T

Marco Billi e Antonino Rotolo

Valutazione della conformità ai sensi dell'articolo 57 dell'AI Act: sinergia con gli spazi di sperimentazione normativa*

Introduzione

Il Regolamento (UE) 2024/1689 sull'intelligenza artificiale (*AI Act*), pubblicato nella Gazzetta Ufficiale dell'Unione europea il 12 luglio 2024 ed entrato in vigore il 1° agosto 2024, segna l'avvio di una nuova fase della regolazione tecnologica europea. La maggior parte delle sue disposizioni diventerà applicabile dal 2 agosto 2026, introducendo un quadro uniforme di requisiti e procedure per lo sviluppo, l'immissione sul mercato e l'utilizzo dei sistemi di intelligenza artificiale nell'Unione.

La portata del Regolamento è ampia: esso si applica a fornitori, distributori, importatori, rappresentanti autorizzati, produttori (*providers*) di prodotti che integrano sistemi di IA e utilizzatori (*deployers*) che operano nel mercato dell'UE o i cui sistemi incidono su soggetti europei, anche se sviluppati al di fuori del territorio unionale (art. 2).

La valutazione di conformità (*conformity assessment*) è disciplinata in particolare dal Capitolo III, Sezione 2, che stabilisce i requisiti per i sistemi di IA ad alto rischio. Secondo l'articolo 3, paragrafo (20), essa è definita come «il processo volto a dimostrare se i requisiti stabiliti nella Sezione 2 del Capitolo III, relativi a un sistema di IA ad alto rischio, siano stati soddisfatti».

La conformità all'AI Act consiste in una valutazione multilivello che dipende da diversi fattori[1]:

* Il presente contributo è stato scritto prima della pubblicazione della bozza dell'atto di esecuzione per l'istituzione di sandbox per l'IA ai sensi dell'AI Act.

[1] Una disamina completa della metodologia da seguire per valutare la conformità di un si-

 DOI: 10.4454/02g1gb68

– La natura del sistema – determinare se il sistema rientri nella definizione di «sistema di intelligenza artificiale» di cui all'art. 3(1), ossia se operi con un certo grado di autonomia e produca output (predizioni, raccomandazioni, decisioni) che influenzano ambienti o persone.
– Il ruolo dell'attore economico – il Regolamento distingue obblighi specifici per fornitori, *deployers*, importatori e distributori. L'individuazione del ruolo è essenziale per definire le responsabilità giuridiche e operative (Capo III).
– L'ambito di applicazione – è necessario verificare se il sistema sia escluso dal campo di applicazione (ad esempio, sistemi destinati esclusivamente a fini militari, di ricerca o uso personale).
– La classificazione del rischio – gli obblighi derivanti dalla classificazione basata sul rischio (*risk-based approach*). I sistemi sono classificati in quattro categorie:
• pratiche vietate (art. 5).
• sistemi ad alto rischio (artt. 6-49 e Allegati I-III).
• sistemi a rischio limitato (art. 50).
• sistemi a rischio minimo o nullo, per i quali non è previsto alcun obbligo specifico.

Nel caso dei sistemi ad alto rischio, la conformità al Regolamento diventa oggetto di valutazione formale, e deve essere completata prima della messa sul mercato o della messa in servizio del sistema.

Questa valutazione mira a dimostrare che il sistema soddisfa i requisiti essenziali previsti dal Capo III, Sezione 2 (affidabilità dei dati, trasparenza, gestione del rischio, sorveglianza umana, robustezza, sicurezza informatica, accuratezza e tracciabilità).

Il processo di valutazione di conformità secondo l'AI Act

La valutazione di conformità rappresenta il processo attraverso il quale un fornitore dimostra che i requisiti per i sistemi di intelligenza artificiale ad alto rischio sono stati pienamente soddisfatti. L'AI Act prevede due principali procedure, la cui scelta o obbligatorietà dipende dalle caratteristiche specifiche del sistema e dall'applicazione di standard armonizzati[2].

stema di IA è fuori dall'obiettivo del presente contributo. La seguente divisione è suggerita da M. Quaranta *et al.*, *AI Compliance Frameworks: A Handbook For Compliance Tools Based on AI Acts' Requirements* (October 02, 2025). Al momento in formato preprint.

[2] Sul ruolo, requisiti, e procedure per verificare la conformità di un sistema di IA, si rimanda a E. Thelisson *et al.*, *Conformity assessment under the EU AI act general approach*, in «AI and

a) Autovalutazione interna (*internal control*)

Questa procedura non richiede l'intervento di un organismo notificato. Il fornitore è responsabile dello svolgimento dell'autovalutazione, che comporta una verifica approfondita del proprio Sistema di Gestione della Qualità (QMS), in conformità con l'articolo 17 dell'AI Act. Il QMS deve garantire la copertura di tutti gli aspetti previsti dal Regolamento, tra cui: conformità normativa, specifiche tecniche, controlli di qualità, testing, gestione dei dati, gestione dei rischi e monitoraggio continuo.

b) Valutazione di terza parte (*third-party assessment*)

Questa procedura, disciplinata dall'articolo 43(1)(b) dell'AI Act, prevede il coinvolgimento di un organismo notificato, al quale il fornitore deve presentare una domanda completa di valutazione del proprio QMS (ex art. 17) e della documentazione tecnica del sistema di IA. L'organismo notificato esegue audit periodici per verificare che il QMS sia mantenuto e applicato correttamente, può effettuare test aggiuntivi sul sistema e, in caso di esito positivo, rilascia un certificato di valutazione valido a livello dell'Unione.

Il processo richiede inoltre che gli operatori economici mantengano una documentazione tecnica completa e aggiornata (art. 11) e applichino un QMS robusto (art. 17), garantendo la tracciabilità di ogni decisione e aggiornamento del sistema. A ciò si aggiunge l'obbligo di registrazione dei sistemi ad alto rischio nella banca dati europea (art. 49) e di sorveglianza post-commercializzazione (art. 72)[3].

Per agevolare l'adempimento di tali obblighi e promuovere l'innovazione regolata, l'AI Act introduce lo strumento delle AI Regulatory Sandboxes (AIRS) (art. 57)[4]. Le sandbox rappresentano ambienti controllati di sperimentazione regolatoria, nei quali fornitori e autorità competenti collaborano alla verifica, sperimentazione e messa a punto di sistemi di IA innovativi in condizioni reali o quasi-reali.

Le sandbox rappresentano dunque un collegamento operativo tra la regolazione sperimentale e la conformità formale, e costituiscono un labo-

Ethics», 4.1 (2024), pp. 113-121; e J. Mökander *et al. Conformity assessments and post-market monitoring: a guide to the role of auditing in the proposed European AI regulation*, in «Minds and Machines» 32.2 (2022), pp. 241-268.

[3] La dottrina sta cominciando a proporre metodologie e strumenti per disegnare un QMS compliant, si veda ad esempio H. Mustroph *et al.*, *Design of a quality management system based on the eu ai act*, in «Legal Knowledge and Information Systems», IOS Press, 2024, pp. 314-320.

[4] Per una contestualizzazione delle AIRS si veda S. Ranchordas, *Experimental regulations for AI: sandboxes for morals and mores*, in «Morals & Machines», 1.1 (2021), pp. 86-100.

ratorio di regulatory learning in cui la conoscenza maturata nel contesto sperimentale può essere tradotta in procedure, standard e prassi conformi al diritto dell'Unione.

Infine, bisogna ricordare che il Regolamento UE sull'intelligenza artificiale (AI Act) introduce un requisito obbligatorio: ciascuno Stato membro deve istituire almeno un sandbox regolatorio per l'IA (operativo entro il 2 agosto 2026)[5]. Questi ambienti controllati consentono ai fornitori di sistemi IA di sviluppare e testare i prototipi sotto la supervisione delle autorità, con l'obiettivo di conciliare innovazione e conformità.

Obiettivi delle sandboxes secondo l'AI Act

Gli articoli 57 e seguenti dell'AI Act delineano gli scopi principali secondo il legislatore europeo, ed in particolare sono[6]:

- Certezza del diritto: aumentare la chiarezza normativa per agevolare la conformità delle imprese.
- Condivisione di *best practice*: favorire lo scambio di esperienze e prassi ottimali tra autorità e operatori.
- Innovazione e competitività: stimolare nuovi servizi IA e rafforzare l'ecosistema europeo dell'IA.
- Apprendimento normativo basato sui dati: raccogliere evidenze concrete dai test per affinare regolamenti e linee guida.
- Accesso accelerato al mercato UE: facilitare l'immissione sul mercato di sistemi IA innovativi (in particolare di PMI e start-up) riducendo gli ostacoli iniziali.

Questi obiettivi si allineano agli scopi di analoghe iniziative nelle fintech e in altri settori, dove i sandbox sono stati introdotti proprio per bilanciare flessibilità applicativa e tutela dei consumatori[7]. Per esempio, nei sandbox in materia di fintech le autorità vigilanti collaborano con banche e start-up

[5] Articolo 57, comma 1.

[6] Tra i tanti, per approfondire si vedano H.J. Allen, *Regulatory sandboxes*, in «Geo. Wash. L. Rev.», 87 (2019), p. 579; e C. Novelli *et al.*, *Getting Regulatory Sandboxes Right: Design and Governance Under the AI Act*, in «Available at SSRN», 5332161 (2025).

[7] Per uno stato delle arte sulle sandbox in materia di fintech, si veda J.J. Goo *et al.*, *The impact of the regulatory sandbox on the fintech industry, with a discussion on the relation between regulatory sandboxes and open innovation*, in «Journal of Open Innovation: Technology, Market, and Complexity», 6.2 (2020), p. 43.

per testare nuovi servizi digitali: i regolatori ottengono informazioni sulle criticità normative delle tecnologie emergenti, mentre gli operatori godono di un ambiente normativo più flessibile durante la sperimentazione.

Ruolo delle sandbox regolatorie nel processo di conformità

Per quanto riguarda nello specifico la valutazione di conformità, le AIRS possono avere un ruolo complementare e di supporto[8]. Come previsto dall'articolo 57, paragrafo 7 dell'AI Act, la partecipazione a una sandbox regolatoria non sostituisce la valutazione di conformità formale, ma la affianca in modo strutturato e complementare. Le autorità competenti forniscono ai fornitori e ai potenziali fornitori orientamento sulle aspettative regolatorie e sulle modalità di adempimento ai requisiti e agli obblighi stabiliti dal Regolamento. Al termine della sperimentazione, esse rilasciano – su richiesta – una prova scritta delle attività completate con successo e un rapporto di uscita che documenta nel dettaglio i test svolti, i risultati ottenuti e gli esiti di apprendimento (*learning outcomes*)[9].

Tali documenti, riconosciuti formalmente dall'articolo 57(7), possono essere utilizzati dai fornitori come evidenze di conformità all'interno del successivo processo di valutazione o nell'ambito delle attività di sorveglianza del mercato. Il medesimo articolo stabilisce che i *Written Proofs* e gli *Exit Reports* devono essere positivamente presi in considerazione dalle autorità di vigilanza del mercato e dagli organismi notificati (*Notified Bodies*), con l'obiettivo di semplificare e accelerare le procedure di valutazione della conformità dei sistemi di IA ad alto rischio[10].

La natura giuridica, secondo il dettame della normativa europea, di questi output è quindi di tipo preparatorio: essi hanno valore probatorio ma non vincolante. La funzione delle sandbox è quindi duplice, operano come

[8] Il legislatore non ha avuto ancora l'occasione di specificare quale sarà il ruolo preciso che le sandbox dovranno avere. In attesa dell'uscita degli standard armonizzati, la dottrina si sta interrogando sul ruolo che questi potranno giocare nel sistema delle sandbox, ad esempio A. Tartaro, E. Panai, *Leveraging technical standards within AI regulatory sandboxes: Challenges and opportunities*, in F. Bagni, F. Seferi (eds), *Regulatory sandboxes for AI and cybersecurity: Questions and answers for stakeholders*, in «National Cybersecurity Lab», 2025, pp. 116-129.

[9] Sempre con riferimento all'articolo 57, comma 7, «The competent authority shall also provide an exit report detailing the activities carried out in the sandbox and the related results and learning outcomes».

[10] Letteralmente «shall be taken positively into account by market surveillance authorities and notified bodies, with a view to accelerating conformity assessment procedures to a reasonable extent».

piattaforme di pre-valutazione, aiutando i fornitori ad allinearsi ai requisiti di conformità, e allo stesso tempo promuovono buone pratiche di sperimentazione regolatoria, favorendo la diffusione di un approccio *compliance-by-design*[11].

L'attività nella sandbox non sostituisce quindi la procedura formale di conformità, ma ne costituisce un passaggio preparatorio dotato di peso documentale significativo, pur privo di valore conclusivo.

Se la sandbox regolatoria fosse interpretata come un'alternativa pienamente sostitutiva alla valutazione di conformità, si correrebbe il rischio di minare l'autorità e le garanzie procedurali proprie del sistema di conformità, oltre a esporre le autorità competenti e gli stakeholder a potenziali responsabilità qualora, in fase di sorveglianza post-mercato, emergesse la non conformità di un sistema precedentemente sperimentato all'interno della sandbox. Pertanto, bisogna domandarsi fino a che punto le sandbox regolatorie devono essere considerate strumenti di supporto e accelerazione, ma non di sostituzione, della valutazione formale.

Funzionamento operativo e buone prassi

Come evidenzia la letteratura, ogni sandbox necessita di obiettivi e parametri di misurazione chiari sin dall'inizio[12]. È importante definire *scopi precisi*, criteri di ingresso/uscita e responsabilità dei partecipanti. Ad esempio, nel processo preparatorio le autorità devono rispondere a domande chiave: quale problema normativo intendono esplorare? Quali stakeholder includere? Che dati utilizzare e come garantire sicurezza e riservatezza? Poiché i sandbox possono coinvolgere più ambiti regolatori (ad es. privacy, sicurezza, leggi settoriali), è spesso necessario associare più autorità competenti[13]. In sintesi, una progettazione trasparente – e comunicata chiaramente agli innovatori – è essenziale affinché l'esperimento produca risultati validi e generalizzabili.

[11] Come già anticipato, questa sembra l'interpretazione più plausibile allo stato attuale della disciplina.

[12] Abbiamo già citato C. Novelli e S. Ranchordas come contributi che chiariscono questo punto.

[13] Il problema è affrontato da S. Ranchordas *et al.*, *Regulatory sandboxes and innovation-friendly regulation: between collaboration and capture*, in «Italian Journal of Public Law», 16 (2024), p. 107.

Il ruolo dei Notified Bodies

Occorre quindi prendere in considerazione il ruolo e le responsabilità degli organismi adibiti a svolgere la funzione di certificatori di conformità secondo l'AI Act, la *Notifying Authority* ed i *Notified Bodies*. Ai sensi dell'articolo 3, punto (21) dell'AI Act, il *conformity assessment body* è definito come «un organismo che svolge attività di valutazione della conformità di terza parte, incluse prove, certificazioni e ispezioni».

Questi organismi notificati (*Notified Bodies* - NB), accreditati secondo il Regolamento (UE) 2019/1020, rappresentano un attore fondamentale nel garantire l'indipendenza e la solidità delle verifiche sui sistemi di IA ad alto rischio. La loro funzione è quindi quella di assicurare che il processo di valutazione sia indipendente, trasparente e ripetibile, fungendo da intermediari tra innovazione tecnologica e tutela dell'interesse pubblico[14].

Il processo di certificazione dei sistemi di intelligenza artificiale ad alto rischio rappresenta l'atto conclusivo del percorso di valutazione di conformità previsto dall'AI Act. Ai sensi della normativa, la certificazione può essere rilasciata unicamente da organismi notificati (i NB) che siano stati accreditati da un'autorità notificante (*notifying authority*) nazionale, la quale ne verifica la competenza, l'indipendenza e la conformità ai requisiti previsti dall'AI Act[15].

È fondamentale chiarire, come precedentemente accennato, che le AI Regulatory Sandboxes non costituiscono in alcun modo un meccanismo di certificazione. Esse possono svolgere una funzione di preparazione alla conformità (*conformity assessment preparation*), fornendo evidenze, analisi e risultati di testing che saranno poi presi in considerazione da un organismo notificato o da un'autorità di vigilanza del mercato, ma non rilasciano né attestati di conformità né certificati di validazione tecnica.

Nel contesto delle sandbox regolatorie, la collaborazione tra autorità competenti e organismi notificati può quindi assumere un valore strategico. Le evidenze raccolte durante la sperimentazione – Written Proof ed Exit Report – possono costituire materiale preliminare di audit e contribuire a ridurre tempi e costi nelle successive fasi di certificazione, a condizione che tali risultati siano riconosciuti o avallati da un NB.

Sebbene l'articolo 31, paragrafo 5 dell'AI Act vieti ai NB di essere direttamente coinvolti nella progettazione, sviluppo, commercializzazione o utilizzo di sistemi di IA ad alto rischio – al fine di preservarne l'indipendenza di giu-

[14] L'AI Act, articoli 28 e seguenti, disciplina ruolo e competenze delle Notifying Authorities e Notifying Bodies

[15] Questa procedura si trova tra gli articoli 28 e 32.

dizio e l'integrità – la loro partecipazione indiretta, in qualità di osservatori tecnici o validatori metodologici, non contrasta con tale principio[16]. In questa prospettiva, i NB potrebbero assumere un ruolo di supervisione o validazione dei protocolli di test adottati nei sandbox, garantendo che le prove sperimentali rispettino criteri di qualità, tracciabilità e rappresentatività conformi agli standard europei armonizzati.

Qualora la partecipazione dei NB fosse formalmente ammessa e regolamentata, le attività svolte nei sandbox acquisirebbero un valore aggiunto all'interno del procedimento di conformità, potendo essere considerate pre-assessment evidence a tutti gli effetti. Anche in assenza di un coinvolgimento diretto, è comunque auspicabile che i test condotti nelle sandbox siano preventivamente avallati o riconosciuti dai NB, in modo da assicurare continuità metodologica tra la fase sperimentale e la successiva valutazione di conformità[17].

Tuttavia, qualora il processo sperimentale di un sandbox sia formalmente accettato da un organismo notificato o da un'autorità notificante, e i test siano condotti dal *main partner* del progetto (ad esempio un *Testing and Experimentation Facility* – TEF), i risultati di tali prove possono essere riconosciuti come input qualificato nel successivo processo di certificazione. In tal caso, il contributo della sandbox diventa una fase di pre-validazione che aumenta l'efficienza della certificazione, riducendo la duplicazione dei test e la ridondanza dei controlli.

Poiché gli organismi notificati operano sul mercato come soggetti economici, l'inclusione di questi nel ciclo di sperimentazione pre-market pone problemi di natura economica, soprattutto in ambito di concorrenza e sostenibilità. Gli NB hanno il diritto esclusivo di condurre le attività di valutazione di terza parte, ma se vengono coinvolti in fase anticipata nei sandbox – offrendo servizi non gratuiti o selettivi – vi è il rischio di distorsione della concorrenza[18]. Per evitare asimmetrie competitive tra NB, la partecipazione deve essere gestita attraverso criteri di trasparenza e rotazione, e possibilmente in coordinamento con le autorità notificanti nazionali e la Commissione europea.

[16] Almeno questo è quanto si può dedurre se si presuppone un'assenza di un conflitto di interessi, ad esempio in assenza in una remunerazione per l'attività svolta che proviene dai provider e dai partecipanti stessi nelle sandbox.

[17] Anche quando questi test non fossero quindi portati a termine dai NB stessi.

[18] Pensiamo, ad esempio, al caso in cui non vengano applicati criteri trasparenti di selezione o accesso: in tale situazione, gli organismi notificati coinvolti all'interno di una sandbox potrebbero ottenere un vantaggio competitivo ingiustificato rispetto a quelli esterni. Ciò contrasterebbe con i principi di indipendenza, imparzialità e parità di trattamento stabiliti dagli articoli 31 e 32 dell'AI Act, che richiedono ai NB di operare senza legami economici o interessi che possano compromettere la neutralità delle valutazioni di conformità.

Una possibile soluzione, ancora da chiarire sul piano giuridico, potrebbe arrivare dall'atto di esecuzione (*implementing act*) ex articolo 58 dell'AI Act, se quest'ultimo disciplinerà:

– le modalità con cui gli NB possono essere coinvolti nei sandbox senza violare il principio di neutralità del mercato.
– i limiti alle tariffe o ai costi che possono essere applicati in fase sperimentale.
– le regole di accesso non discriminatorio per tutti gli NB interessati, anche con riferimento ai settori verticali (sanità, mobilità, finanza, sicurezza).

Dal punto di vista economico, emerge inoltre un problema di sostenibilità: se un fornitore paga sia per i servizi sandbox (pre-market testing) sia per la certificazione formale da parte dell'NB, si genera un rischio di doppio pagamento, che potrebbe disincentivare la partecipazione alle sandbox e aumentare i costi complessivi di accesso al mercato.

Ciò potrebbe perturbare il mercato e innalzare i costi di transazione per i provider di IA, soprattutto per le PMI, riducendo la competitività del sistema europeo rispetto ad altri contesti regolatori meno onerosi. Questo contributo intende anche domandarsi se questa concezione di divisione delle responsabilità tra sandbox e NB possa venir superata dagli interessi economici e di sviluppo tecnologico che regolano il mondo dell'Intelligenza Artificiale.

Il progetto EUSAiR

Ora esaminiamo come questa sfida viene affrontata a livello operativo, e in particolare come il progetto EUSAiR[19] sta contribuendo a sostenere e coordinare gli sforzi dell'Unione Europea nello sviluppo delle sandbox regolatorie per l'intelligenza artificiale.

EUSAiR (European Sandboxes for AI Regulation) è un'iniziativa biennale finanziata dal programma Digital Europe che mira a creare un quadro armonizzato per la sperimentazione regolatoria in ambito IA. Il progetto supporta gli Stati membri nell'attuazione delle sandbox, fornendo modelli e strumenti comuni per la documentazione, la valutazione e il monitoraggio dei progetti, al fine di favorire l'innovazione, la certezza e la conformità all'AI Act.

[19] Si rimanda al sito del progetto https://eusair-project.eu/.

Aspetti operativi e tecnici

Dal punto di vista operativo, tre elementi risultano determinanti per massimizzare l'utilità dei risultati prodotti dalle sandbox regolatorie[20]. In primo luogo, il livello di maturità tecnologica del sistema, o Technology Readiness Level (TRL), incide direttamente sulla solidità delle evidenze raccolte e sulla preparazione alla conformità: sistemi più maturi offrono esiti più affidabili e immediatamente utilizzabili nel percorso regolatorio, mentre quelli in fase iniziale traggono maggiore beneficio dal confronto consulenziale e formativo. In secondo luogo, un approccio modulare alla conformità consente di impiegare la sandbox sia come strumento pre-certificativo dedicato a specifici requisiti, sia come spazio di valutazione preliminare più ampia e olistica, utile alla pianificazione delle verifiche successive. Infine, l'adozione di standard armonizzati e di formati interoperabili, come previsto dagli articoli 17 e 40 dell'AI Act, garantisce coerenza documentale e facilita la collaborazione tra autorità nazionali, organismi notificati e operatori economici.

Un ulteriore aspetto chiave riguarda i test in condizioni reali, riconosciuti dall'articolo 60 dell'AI Act come strumento essenziale per verificare l'efficacia dei controlli tecnici e procedurali in contesti rappresentativi dell'uso effettivo dei sistemi di IA. Nel quadro delle Testing and Experimentation Facilities e delle AI Regulatory Sandboxes, queste attività consentono di valutare il grado di conformità ai requisiti normativi e di misurare la maturità regolatoria dei partecipanti attraverso modelli di avanzamento progressivo. Il progetto EUSAiR ha predisposto un framework di sandbox multilivello, un primo basato sulle attività base di supporto documentale ai provider, ed un secondo, denominato AIRS Add-ons, che offre inoltre servizi specializzati per il supporto alla conformità normativa, alla sicurezza informatica e alla protezione dei dati, integrando la collaborazione con le Autorità per la Protezione dei Dati e con gli organismi di valutazione della conformità. Per garantire la qualità metodologica e la credibilità dei risultati, è necessario il coinvolgimento di esperti altamente qualificati, tra cui specialisti di protezione dei dati, coordinatori di test in ambienti reali e professionisti della conformità dotati di competenze specifiche sugli standard armonizzati e sulle metodologie di valutazione.

All'interno di questo quadro operativo, due strumenti documentali assumono un ruolo centrale: la Written Proof e l'Exit Report[21]. La Written

[20] Per i risultati del progetto si rimanda ai deliverable, che presto saranno pubblici.

[21] I seguenti modelli di Exit Report e Written Proof preparati da EUSAiR vanno a com-

Proof, rilasciata dall'autorità competente al termine della partecipazione, costituisce la prova ufficiale delle attività svolte e dei risultati conseguiti. Essa descrive in modo sintetico le attività completate, i requisiti normativi affrontati e lo stato finale del progetto rispetto agli obiettivi fissati nel piano della sandbox. Il progetto EUSAiR ha sviluppato un modello standardizzato di Written Proof per garantire coerenza tra Stati membri e facilitare l'utilizzo di questo documento all'interno dei processi di conformità e sorveglianza di mercato.

L'Exit Report, invece, rappresenta la documentazione più estesa e analitica dell'intera esperienza in sandbox. Esso riassume i servizi di consulenza e supporto forniti, le sfide regolatorie incontrate, la valutazione dei risultati ottenuti rispetto agli indicatori di performance, nonché eventuali incidenti, reclami o segnalazioni registrate durante la sperimentazione. Include infine una riflessione sugli insegnamenti tratti (*lesson learned)* e sui contributi forniti all'evoluzione delle pratiche regolatorie europee. Secondo quanto previsto dall'articolo 57, paragrafo 7, dell'AI Act, le autorità nazionali dovrebbero fornire la versione destinata ai partecipanti entro un mese dalla conclusione della sandbox e quella destinata alla diffusione pubblica entro due mesi. Il modello operativo elaborato da EUSAiR per la redazione di questi report contribuisce a rendere omogenee le procedure documentali e a consolidare la base probatoria utile per la successiva valutazione di conformità.

Written Proof

La Written Proof ha una duplice natura: da un lato, costituisce un documento tecnico-normativo che registra in modo sintetico le attività e i risultati conseguiti durante la sandbox; dall'altro, è uno strumento di trasferimento di conoscenze e di raccordo tra la sperimentazione e la valutazione formale di conformità.

Il modello elaborato da EUSAiR si articola in cinque sezioni principali. La Parte A, dedicata alle informazioni generali, identifica il partecipante, la durata e il riferimento amministrativo della sperimentazione, garantendo la tracciabilità amministrativa del progetto. La Parte B descrive in modo analitico le attività tecniche condotte – sviluppo, addestramento, validazione e test –

plementare quanto descritto nella bozza dell'atto di implementazione (https://digital-strategy.ec.europa.eu/en/consultations/commission-seeks-feedback-draft-implementing-act-establish-ai-regulatory-sandboxes-under-ai-act).

fornendo la base empirica per la valutazione della robustezza del sistema di IA e per la successiva compilazione del fascicolo tecnico. La Parte C è dedicata alle questioni regolatorie affrontate e ai requisiti dell'AI Act oggetto di analisi o verifica, documentando le aree di incertezza interpretativa chiarite durante la sperimentazione e contribuendo all'apprendimento istituzionale.

La Parte D riassume lo stato di avanzamento del progetto, con una valutazione sintetica dell'idoneità del sistema a proseguire verso la fase di conformità formale. Questa sezione rappresenta un punto di raccordo tra la dimensione sperimentale e quella regolatoria, offrendo alle autorità elementi oggettivi per stimare la maturità regolatoria del partecipante. Infine, la Parte E raccoglie gli elementi formali di approvazione e sottoscrizione, assicurando la validità e l'autenticità del documento ai fini del riconoscimento ufficiale.

La Written Proof, così strutturata, assolve una funzione di ponte tra sperimentazione e conformità. Per i fornitori, costituisce una prova della solidità del proprio sistema di gestione della qualità e del rischio; per le autorità, rappresenta un documento interoperabile che può essere utilizzato come evidenza tecnica nel processo di conformity assessment, in linea con l'articolo 57, paragrafo 7, dell'AI Act. La sua efficacia dipende dal mantenimento di un equilibrio costante tra valore probatorio e valore conoscitivo: se intesa unicamente come certificazione preliminare, ridurrebbe la sandbox a un audit; se considerata solo come esercizio di apprendimento, perderebbe la propria rilevanza regolatoria. Trovare la giusta complementarità tra queste due dimensioni sarà oggetto di test nel corso del progetto.

Exit Report

L'Exit Report rappresenta il secondo asse documentale della fase post-sperimentazione nelle sandbox e ha una funzione più estesa rispetto alla Written Proof. È concepito per raccogliere, interpretare e sistematizzare tutte le evidenze emerse durante la partecipazione alla sandbox, offrendo una visione completa del percorso sperimentale e dei risultati ottenuti.

Anche in questo caso, il modello sviluppato da EUSAiR adotta una struttura articolata. La Parte A include le informazioni generali, garantendo la tracciabilità amministrativa e temporale. La Parte B descrive i servizi e il supporto forniti all'interno della sandbox – inclusi consulenza regolatoria, accesso a infrastrutture europee e assistenza tecnica – documentando il grado di interazione tra autorità e partecipante. La Parte C riassume le attività

svolte e i risultati raggiunti, con particolare attenzione ai rischi gestiti nelle fasi di sviluppo, addestramento e test, e quindi fornisce elementi di prova sulla sicurezza e affidabilità del sistema.

La Parte D analizza le sfide regolatorie incontrate e le soluzioni adottate, delineando un quadro delle barriere normative affrontate e delle strategie di adattamento sperimentate. Questa sezione contribuisce direttamente alla chiarificazione dei requisiti regolatori, offrendo un patrimonio di conoscenze utile sia alle autorità sia ai legislatori. La Parte E valuta le performance del progetto sulla base di indicatori chiave, misurando il rispetto dei tempi, la qualità dei risultati e la coerenza con gli obiettivi progettuali: tali elementi costituiscono un riferimento quantitativo per la verifica di conformità.

La Parte F raccoglie eventuali incidenti, reclami o malfunzionamenti, insieme alle azioni correttive intraprese. La loro documentazione introduce una dimensione di *learning by error* che anticipa, in ambiente controllato, il meccanismo di monitoraggio post-market. La Parte G sintetizza le *lesson learned* e gli insegnamenti regolatori, costituendo il punto di contatto tra la sperimentazione e la normazione: qui l'esperienza del singolo progetto si trasforma in conoscenza sistemica utile al miglioramento del quadro normativo europeo. La Parte H, infine, contiene gli elementi formali di approvazione e sottoscrizione, che assicurano la validità istituzionale del documento e ne consentono l'utilizzo ufficiale ai fini della conformità.

Grazie a questa articolazione, l'Exit Report svolge una funzione di raccordo strutturale tra apprendimento e conformità. Da un lato, fornisce alle autorità un archivio tecnico e organizzativo che semplifica le verifiche successive, riducendo i tempi di valutazione e migliorando la qualità del processo di *conformity assessment*. Dall'altro, alimenta il *regulatory learning*, consentendo di individuare aree di ambiguità normativa e di orientare la futura evoluzione degli standard armonizzati e delle prassi di supervisione.

Complementarità con audit e monitoraggio post-commercializzazione

Nel quadro complessivo di vigilanza, il percorso di sandbox si collega sempre alla logica *ex-ante* di valutazione di conformità. Come già affermato, l'AI Act conferma che i sistemi IA ad alto rischio potranno essere immessi sul mercato solo se sottoposti ed aver passato con successo una valutazione di conformità (interno o con organismo notificato). Tuttavia, la fase successiva all'entrata in commercio prevede un monitoraggio post-mercato che integra e completa la conformità iniziale.

Dal punto di vista dei controlli, la letteratura distingue tre approcci di auditing IA[22]: *auditing di funzionalità* (verifica degli scopi e delle motivazioni d'uso del sistema), *auditing del codice* (analisi del software sottostante) e *auditing d'impatto* (indagine sugli effetti reali del sistema). Le procedure di conformità previste dall'AI Act combinano elementi di auditing di funzionalità e di codice (es. analisi della documentazione, test di robustness, verifiche di governance). Il monitoraggio post-commercializzazione, per sua natura, aggiunge invece la dimensione dell'audit d'impatto (valutazione di outcome effettivi, bias, rischi emergenti). Questo è particolarmente importante per sistemi IA che continuano ad apprendere o a evolversi dopo il rilascio: il post-market monitoring rileva eventuali derive non previste e alimenta interventi correttivi basati sui dati reali di utilizzo. In altre parole, la fase dopo il sandbox e l'entrata in servizio non diminuisce le responsabilità del fornitore, ma garantisce che l'aderenza alle norme rimanga costante lungo tutto il ciclo di vita del sistema.

I sistemi di intelligenza artificiale ad alto rischio devono inoltre rispettare rigorosi requisiti di cybersecurity, come previsto dall'AI Act e in linea con le indicazioni di ENISA e degli standard europei emergenti. La sicurezza informatica è un elemento critico del conformity assessment, influenzando direttamente la robustezza, l'affidabilità e la tracciabilità dei sistemi IA.

All'interno dei sandbox regolatori, la verifica dei requisiti di cybersecurity può essere condotta in modalità simulata o controllata, consentendo ai fornitori di identificare vulnerabilità, testare misure di mitigazione e integrare protocolli di protezione già nella fase di sviluppo. Tuttavia, tali test preliminari non sostituiscono le attività formali di audit dei NB, che rimangono necessarie per il rilascio della certificazione.

Un aspetto innovativo è l'opportunità di creare un regulatory learning per gli NB in materia di cybersecurity. Partecipando a sandbox che simulano attacchi informatici, gestione di incidenti e resilienza dei sistemi, gli organismi notificati possono acquisire familiarità con le minacce emergenti, aggiornare le checklist di valutazione e affinare le linee guida per la certificazione. Questo processo può ridurre i tempi di audit e aumentare la qualità complessiva delle valutazioni, contribuendo a un ecosistema più sicuro e affidabile.

[22] In questo caso si fa riferimento soprattutto agli approcci previsti da J. Mökander, *Auditing of AI: legal, ethical and technical approaches*, in «Digital Society» 2.3 (2023), p. 49 e D. Hartmann *et al.*, *Addressing the regulatory gap: moving towards an EU AI audit ecosystem beyond the AI Act by including civil society*, in «AI and Ethics», 5.4 (2025), pp. 3617-3638.

Dal punto di vista operativo, è essenziale definire una roadmap integrata tra sandbox, TEF e organismi notificati per, in primis, garantire che i test di cybersecurity eseguiti in ambiente controllato siano statisticamente rappresentativi dei test richiesti per la conformità; ed in secundis, allineare i criteri di vulnerabilità e mitigazione agli standard europei armonizzati.

In pratica, un sistema IA che ha superato con successo i test di cybersecurity all'interno del sandbox, e che mantiene condizioni di *compliance* con la documentazione tecnica richiesta, sarà altamente predisposto a superare la successiva valutazione di conformità formale. Questo approccio riduce i rischi di errori procedurali, minimizza i costi di ripetizione dei test e contribuisce alla sostenibilità del processo di certificazione, senza generare conflitti con i requisiti di mercato o duplicazioni economiche.

In sintesi, l'integrazione dei test di cybersecurity nei sandbox regolatori costituisce un collegamento operativo e cognitivo tra sperimentazione e certificazione formale, rafforzando sia la robustezza tecnica dei sistemi IA sia l'efficacia complessiva del compliance ecosystem europeo.

Una delle criticità più rilevanti è la frammentazione istituzionale. L'idea di un «*one-stop shop*» per la valutazione della conformità dell'IA appare difficilmente realizzabile nel breve termine: i partner coinvolti – NB, ENISA, TEF, autorità nazionali, AI Office – operano con mandati distinti e competenze differenziate.

Per mitigare questi rischi e promuovere una cooperazione, è auspicabile che le sandbox stabiliscano meccanismi di coordinamento stabile con:

- gli organismi notificati (NB), per l'allineamento metodologico e la mutua riconoscibilità dei risultati di test.
- ENISA, per la definizione dei requisiti di cybersicurezza dei sistemi di IA (ancora in fase di consolidamento).
- i TEF settoriali, per la conduzione di prove tecniche in condizioni reali e l'utilizzo di dataset di riferimento.
- le autorità nazionali competenti (NCA) e l'AI Office, per assicurare legittimità istituzionale.

Tale cooperazione non si esaurisce nella sola raccolta di evidenze tecniche, ma può dar vita a un vero e proprio regulatory learning condiviso. Anche gli NB possono trarre beneficio da questa interazione, acquisendo familiarità con le nuove categorie di sistemi di IA, le metriche di rischio, le architetture di rete neurale o i modelli general-purpose (GPAI). In tal senso, la sandbox non è solo un *testing ground* per le imprese, ma anche un

laboratorio di apprendimento regolatorio per le autorità e gli organismi di valutazione della conformità.

La cooperazione deve quindi basarsi su protocolli di interoperabilità e formati standardizzati di reporting (ad esempio per *Written Proof* ed *Exit Report*), che consentano di ridurre la distanza tra test in ambiente controllato e valutazione formale.

Conclusioni

I sandbox regolatori per l'IA incarnano un approccio agile di compliance by design: integrando supporto normativo e verifica tecnica in un contesto controllato, aiutano i fornitori ad allineare precocemente i propri sistemi ai requisiti UE. La disciplina dell'AI Act valorizza questa esperienza: le prove documentali dei sandbox (report e attestati) sono ufficialmente riconosciute come evidenza di conformità, con vantaggi concreti nel successivo processo di certificazione. Nel complesso, il modello sandbox consolida la sinergia tra innovazione e regolazione: da un lato facilita l'accesso precoce al mercato europeo, e dall'altro offre alle autorità insight pratici per ottimizzare le norme. In tal modo, si costruisce gradualmente un ecosistema di audit e vigilanza IA più informato e partecipato, capace di gestire in modo sostenibile i rischi delle tecnologie emergenti.

Con l'entrata in piena applicazione dell'AI Act nel 2026, la valutazione di conformità diventerà un asse portante della governance tecnologica europea. L'integrazione tra sandbox, organismi notificati e autorità competenti segnerà il passaggio da un modello puramente reattivo a uno proattivo e cooperativo.

In tale contesto, le sandbox regolatorie non solo accelerano la certificazione dei sistemi ad alto rischio, ma contribuiscono a un ecosistema di fiducia multilivello, in cui la conformità è il risultato di un processo di apprendimento reciproco tra industria, ricerca e regolazione.

English title: Conformity assessment under article 57 AI Act: Synergy with Regulatory Sandboxes

Abstract

This article examines the evolving architecture of conformity assessment under the EU Artificial Intelligence Act (AI Act), focusing on the interaction between regulatory sandboxes, notified bodies, and the formal certification process. It explores how Written Proofs and Exit Reports, developed within sandbox environments, serve as structured documentation bridging experimental testing and market conformity procedures. The analysis highlights the dual nature of these instruments: as technical evidence supporting compliance and tools for regulatory learning that inform institutional guidance and standards development. Particular attention is given to the role of notified bodies, their potential engagement within sandbox testing under Article 31(5), and the risks of overlap or market distortion. The study argues for a balanced integration of sandbox outcomes into conformity workflows, proposing a cooperative model where regulatory experimentation closely interacts with conformity procedures.

Keywords: regulatory sandboxes; conformity assessment; notified bodies; exit report; cybersecurity.

Marco Billi
University of Bologna
marco.billi3@unibo.it

Antonino Rotolo
University of Bologna
antonio.rotolo@unibo.it

La presente ricerca è stata svolta nell'ambito del progetto EUSAIR. Le opinioni espresse nel presente contributo sono esclusivamente degli autori e non riflettono necessariamente la posizione dell'Unione Europea, che non può essere ritenuta responsabile dei contenuti qui riportati. Grant Agreement n. 101195535.