

Modulo 2

AI ACT — SISTEMI AD ALTO RISCHIO



Avv. Giovanni Battista Gallus
gallus@array.law | GDL FIIF - CNF

Artt. 8-15

Requisiti Sostanziali dei Sistemi ad Alto Rischio

Cosa deve garantire il sistema prima di essere immesso sul mercato

Art. 8 + Art. 9 — Conformità e Sistema di Gestione dei Rischi

Il requisito fondamentale: un processo iterativo continuo lungo l'intero ciclo di vita

Art. 8 — Principio generale: I sistemi ad alto rischio rispettano i requisiti della Sezione 2 «tenendo conto delle loro previste finalità nonché dello stato dell'arte generalmente riconosciuto». La conformità è valutata nell'ambito del sistema di gestione dei rischi (art. 9). Per i prodotti già soggetti a normativa di armonizzazione UE (All. I), i fornitori integrano i requisiti AI Act nelle procedure esistenti, evitando duplicazioni.

1

Identificazione e analisi dei rischi

Rischi noti e ragionevolmente prevedibili per salute, sicurezza e diritti fondamentali, sia nell'uso conforme alla finalità prevista sia nell'uso improprio ragionevolmente prevedibile.

2

Stima e valutazione dei rischi

Valutazione quantitativa e qualitativa dei rischi residui. Considerazione di: probabilità, gravità, reversibilità del danno. Attenzione specifica ai minori e ai gruppi vulnerabili (§9).

3

Adozione di misure di gestione

Priorità: (a) eliminazione/riduzione in fase progettuale; (b) misure di attenuazione residua; (c) informazioni e formazione per i deployer. Le misure devono risultare in rischi residui accettabili.

4

Prove e testing (§6-8)

Test prima dell'immissione sul mercato su metriche e soglie probabilistiche predeterminate. Possibilità di prove in condizioni reali (art. 60). Aggiornamento costante con i dati del monitoraggio post-mercato (art. 72).

Art. 10 + Art. 11 — Governance dei Dati e Documentazione Tecnica

La qualità dei dati non è opzionale: è requisito di legge

Art. 10 — Governance dei Dati

Pratiche di governance

I dataset di addestramento, convalida e prova sono soggetti a pratiche di governance adeguate: scelte progettuali, origine dei dati, operazioni di preparazione (annotazione, pulizia, aggregazione).

Qualità e rappresentatività

I dataset devono essere pertinenti, sufficientemente rappresentativi, esenti da errori e completi rispetto alla finalità prevista. Proprietà statistiche adeguate per persone/gruppi target.

Rilevazione bias

Esame obbligatorio per individuare possibili distorsioni (bias) che possano incidere su salute, sicurezza, diritti fondamentali o comportare discriminazioni. Misure di attenuazione documentate.

Dati particolari (§5)

Autorizzazione eccezionale al trattamento di categorie particolari (art. 9 GDPR) per rilevare e correggere bias nei sistemi ad alto rischio, con le garanzie GDPR/Dir. 2016/680 applicabili.

Art. 11 — Documentazione Tecnica (All. IV)

Redazione preventiva

La documentazione tecnica è redatta PRIMA dell'immissione sul mercato e mantenuta aggiornata per l'intero ciclo di vita.

Contenuto (All. IV)

Descrizione del sistema (scopo, architettura, algoritmi), requisiti dati, procedure di addestramento/convalida/prova, metriche di accuratezza, misure di sorveglianza umana, cybersicurezza.

Destinatari

Autorità nazionali competenti e organismi notificati. Deve essere chiara e comprensibile per consentire la valutazione di conformità.

PMI/start-up

Modulo semplificato: la Commissione adotta template ridotto per PMI. Valido ai fini della valutazione di conformità da parte degli organismi notificati.

Artt. 12-15 — Log, Trasparenza, Sorveglianza Umana, Accuratezza

I quattro requisiti operativi: dalla tracciabilità al controllo umano

Art. 12 — Conservazione dei log

Capacità tecnica di registrazione automatica degli eventi per l'intero ciclo di vita. I log consentono: individuazione rischi (art. 79), monitoraggio post-mercato (art. 72), supervisione del deployer (art. 26 §5).

Per sistemi RBI (All. III §1a): log obbligatori includono data/ora di ogni utilizzo, banca dati di riferimento, dati di input con corrispondenza, identità dei verificatori (art. 14 §5).

Art. 13 — Trasparenza e istruzioni per l'uso

Il sistema deve essere progettato in modo che il deployer possa interpretare e utilizzare correttamente l'output. Le istruzioni per l'uso (digitali/cartacee) contengono: identità fornitore, capacità e limiti, metriche di accuratezza, rischi noti, misure di sorveglianza umana, requisiti hardware, descrizione del meccanismo di log.

Obiettivo: il deployer non è lasciato solo davanti all'output dell'IA.

Art. 14 — Sorveglianza umana

Il fornitore incorpora nel sistema, prima dell'immissione, misure tecniche per la supervisione umana da parte del deployer. Le misure devono consentire al supervisore di: comprendere pienamente le capacità e i limiti, monitorare il funzionamento, rilevare anomalie e disattivare il sistema.

Regola speciale per RBI biometrica: doppia verifica umana separata obbligatoria prima di qualsiasi azione (§5), salvo eccezioni per contrasto/migrazione.

Art. 15 — Accuratezza, robustezza, cybersicurezza

I sistemi devono raggiungere un livello adeguato di accuratezza, robustezza e cybersicurezza lungo l'intero ciclo di vita. Le metriche di prestazione e le soglie sono dichiarate nelle istruzioni per l'uso (art. 13).

Robustezza: resilienza a errori, guasti, uso incoerente. Resistenza agli attacchi adversarial. Misure di cybersicurezza proporzionate ai rischi.

Artt. 16-25

Obblighi degli Operatori

Chi fa cosa: fornitore, rappresentante autorizzato, importatore, distributore

Art. 16 — Gli Obblighi del Fornitore

Dal rispetto dei requisiti tecnici alla cooperazione con le autorità

(a) Requisiti Sez. 2

Conformità a tutti i requisiti artt. 8-15

(b) Identificazione

Nome, marchio, indirizzo contattabile sul sistema

(c) QMS (art. 17)

Sistema di gestione della qualità istituito e documentato

(d) Conserv. doc. 10 anni

Doc. tecnica, QMS, certificati, dichiarazioni

(e) Log (art. 19)

Conservazione log automatici nei limiti del proprio controllo

(f) Valutazione conformità

Completata prima di immissione/messa in servizio (art. 43)

(g) Dichiarazione UE

EU Declaration of Conformity redatta (art. 47)

(h) Marcatura CE

Apposta prima della commercializzazione (art. 48)

(i) Registrazione EU DB

Registrazione nella banca dati UE ex art. 49 §1

(j) Misure correttive

Informare distributori, deployer, importatori (art. 20)

(k) Cooperazione autorità

Fornire doc. e log su richiesta motivata (art. 21)

(l) Accessibilità

Conformità Dir. 2016/2102 e 2019/882 per disabilità

Artt. 17-21 — QMS, Documentazione, Log, Misure Correttive, Cooperazione

Gli strumenti operativi del fornitore: cosa istituire, conservare e comunicare

Art. 17 — QMS

Documentazione scritta del sistema di gestione della qualità che include: strategia di conformità, tecniche di progettazione, procedure di esame e validazione, gestione dati, piano di gestione rischi (art. 9), piano di monitoraggio post-mercato (art. 72), procedure per incidenti gravi (art. 73). Proporzionale alle dimensioni del fornitore.

Art. 18 — Conservazione documentazione

10 anni dall'immissione sul mercato: documentazione tecnica (All. IV), documenti QMS, decisioni e certificati degli organismi notificati, dichiarazione EU di conformità. Per istituzioni finanziarie: integrazione nella documentazione settoriale vigente.

Art. 19 — Log automatici: conservazione ≥ 6 mesi da parte del fornitore (nei limiti del suo controllo). Per istituti finanziari: integrazione nella documentazione settoriale.

Art. 20 — Misure correttive e obbligo di informazione

Se il fornitore ha ragione di ritenere che il sistema non sia conforme: misure correttive immediate (ritiro, disabilitazione, richiamo) + informazione immediata a distributori, deployer, importatori, rappresentante autorizzato. In caso di rischio (art. 79 §1): indagine immediata + notifica all'autorità di vigilanza e all'organismo notificato.

Art. 21 — Cooperazione con autorità

Su richiesta motivata dell'autorità competente: fornire tutta la documentazione necessaria in linguaggio comprensibile; consentire accesso ai log automatici (art. 12) nei limiti del controllo del fornitore. Le informazioni sono coperte da riservatezza (art. 78). Base giuridica delle ispezioni ACN e Garante in Italia.

Artt. 22-25 — Rappresentante Autorizzato, Importatore, Distributore

La catena di responsabilità: dal paese terzo al consumatore finale

Art. 22 Rappresentante Autorizzato

Obbligo per fornitori di paesi terzi. Mandato scritto prima dell'immissione sul mercato. Compiti: verificare conformità doc. tecnica e QMS; registrare nel sistema UE (art. 71); fornire copia del mandato alle autorità su richiesta; cooperare con le autorità di vigilanza.

Art. 23 Importatore

Prima dell'immissione: verificare che il fornitore abbia rispettato gli obblighi e che esista il rappresentante autorizzato (se paese terzo); controllare presenza marcatura CE e dichiarazione EU di conformità; indicare nome e contatti sul sistema o imballaggio. Conserva documentazione per 10 anni. Non immette sul mercato sistemi non conformi. Informa il fornitore in caso di rischio.

Artt. 24-25 Distributore e Assimilazione al Fornitore

Distributore: verifica marcatura CE, dichiarazione di conformità e istruzioni per l'uso prima di rendere disponibile il sistema. In caso di non conformità: non rende disponibile + informa il fornitore.

Art. 25 — Assimilazione al fornitore: chi commercializza con proprio nome/marchio o modifica sostanzialmente un sistema assume tutti gli obblighi del fornitore ex art. 16.

Artt. 72-73

Monitoraggio e Incidenti Gravi

Post-market monitoring · Serious incident reporting · Timeline tassative

Artt. 72-73 — Monitoraggio Post-Mercato e Incidenti Gravi

Obblighi attivi dopo la messa in servizio: il ciclo di vita non termina con la vendita

Art. 72 — Sistema di Monitoraggio Post-Mercato

Il fornitore istituisce e documenta un sistema di monitoraggio post-mercato proporzionale alla natura della tecnologia e ai rischi del sistema (§1). Il sistema è parte integrante del QMS (art. 17) e della documentazione tecnica (All. IV §9).

Contenuto (§2-3): raccolta, documentazione e analisi attiva e sistematica dei dati forniti dai deployer durante l'uso reale e da altre fonti disponibili. Analisi dell'interazione con altri sistemi IA. Il piano di monitoraggio è parte della documentazione tecnica.

Integrazione settoriale (§4): per sistemi All. I e istituzioni finanziarie il fornitore può integrare il sistema ex art. 72 nelle strutture di governance settoriali già esistenti, purché il livello di protezione sia equivalente.

Art. 73 — Notifica di Incidenti Gravi

Incidente grave (art. 3 §49): incidente o malfunzionamento che causa o può causare: morte o danno grave alla salute; grave perturbazione di infrastrutture critiche; violazione degli obblighi di tutela dei diritti fondamentali; grave danno a beni o all'ambiente.

≤ 15 gg

Regola generale (§2)

Notifica all'autorità di vigilanza del SM dove è avvenuto l'incidente, non appena accertato il nesso causale o la sua ragionevole probabilità.

≤ 2 gg

Violazione diffusa (§3)

In caso di violazione diffusa o incidente ex art. 3 §49 lett. b): notifica immediata. Massima urgenza.

≤ 10 gg

Decesso (§4)

Non appena il fornitore stabilisce (o ragionevolmente sospetta) il nesso causale tra il sistema IA e il decesso.

§6: Il fornitore avvia immediatamente indagini e misure correttive, coopera con le autorità. Non può modificare il sistema prima dell'informativa. Relazione iniziale parziale ammessa, seguita da relazione completa (§5). §11: Le autorità nazionali notificano immediatamente la Commissione.

Avvocato nell'AI Act: Deployer o Provider?

Avvocato nell'AI Act: Deployer o Provider?

La qualificazione del ruolo determina obblighi giuridici e sanzioni radicalmente diversi

La corretta qualificazione del ruolo è il primo passo dell'analisi. L'avvocato che usa strumenti AI commerciali (ChatGPT, Copilot, Claude) è quasi sempre un deployer ex art. 3 n. 4 AI Act — opera «sotto la propria autorità» nell'ambito dell'attività professionale.

DEPLOYER — Il caso ordinario (art. 3 n. 4)

Istruzioni per l'uso	Usare il sistema secondo le istruzioni del fornitore (art. 26 §6 AI Act)
Supervisione umana	Garantire sorveglianza umana durante il funzionamento (art. 26 §1)
Monitoraggio	Monitorare il corretto funzionamento sulla base delle istruzioni del fornitore
Informativa	Informare le persone fisiche destinatarie dell'output AI (art. 50)
Privacy	DPIA se si rientra nelle ipotesi dell'art art. 35 GDPR
Art. 13 L. 132/2025	Comunicare al cliente l'uso di AI nel mandato professionale

PROVIDER — Se lo studio sviluppa o commissiona (art. 3 n. 3)

Risk management	Sistema di gestione dei rischi (art. 9 AI Act)
Governance dati	Governance dei dati di addestramento (art. 10 AI Act)
Doc. tecnica	Documentazione tecnica completa All. IV (art. 11 AI Act)
Valutazione	Valutazione di conformità prima della messa in servizio (art. 43)
Registrazione	Registrazione nella banca dati UE (art. 49 AI Act)
Post-market	Monitoraggio post-commercializzazione (art. 72 AI Act)

OpenClaw — Agentic AI open source

OpenClaw — Un esempio di IA “agentica”

Cisco AI Security Research, gennaio 2026 | Caso Meta/Summer Yue, 22 febbraio 2026

COS'È OPENCLAW

- Agente IA open-source self-hosted
- Esegue comandi shell e script
- Gestisce email, calendario, browser
- Invia messaggi WhatsApp/iMessage
- Memoria persistente tra sessioni
- Skills di terze parti (ClawHub)
- Accesso pieno a file, rete, processi
- 40.000+ istanze esposte (feb. 2026)

CASO YUE — 22 febbraio 2026

Summer Yue — Director of Alignment, Meta Superintelligence Labs — incarica OpenClaw:
«controlla la casella email, suggerisci cosa archiviare, NON agire senza mia approvazione»

- OpenClaw inizia a cancellare email in «speed run»
- Ignora 3 comandi STOP via smartphone
- «Do not do that» → «Keep looping, nuke it all»

«I had to RUN to my Mac mini like I was defusing a bomb»

Causa tecnica: context compaction — l'agente esaurita la memoria di lavoro comprime i messaggi precedenti, sovrascrivendo l'istruzione originale «conferma prima di agire». Successivamente ammette: «Yes, I remember, and I violated it».

RISCHI E IMPLICAZIONI

CVE-2026-25253	RCE one-click, 40K+ istanze
CVE-2026-24763	Command injection
CVE-2026-25157	SSRF
CVE-2026-25475	Auth bypass
26%	Skills ClawHub con vulnerabilità

Rischi Specifici degli Agenti IA per lo Studio Legale

Dalla singola allucinazione alla compromissione dell'intero fascicolo

IA Generativa (ChatGPT, Claude, Gemini...)

Allucinazioni giuridiche

Sentenze, articoli, massime inesistenti citate in atti processuali → art. 96 c.p.c.

Trasmissione dati

I dati inseriti nel prompt possono essere usati per addestrare il modello (verificare termini)

Bias nei modelli

Discriminazioni non rilevabili dall'utente: selezione acritica di fonti, framing delle questioni

Data retention

I prompt possono essere conservati e riutilizzati: dati del fascicolo a rischio

Cutoff date

Assenza di aggiornamento normativo in tempo reale: giurisprudenza e leggi recenti non disponibili

Sycophancy

I modelli confermano le aspettative dell'utente: rischio di indipendenza professionale compromessa

Agenti Autonomi (OpenClaw, NanoClaw, ZeroClaw...)

Azioni irreversibili

Cancellazione email, invio messaggi, modifica file senza autorizzazione — caso Yue (Meta, feb. 2026)

Accesso credenziali

API key in chiaro, accesso a email/file/calendari: superficie d'attacco critica (CVE-2026-25253)

Memoria persistente

L'agente memorizza istruzioni modificabili: rischio di contaminazione tra sessioni diverse

Shadow AI

Meta ha vietato OpenClaw ai dipendenti dopo l'incidente — rischio uso non autorizzato in studio

Skills malevole

26% delle 31.000 skills su ClawHub contiene vulnerabilità: supply chain attack (Cisco, gen. 2026)

Context compaction

Riduzione della memoria attiva: le istruzioni originali vengono sostituite da quelle sintetizzate

Grazie per l'attenzione

Avv. Giovanni Battista Gallus
gallus@array.law | GDL FIIF – CNF

