turni o sedi sono sovrarappresentati tra i casi segnalati in passato, il sistema tende a perpetuare lo squilibrio, e vanno misurati i tassi di segnalazione e di falsi positivi per gruppo, oltre a precision e recall aggregate (Modulo 2, approfondimento 1). Quanto alla **proporzionalità**, in Italia il controllo a distanza dei lavoratori è disciplinato dall'art. 4 della L. 300/1970 (Statuto dei lavoratori), cui si aggiunge il principio di minimizzazione del GDPR. La progettazione deve stabilire se il monitoraggio è necessario e proporzionato rispetto alla finalità di sicurezza e come viene reso trasparente verso chi è osservato.

Scenario didattico

Un sistema per le minacce interne segnala con frequenza molto maggiore i dipendenti di un reparto. La verifica per gruppo mostra che il reparto lavora con orari irregolari per ragioni legittime, sovrarappresentati nei dati storici sulle anomalie, ed evita così indagini disciplinari ingiustificate.

Cautele

Ai lavoratori monitorati va data un'informativa chiara, anche sull'uso dell'AI (L. 132/2025, art. 11), e dove l'art. 4 dello Statuto lo richiede servono l'accordo sindacale o l'autorizzazione. Per i trattamenti ad alto rischio si svolge la valutazione d'impatto. I tassi di segnalazione per gruppo si verificano periodicamente, e ogni segnalazione del sistema avvia una verifica umana senza costituire un accertamento.

Governance dell'AI nel SOC

Gli obblighi esaminati nel modulo si traducono in una **governance** concreta, che comprende almeno un inventario dei sistemi AI con la relativa classificazione di rischio (approfondimento 1) e una **policy d'uso** che stabilisca strumenti ammessi, attività consentite e dati trasferibili a servizi esterni (Modulo 6). Si aggiungono un titolare del rischio per ciascun sistema, valutazioni e audit periodici collegati al golden set (Modulo 14) e la documentazione che lega questi elementi. L'art. 4 dell'AI Act, come modificato dal Reg. (UE) 2026/1744, richiede inoltre a fornitori e deployer misure per sostenere lo sviluppo dell'alfabetizzazione in materia di AI del personale.

Scenario didattico

Un'ispezione chiede conto dell'uso dell'AI nel SOC.
L'organizzazione presenta inventario e classificazione dei sistemi, policy d'uso approvata, valutazioni d'impatto dei sistemi ad alto rischio, risultati periodici del golden set e registro delle decisioni automatizzate con gli interventi umani.

Metodo

Il project work (Modulo 16) impiega insieme gli elementi della governance trattati nel corso, alcuni dei quali nelle letture di raccordo: playbook (Modulo 1), audit trail dell'automazione (Modulo 5), catena di custodia (Modulo 7), spiegabilità (Modulo 8), threat model dei sistemi AI (Modulo 13), osservabilità (Modulo 14).

Policy d'uso dell'AI per un SOC fittizio

Svolgimento (in gruppi)

1. Inventario dei sistemi AI del SOC fittizio assegnato e classificazione secondo l'AI Act
2. Matrice di responsabilità: chi approva, chi revisiona, chi risponde
3. Requisiti di logging e di revisione umana per ciascun sistema
4. Discussione plenaria con revisione tra pari delle policy prodotte

Consegna

- Policy di 3–4 pagine, firmabile da un dirigente
- Tabella di classificazione dei sistemi, motivata
- Una criticità lasciata aperta, con le opzioni sul tavolo

Lecture e autovalutazione

Lecture obbligatorie

- Reg. (UE) 2024/1689 (AI Act), artt. 4, 5, 6, 12, 14, 26, 50 e allegato III, nel testo modificato dal Reg. (UE) 2026/1744 (Omnibus digitale sull'IA)
- WP29, *Linee guida sul processo decisionale automatizzato* (WP251 rev.01), fatte proprie dall'EDPB; D.Lgs. 138/2024
- NIST AI RMF 1.0 (2023); Floridi et al., *AI4People*, Minds and Machines, 2018

Domande di autovalutazione

1. In quali casi un sistema di sicurezza AI ricade nell'alto rischio dell'AI Act?
2. Quando un contenimento automatico integra una decisione ex art. 22 GDPR?
3. Chi risponde del danno causato da un blocco automatico errato?
4. Quali dati di telemetria richiedono una DPIA?
5. Perché la spiegabilità è anche un requisito probatorio?

FINE DELLA LEZIONE · MODULO 15

Grazie

Governance, etica e diritto · Corso «Incident Response potenziata dall'AI»

Domande e materiali: ricevimento del docente. Profilo e contatti:
vincenzocalabro.it



Vincenzo Calabrò

CYBER SECURITY · DIGITAL FORENSICS



VINCENZOCALABRO.IT