



Il presente documento è pubblicato sotto la responsabilità del Segretario generale dell'OCSE. Esso riproduce uno strumento giuridico dell'OCSE e può contenere materiale aggiuntivo. Le opinioni espresse e le argomentazioni utilizzate nel materiale aggiuntivo non riflettono necessariamente le posizioni ufficiali dei paesi membri dell'OCSE.

Il presente documento, così come tutti i dati e le mappe in esso contenuti, non pregiudica lo status o la sovranità su alcun territorio, la delimitazione delle frontiere e dei confini internazionali e il nome di alcun territorio, città o area.

Per accedere ai testi ufficiali e aggiornati degli strumenti giuridici dell'OCSE, nonché ad altre informazioni correlate, si prega di consultare il Compendio degli strumenti giuridici dell'OCSE all'indirizzo <http://legalinstruments.oecd.org>.

Si prega di citare questo documento come:

OCSE, *Raccomandazione del Consiglio sull'Intelligenza Artificiale*, OCSE/LEGAL/0449

Serie: Strumenti giuridici dell'OCSE

Crediti fotografici: © kras99/Shutterstock.com

© OCSE 2025

Questo documento è fornito gratuitamente. Può essere riprodotto e distribuito gratuitamente senza necessità di ulteriori autorizzazioni, purché non venga modificato in alcun modo. Non può essere venduto.

Il presente documento è disponibile nelle due lingue ufficiali dell'OCSE (inglese e francese). Può essere tradotto in altre lingue, a condizione che la traduzione sia etichettata come "traduzione non ufficiale" e includa la seguente dichiarazione di non responsabilità: "*Questa traduzione è stata preparata da [NOME DELL'AUTORE DELLA TRADUZIONE] a solo scopo informativo e la sua accuratezza non può essere garantita dall'OCSE. Le uniche versioni ufficiali sono i testi in inglese e francese disponibili sul sito web dell'OCSE [http:// legalinstruments.oecd.org](http://legalinstruments.oecd.org)".*

Informazioni di base

La Raccomandazione sull'Intelligenza Artificiale (IA) (di seguito la "Raccomandazione") – il primo standard intergovernativo sull'IA – è stata adottata dal Consiglio dell'OCSE, riunitosi a livello ministeriale il 22 maggio 2019, su proposta del Comitato per le Politiche Digitali (DPC, precedentemente Comitato per le Politiche dell'Economia Digitale, CDEP). La Raccomandazione mira a promuovere l'innovazione e la fiducia nell'IA, favorendo una gestione responsabile di un'IA affidabile e garantendo al contempo il rispetto dei diritti umani e dei valori democratici. Nel giugno 2019, al Vertice di Osaka, i leader del G20 hanno accolto con favore i Principi del G20 sull'IA, derivati dalla Raccomandazione.

La Raccomandazione è stata rivista dal Consiglio dell'OCSE l'8 novembre 2023 per aggiornare la definizione di "sistema di IA", al fine di garantire che la Raccomandazione continui ad essere tecnicamente accurata e a riflettere gli sviluppi tecnologici, anche per quanto riguarda l'IA generativa. Sulla base della Relazione al Consiglio del 2024 sulla sua attuazione, diffusione e continua rilevanza, la Raccomandazione è stata rivista dal Consiglio dell'OCSE, riunitosi a livello ministeriale il 3 maggio 2024, per riflettere gli sviluppi tecnologici e politici, anche per quanto riguarda l'IA generativa, e per facilitarne ulteriormente l'attuazione.

Il lavoro dell'OCSE sull'intelligenza artificiale

L'intelligenza artificiale (IA) è una tecnologia di uso generale che ha il potenziale per: migliorare il benessere delle persone, contribuire a un'attività economica globale positiva e sostenibile, aumentare l'innovazione e la produttività e aiutare a rispondere alle principali sfide globali. Viene impiegata in molti settori, dalla produzione, all'istruzione, alla finanza e ai trasporti, fino alla sanità e alla sicurezza.

Oltre ai benefici, l'intelligenza artificiale pone anche delle sfide per le nostre società e le nostre economie, in particolare per quanto riguarda i cambiamenti e le disuguaglianze economiche, la concorrenza, le transizioni nel mercato del lavoro e le implicazioni per la democrazia e i diritti umani.

Dal 2016, l'OCSE ha intrapreso attività empiriche e politiche sull'IA a sostegno del dibattito politico, a partire dal Technology Foresight Forum sull'IA tenutosi nello stesso anno, seguito dalla conferenza internazionale sull'IA: Macchine intelligenti, politiche intelligenti nel 2017. L'Organizzazione ha inoltre svolto attività di analisi e misurazione che forniscono una panoramica del panorama tecnico dell'IA, mappano gli impatti economici e sociali delle tecnologie di IA e delle loro applicazioni, individuano le principali considerazioni politiche e descrivono le iniziative in materia di IA promosse da governi e altri soggetti interessati a livello nazionale e internazionale.

Questo lavoro ha dimostrato la necessità di delineare un quadro politico stabile a livello internazionale per promuovere la fiducia e l'adozione dell'IA nella società. In questo contesto, il Consiglio dell'OCSE ha adottato, su proposta del DPC, una Raccomandazione per promuovere un approccio all'IA affidabile incentrato sull'uomo, che favorisca la ricerca, preservi gli incentivi economici all'innovazione e si applichi a tutte le parti interessate.

Un processo inclusivo e partecipativo per l'elaborazione della Raccomandazione

Lo sviluppo della Raccomandazione è stato di natura partecipativa, incorporando contributi provenienti da un'ampia gamma di fonti durante tutto il processo. Nel maggio 2018, il Comitato per la protezione dei dati (DPC) ha concordato di costituire un gruppo di esperti per definire i principi volti a promuovere la fiducia e l'adozione dell'IA, con l'obiettivo di elaborare una bozza di Raccomandazione nel corso del 2019. Successivamente è stato istituito il Gruppo informale di esperti sull'IA presso l'OCSE, composto da oltre 50 esperti provenienti da diverse discipline e settori (governo, industria, società civile, sindacati, comunità tecnica e mondo accademico) - si veda <http://www.oecd.org/going-digital/ai/oecd-aigo-membership-list.pdf> per l'elenco completo. Tra settembre 2018 e febbraio 2019 il gruppo ha tenuto quattro riunioni. Il lavoro ha beneficiato della diligenza, dell'impegno e dei contributi sostanziali degli esperti partecipanti al gruppo, nonché delle loro esperienze multidisciplinari e provenienti da diversi ambiti di competenza.

Sulla base del documento finale del gruppo informale, è stata elaborata una bozza di raccomandazione in seno al DPC, previa consultazione con altri organi competenti dell'OCSE, e approvata in una riunione speciale il 14-15 marzo 2019. Il Consiglio dell'OCSE ha adottato la raccomandazione nella sua riunione a livello ministeriale del 22-23 maggio 2019.

Ambito di applicazione della raccomandazione

A complemento degli standard OCSE già esistenti e pertinenti all'IA, come quelli in materia di privacy e protezione dei dati, gestione del rischio per la sicurezza digitale e condotta aziendale responsabile, la Raccomandazione si concentra su questioni politiche specifiche dell'IA e si propone di definire uno standard attuabile e sufficientemente flessibile da resistere alla prova del tempo in un settore in rapida evoluzione. La Raccomandazione contiene cinque principi di alto livello basati su valori e cinque raccomandazioni per le politiche nazionali e la cooperazione internazionale. Propone inoltre una comprensione comune di termini chiave, come "sistema di IA", "ciclo di vita del sistema di IA" e "attori dell'IA", ai fini della Raccomandazione stessa.

Più nello specifico, la raccomandazione comprende due sezioni sostanziali:

1. **Principi per una gestione responsabile dell'IA affidabile:** la prima sezione definisce cinque principi complementari rilevanti per tutte le parti interessate: i) crescita inclusiva, sviluppo sostenibile e benessere; ii) rispetto dello stato di diritto, dei diritti umani e dei valori democratici, inclusi equità e privacy; iii) trasparenza e spiegabilità; iv) robustezza, sicurezza e protezione; e v) responsabilità. Questa sezione invita inoltre gli attori dell'IA a promuovere e attuare questi principi in base ai loro ruoli.
2. **Politiche nazionali e cooperazione internazionale per un'IA affidabile:** in linea con i cinque principi sopracitati, la seconda sezione fornisce cinque raccomandazioni ai Membri e ai non Membri che hanno aderito alla Raccomandazione (di seguito "gli Aderenti") da attuare nelle loro politiche nazionali e nella cooperazione internazionale: i) investire nella ricerca e nello sviluppo dell'IA; ii) promuovere un ecosistema inclusivo che favorisca l'IA; iii) plasmare un contesto di governance e politico interoperabile e favorevole all'IA; iv) sviluppare le capacità umane e prepararsi alla trasformazione del mercato del lavoro; e v) cooperare a livello internazionale per un'IA affidabile.

Revisioni della raccomandazione per il 2023 e il 2024

Nel 2023, è stata individuata un'opportunità per mantenere la rilevanza della Raccomandazione aggiornando la sua definizione di "sistema di IA", e il DPC ha approvato una bozza di definizione rivista in una sessione congiunta del Comitato e del suo Gruppo di lavoro sulla governance dell'IA (AIGO) il 16 ottobre 2023. Il Consiglio dell'OCSE ha adottato la definizione rivista di "sistema di IA" nella sua riunione dell'8 novembre 2023. L'aggiornamento della definizione includeva modifiche volte a:

- chiarire gli obiettivi di un sistema di intelligenza artificiale (che possono essere espliciti
- impliciti);
- sottolineare il ruolo dell'input che può essere fornito da esseri umani o macchine;
- chiarire che la Raccomandazione si applica ai sistemi di intelligenza artificiale generativa, che producono "contenuti";
- sostituire la parola "reale" con "fisico" per chiarezza e allineamento con altri processi internazionali; riflettere sul fatto che alcuni sistemi di intelligenza artificiale possono continuare ad evolversi dopo la loro progettazione e implementazione.

In linea con le conclusioni della [relazione al Consiglio del 2024](#), la Raccomandazione è stata ulteriormente rivista, in occasione della riunione del Consiglio del 2024, a livello ministeriale per mantenerne la continua rilevanza e facilitarne l'attuazione cinque anni dopo la sua adozione. Gli aggiornamenti specifici miravano a:

- riflettere sulla crescente importanza di affrontare la disinformazione e la misinformazione e di salvaguardare l'integrità delle informazioni nel contesto dell'IA generativa;
- affrontare gli usi al di fuori dello scopo previsto, l'uso improprio intenzionale o l'uso improprio non intenzionale;
- chiarire le informazioni che gli attori dell'IA dovrebbero fornire in merito ai sistemi di IA per garantire trasparenza e divulgazione responsabile;
- affrontare i problemi di sicurezza, in modo da poter sovrascrivere, riparare e/o dismettere in sicurezza, tramite interazione umana, i sistemi di IA che rischiano di causare danni indebiti o mostrano comportamenti indesiderati;
- enfatizzare una condotta aziendale responsabile durante tutto il ciclo di vita del sistema di IA, che coinvolge la cooperazione con i fornitori di conoscenza e risorse di IA, gli utenti del sistema di IA e altri stakeholder, sottolineando la necessità che le giurisdizioni collaborino per promuovere ambienti politici e di governance interoperabili per l'IA, contro l'aumento delle iniziative politiche sull'IA in tutto il mondo e introducendo un riferimento esplicito alla sostenibilità ambientale, di cui l'importanza è cresciuta considerevolmente dall'adozione della Raccomandazione nel 2019.

Inoltre, alcune delle intestazioni dei principi e delle raccomandazioni sono state ampliate per maggiore chiarezza, e il testo sulla tracciabilità e la gestione del rischio è stato ulteriormente elaborato e spostato nel principio "Responsabilità" in quanto principio più appropriato per questi concetti.

Per ulteriori informazioni, consultare il sito oecd.ai. Informazioni di contatto: ai@oecd.org

Implementazione

La raccomandazione impone al DPC di riferire al Consiglio in merito alla sua attuazione, diffusione e continua rilevanza cinque anni dopo la sua adozione e successivamente con cadenza regolare.

Relazione al Consiglio del 2024

Il DPC, tramite AIGO, ha elaborato un [rapporto](#) al Consiglio in merito all'attuazione, alla diffusione e alla continua rilevanza della Raccomandazione a cinque anni dalla sua attuazione, nonché alle proposte di revisione basate sulle sue conclusioni.

Il Rapporto del 2024 ha concluso che la Raccomandazione fornisce un riferimento internazionale significativo e utile per l'elaborazione delle politiche nazionali in materia di IA. La Raccomandazione viene attuata dai Paesi aderenti, è ampiamente diffusa e rimane pienamente rilevante, anche come solido quadro di riferimento per analizzare le evoluzioni tecnologiche, come quelle relative all'IA generativa.

Tuttavia, il Rapporto del 2024 ha rilevato la necessità di aggiornamenti per chiarire la sostanza di alcune disposizioni della Raccomandazione, per facilitarne l'attuazione, aumentarne la rilevanza e garantire che la Raccomandazione rifletta gli importanti sviluppi tecnologici, anche per quanto riguarda l'intelligenza artificiale generativa.

Ulteriori attività a supporto dell'attuazione della raccomandazione.

Oltre a riferire al Consiglio sull'attuazione della Raccomandazione, il DPC è anche incaricato di proseguire il proprio lavoro sull'IA, basandosi su questa Raccomandazione e tenendo conto del lavoro svolto in altri forum internazionali, come l'UNESCO, l'Unione europea, il Consiglio d'Europa e l'iniziativa per la creazione di un Panel internazionale sull'IA (vedere <https://pm.gc.ca/eng/news/2018/12/06/mandate-international-panel-artificial-intelligence> e <https://www.gouvernement.fr/en/france-and-canada-create-new-expert-international-panel-on-artificial-intelligence>).

Al fine di sostenere l'attuazione della Raccomandazione, il Consiglio ha incaricato il DPC di elaborare linee guida pratiche per l'attuazione, di fornire un forum per lo scambio di informazioni sulle politiche e le attività in materia di IA e di promuovere un dialogo interdisciplinare e multilaterale.

Al fine di fornire un forum inclusivo per lo scambio di informazioni sulle politiche e le attività relative all'IA e per promuovere un dialogo interdisciplinare e multistakeholder, l'OCSE ha lanciato nel febbraio 2020 i) [l' Osservatorio sulle politiche in materia di IA](#) (OECD.AI) e ii) la rete informale di esperti dell'OCSE sull'IA (ONE AI).

OECD.AI è un hub inclusivo per le politiche pubbliche sull'IA che mira ad aiutare i paesi a incoraggiare, promuovere e monitorare lo sviluppo responsabile di sistemi di intelligenza artificiale affidabili a beneficio della società. Unisce le risorse di tutta l'OCSE con quelle dei partner di tutti i gruppi di stakeholder per fornire analisi politiche multidisciplinari e basate su dati concreti in materia di IA. L'Osservatorio include un database in tempo reale di strategie, politiche e iniziative sull'IA che i paesi e altri stakeholder possono condividere e aggiornare, consentendo il confronto interattivo dei

loro elementi chiave. Viene costantemente aggiornato con metriche, misurazioni, politiche e buone pratiche sull'IA che portano a ulteriori aggiornamenti nelle linee guida pratiche per l'implementazione. La rete di esperti sull'IA dell'OCSE (ONE AI) è un gruppo informale di esperti di IA provenienti da enti governativi, imprese, mondo accademico e società civile che fornisce competenze e consulenza politica specifiche sull'IA al OCSE. La rete fornisce uno spazio alla comunità internazionale dell'IA per avere discussioni approfondite riguardo alle opportunità e alle sfide condivise in materia di politiche sull'IA.

Nel 2022, in riconoscimento della crescente importanza e portata del suo lavoro sull'IA, il DPC ha creato un Gruppo di lavoro dedicato all'argomento: il Gruppo di lavoro sulla governance dell'IA (AIGO).

Il partito sostiene l'attuazione della raccomandazione dell'OCSE sull'intelligenza artificiale e sviluppa una guida pratica per l'implementazione, fornisce un forum e un hub online per lo scambio di informazioni su politica e attività sull'IA attraverso l'Osservatorio sulle politiche sull'IA dell'OCSE e promuovere la collaborazione tra più parti interessate e dialogo interdisciplinare attraverso i gruppi di esperti della rete di esperti OECD.AI.

IL CONSIGLIO,

CON RIGUARDO all'articolo 5 b) della Convenzione sull'Organizzazione per la cooperazione economica e Sviluppo del 14 dicembre 1960;

TENENDO CONTO degli standard elaborati dall'OCSE nei settori della privacy, della sicurezza digitale, della tutela dei consumatori e della condotta aziendale responsabile;

CONSIDERANDO gli obiettivi di sviluppo sostenibile stabiliti nell'Agenda 2030 per lo sviluppo sostenibile Risoluzione adottata dall'Assemblea generale delle Nazioni Unite ([A/RES/70/1](#)) nonché il 1948 Dichiarazione universale dei diritti umani;

CONSIDERATO l'importante lavoro svolto sull'intelligenza artificiale (di seguito, "IA") in altri forum internazionali governativi e non governativi;

RICONOSCENDO che l'IA ha implicazioni pervasive, di vasta portata e globali che stanno trasformando le società, i settori economici e il mondo del lavoro, e che probabilmente lo farà sempre di più in futuro;

RICONOSCENDO che l'IA ha il potenziale per migliorare il benessere delle persone, per contribuire a un'attività economica globale positiva e sostenibile, per aumentare l'innovazione e la produttività e per aiutare a rispondere alle principali sfide globali;

RICONOSCENDO che, al contempo, tali trasformazioni possono avere effetti diversi all'interno e tra società ed economie, in particolare per quanto riguarda i cambiamenti economici, la concorrenza, le transizioni nel mercato del lavoro, le disuguaglianze e le implicazioni per la democrazia e i diritti umani, la privacy e la protezione dei dati, nonché la sicurezza digitale;

RICONOSCENDO che la fiducia è un fattore chiave per la trasformazione digitale; che, sebbene la natura delle future applicazioni dell'IA e le loro implicazioni possano essere difficili da prevedere, l'affidabilità dei sistemi di IA è un fattore chiave per la diffusione e l'adozione dell'IA; e che un dibattito pubblico ben informato che coinvolga l'intera società è necessario per cogliere il potenziale benefico della tecnologia, limitando al contempo i rischi ad essa associati;

SOTTOLINEANDO che alcuni quadri giuridici, regolamentari e politici nazionali e internazionali esistenti sono già rilevanti per l'IA, compresi quelli relativi ai diritti umani, alla protezione dei dati personali e dei consumatori, ai diritti di proprietà intellettuale, alla condotta aziendale responsabile e alla concorrenza, pur rilevando che potrebbe essere necessario valutare l'adeguatezza di alcuni quadri e sviluppare nuovi approcci;

RICONOSCENDO che, dato il rapido sviluppo e l'implementazione dell'IA, è necessario un contesto politico stabile che promuova un approccio incentrato sull'uomo per un'IA affidabile, che favorisca la ricerca, preservi gli incentivi economici all'innovazione e che si applichi a tutte le parti interessate in base al loro ruolo e al contesto;

CONSIDERANDO che cogliere le opportunità offerte e affrontare le sfide poste dalle applicazioni di IA, nonché responsabilizzare le parti interessate, è essenziale per promuovere l'adozione di un'IA affidabile nella società e per trasformare l'affidabilità dell'IA in un parametro competitivo nel mercato globale.

In merito alla proposta del Comitato per le politiche digitali:

- I. ACCETTA che, ai fini della presente Raccomandazione, i seguenti termini debbano essere intesi come segue:

- *Sistema di intelligenza artificiale*: un sistema di intelligenza artificiale è un sistema basato su macchine che, per obiettivi espliciti o impliciti, deduce, a partire dagli input ricevuti, come generare output quali previsioni, contenuti, raccomandazioni o decisioni in grado di influenzare ambienti fisici o virtuali. I diversi sistemi di intelligenza artificiale variano nei loro livelli di autonomia e adattabilità dopo l'implementazione.
- *Ciclo di vita di un sistema di IA*: Il ciclo di vita di un sistema di IA in genere comprende diverse fasi, tra cui: pianificazione e progettazione; raccolta ed elaborazione dei dati; creazione di modelli e/o adattamento di modelli esistenti a compiti specifici; test, valutazione, verifica e convalida; messa a disposizione/implementazione; gestione e monitoraggio; e dismissione/smantellamento. Queste fasi si svolgono spesso in modo iterativo e non sono necessariamente sequenziali. La decisione di dismettere un sistema di IA può essere presa in qualsiasi momento durante la fase di gestione e monitoraggio.
- *Attori dell'IA*: Gli attori dell'IA sono coloro che svolgono un ruolo attivo nel ciclo di vita del sistema di IA, comprese le organizzazioni e gli individui che implementano o gestiscono l'IA.
- *Conoscenza dell'IA*: La conoscenza dell'IA si riferisce alle competenze e alle risorse, quali dati, codice, algoritmi, modelli, ricerca, know-how, programmi di formazione, governance, processi e migliori pratiche, necessarie per comprendere e partecipare al ciclo di vita dei sistemi di IA, inclusa la gestione dei rischi.
- *Parti interessate*: Le parti interessate comprendono tutte le organizzazioni e gli individui coinvolti nei sistemi di intelligenza artificiale o da essi influenzati, direttamente o indirettamente. Gli attori dell'IA costituiscono un sottoinsieme delle parti interessate.

Sezione 1: Principi per una gestione responsabile dell'IA affidabile

- II. **RACCOMANDA** che i membri e i non membri che aderiscono alla presente Raccomandazione (di seguito "gli aderenti") promuovano e attuino i seguenti principi, che sono rilevanti per tutte le parti interessate, per una gestione responsabile dell'IA affidabile.
- III. **INVITA** tutti gli attori dell'IA a promuovere e attuare, secondo i rispettivi ruoli, i seguenti principi per una gestione responsabile dell'IA affidabile.
- IV. **SOTTOLINEA** che i seguenti principi sono complementari e dovrebbero essere considerati come un intero.

1.1 Crescita inclusiva, sviluppo sostenibile e benessere

Le parti interessate dovrebbero impegnarsi attivamente in una gestione responsabile dell'IA affidabile, al fine di perseguire risultati positivi per le persone e per il pianeta, come ad esempio potenziare le capacità umane e migliorare la creatività, promuovere l'inclusione delle popolazioni sottorappresentate, ridurre le disuguaglianze economiche, sociali, di genere e di altro tipo e proteggere gli ambienti naturali, favorendo così una crescita inclusiva, il benessere, lo sviluppo sostenibile e la sostenibilità ambientale.

1.2 Rispetto dello stato di diritto, dei diritti umani e dei valori democratici, compresi l'equità e la privacy.

- a) Gli attori che operano nel campo dell'intelligenza artificiale devono rispettare lo stato di diritto, i diritti umani, i valori democratici e centrati sulla persona durante l'intero ciclo di vita del sistema di IA. Ciò include la non discriminazione e l'uguaglianza, la libertà, la dignità, l'autonomia degli individui, la privacy e la protezione dei dati, la diversità, l'equità, la giustizia sociale e i diritti dei lavoratori riconosciuti a livello internazionale. Questo include anche la lotta alla disinformazione e alla disinformazione amplificate dall'IA, nel rispetto della libertà di espressione e degli altri diritti e libertà tutelati dal diritto internazionale applicabile.
- b) A tal fine, gli attori dell'IA dovrebbero implementare meccanismi e garanzie, come la capacità di intervento e supervisione umana, anche per affrontare i rischi derivanti da usi al di fuori dello scopo previsto, da abusi intenzionali o da abusi non intenzionali, in modo appropriato al contesto e coerente con lo stato dell'arte.

1.3. Trasparenza e spiegabilità

Gli attori che operano nel settore dell'IA dovrebbero impegnarsi a garantire trasparenza e divulgazione responsabile delle informazioni relative ai sistemi di IA. A tal fine, dovrebbero fornire informazioni significative, appropriate al contesto e coerenti con lo stato dell'arte:

- i. per promuovere una comprensione generale dei sistemi di intelligenza artificiale, comprese le loro capacità e limitazioni,
- ii. per rendere le parti interessate consapevoli delle loro interazioni con i sistemi di intelligenza artificiale, anche sul posto di lavoro,
- iii. laddove possibile e utile, fornire informazioni chiare e di facile comprensione sulle fonti dei dati/input, sui fattori, sui processi e/o sulla logica che hanno portato alla previsione, al contenuto, alla raccomandazione o alla decisione, per consentire a coloro che sono interessati da un sistema di IA di comprenderne l'output e,
- iv. fornire informazioni che permettano a coloro che sono stati influenzati negativamente da un sistema di intelligenza artificiale di contestarne i risultati.

1.4. Robustezza, sicurezza e protezione

- a) I sistemi di intelligenza artificiale devono essere robusti, sicuri e protetti durante tutto il loro ciclo di vita, in modo che, in condizioni di utilizzo normale, utilizzo o abuso prevedibili o altre condizioni avverse, funzionino correttamente e non presentino rischi irragionevoli per la sicurezza.
- b) Dovrebbero essere predisposti, ove opportuno, meccanismi per garantire che, qualora i sistemi di IA rischiano di causare danni eccessivi o mostrino comportamenti indesiderati, possano essere disattivati, riparati e/o dismessi in modo sicuro, se necessario.

1.5 Responsabilità

- a) Gli attori che operano nel campo dell'IA dovrebbero essere responsabili del corretto funzionamento dei sistemi di IA e del rispetto dei principi sopracitati, in base ai loro ruoli, al contesto e in linea con lo stato dell'arte.
- b) A tal fine, gli attori dell'IA dovrebbero garantire la tracciabilità, anche in relazione ai set di dati, ai processi e alle decisioni prese durante il ciclo di vita del sistema di IA, per consentire l'analisi degli output del sistema di IA e delle risposte alle interrogazioni, appropriata al contesto e coerente con lo stato dell'arte.
- c) Gli attori dell'IA, in base ai loro ruoli, al contesto e alla loro capacità di agire, dovrebbero applicare un approccio sistematico alla gestione del rischio in ogni fase del ciclo di vita del sistema di IA su base continuativa e adottare una condotta aziendale responsabile per affrontare i rischi relativi ai sistemi di IA, anche attraverso la cooperazione tra i diversi attori dell'IA, i fornitori di conoscenze e risorse di IA, gli utenti dei sistemi di IA e le altre parti interessate. I rischi includono quelli relativi a pregiudizi dannosi, diritti umani, tra cui sicurezza, protezione e privacy, nonché diritti del lavoro e di proprietà intellettuale.

Sezione 2: Politiche nazionali e cooperazione internazionale per un'IA affidabile

V. **RACCOMANDA** agli aderenti di attuare le seguenti raccomandazioni, in conformità con i principi di cui alla sezione 1, nelle proprie politiche nazionali e nella cooperazione internazionale, con particolare attenzione alle piccole e medie imprese (PMI).

2.1. Investire nella ricerca e nello sviluppo dell'IA

- a) I governi dovrebbero prendere in considerazione investimenti pubblici a lungo termine e incoraggiare investimenti privati nella ricerca e sviluppo e nella scienza aperta, compresi gli sforzi interdisciplinari, per stimolare l'innovazione nell'IA affidabile, concentrandosi su questioni tecniche complesse e sulle implicazioni sociali, legali ed etiche e sulle questioni politiche legate all'IA.
- b) I governi dovrebbero inoltre considerare gli investimenti pubblici e incoraggiare gli investimenti privati in strumenti open source e set di dati aperti che siano rappresentativi e rispettino la privacy e la protezione dei dati, al fine di sostenere un ambiente per la ricerca e lo sviluppo dell'IA privo di pregiudizi dannosi e per migliorare l'interoperabilità e l'uso degli standard.

2.2. Promuovere un ecosistema inclusivo che favorisca l'intelligenza Artificiale

I governi dovrebbero promuovere lo sviluppo e l'accesso a un ecosistema digitale inclusivo, dinamico, sostenibile e interoperabile per un'intelligenza artificiale affidabile. Tale ecosistema comprende, tra l'altro, dati, tecnologie di intelligenza artificiale, infrastrutture computazionali e di connettività e, ove opportuno, meccanismi per la condivisione delle conoscenze in materia di intelligenza artificiale. A tal proposito, i governi dovrebbero valutare la possibilità di promuovere meccanismi, come i trust di dati, a supporto della condivisione sicura, equa, legale ed etica dei dati.

2.3. Creare un contesto di governance e politiche interoperabili e favorevoli all'IA

- a) I governi dovrebbero promuovere un contesto politico agile che supporti la transizione dalla fase di ricerca e sviluppo alla fase di implementazione e gestione di sistemi di intelligenza artificiale affidabili. A tal fine, dovrebbero valutare l'utilizzo della sperimentazione per fornire un ambiente controllato in cui i sistemi di intelligenza artificiale possano essere testati e, se del caso, ampliati. Dovrebbero inoltre adottare approcci basati sui risultati che offrano flessibilità nel raggiungimento degli obiettivi di governance e cooperare all'interno e tra le giurisdizioni per promuovere, ove opportuno, contesti di governance e politici interoperabili.
- b) I governi dovrebbero rivedere e adattare, se del caso, i propri quadri normativi e politici e i meccanismi di valutazione relativi ai sistemi di IA, al fine di incoraggiare l'innovazione e la concorrenza per un'IA affidabile.

2.4. Rafforzare le capacità umane e prepararsi alla trasformazione del mercato del lavoro

- a) I governi dovrebbero collaborare strettamente con le parti interessate per prepararsi alla trasformazione del mondo del lavoro e della società. Dovrebbero mettere le persone in condizione di utilizzare e interagire efficacemente con i sistemi di intelligenza artificiale in un'ampia gamma di applicazioni, anche fornendo loro le competenze necessarie.
- b) I governi dovrebbero adottare misure, anche attraverso il dialogo sociale, per garantire una transizione equa per i lavoratori durante l'implementazione dell'IA, ad esempio attraverso programmi di formazione lungo tutto l'arco della vita lavorativa, sostegno per coloro che sono colpiti dalla perdita del posto di lavoro, anche attraverso la protezione sociale, e accesso a nuove opportunità nel mercato del lavoro.
- c) I governi dovrebbero inoltre collaborare strettamente con le parti interessate per promuovere l'uso responsabile dell'IA sul lavoro, per migliorare la sicurezza dei lavoratori, la qualità dei posti di lavoro e dei servizi pubblici, per favorire l'imprenditorialità e la produttività e per garantire che i benefici derivanti dall'IA siano condivisi in modo ampio ed equo.

2.5. Cooperazione internazionale per un'IA affidabile

- a) I governi, compresi i paesi in via di sviluppo e le parti interessate, dovrebbero cooperare attivamente per promuovere questi principi e per progredire nella gestione responsabile di un'IA affidabile.
- b) I governi dovrebbero collaborare in seno all'OCSE e ad altri forum globali e regionali per promuovere, ove opportuno, la condivisione delle conoscenze sull'IA. Dovrebbero incoraggiare iniziative internazionali, intersettoriali e aperte che coinvolgano più parti interessate, al fine di acquisire competenze a lungo termine sull'IA.
- c) I governi dovrebbero promuovere lo sviluppo di un approccio globale basato sul consenso e sulla partecipazione di più parti interessate. Standard tecnici per un'intelligenza artificiale interoperabile e affidabile.

- d) I governi dovrebbero inoltre incoraggiare lo sviluppo e l'utilizzo, da parte propria, di indicatori comparabili a livello internazionale per misurare la ricerca, lo sviluppo e l'implementazione dell'IA, e raccogliere le prove necessarie per valutare i progressi nell'attuazione di questi principi.

- VI. INVITA** il Segretario Generale e gli Aderenti a diffondere la presente Raccomandazione
- VII. INVITA** i non aderenti a prendere debitamente in considerazione la presente Raccomandazione e ad aderirvi.
- VIII. INCARICA** il Comitato per le politiche digitali, tramite il suo gruppo di lavoro sulla governance dell'IA, di:
- a) continuare il suo importante lavoro sull'intelligenza artificiale basandosi su questa raccomandazione e tenendo conto del lavoro svolto in altri forum internazionali, e sviluppare ulteriormente il quadro di misurazione per le politiche sull'IA basate su dati concreti;
 - b) sviluppare e perfezionare ulteriormente le linee guida pratiche sull'attuazione della presente Raccomandazione per far fronte agli sviluppi in continua evoluzione e alle nuove priorità politiche;
 - c) fornire un forum per lo scambio di informazioni sulle politiche e le attività relative all'IA, comprese le esperienze con l'attuazione della presente Raccomandazione, e per promuovere un dialogo multidisciplinare e tra le parti interessate al fine di promuovere la fiducia e l'adozione dell'IA;
 - d) presentare al Consiglio una relazione, in consultazione con gli altri comitati competenti, sull'attuazione, la diffusione e la continua rilevanza della presente raccomandazione entro e non oltre cinque anni dalla sua revisione e almeno ogni dieci anni successivamente.

N.B. Il testo è stato tradotto automaticamente, dalla versione inglese, con Google Traduttore e rivisto redazionalmente.

<https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>