

INIZIATIVE, NORME E POLITICHE SULL'INTELLIGENZA ARTIFICIALE IN EUROPA E ITALIA

Report 18/06/2024

A cura di Alessio De Marco



INDICE

1. Iniziative e Norme Europee sull'Intelligenza Artificiale.....	3
Premessa	3
L'AI Act.....	3
AI Pact: Iniziativa Volontaria per l'Implementazione Anticipata dell'AI Act.....	4
Convenzione del Consiglio d'Europa sull'Intelligenza Artificiale: Un Approccio Basato sui Diritti Umani	4
2. Iniziative e Politiche italiane sull'Intelligenza Artificiale.....	5
Posizione Italiana sul Regolamento sull'Intelligenza Artificiale nell'Unione Europea.....	5
Il Piano triennale per l'informatica nella Pubblica Amministrazione 2024-2026.....	6
La Strategia Italiana per l'Intelligenza Artificiale 2024-2026.....	7
La posizione del Garante Privacy nella Regolamentazione dell'Intelligenza Artificiale in Italia.....	8
Disegno di Legge Italiano sull'Intelligenza Artificiale: Principi e Obiettivi.....	9
3. Regolamentazione delle AI da Parte delle Amministrazioni Locali statunitensi	10
Politica di Seattle sull'AI Generativa.....	10
Linee Guida di Boston sull'AI Generativa	10
4. Intelligenza Artificiale nei Servizi Pubblici italiani: alcuni esempi	11
L'intelligenza artificiale nella gestione urbana	11
La Smart Control Room del Comune di Venezia.....	11
La Smart City Control Room del Comune di Firenze	11
Il monitoraggio del traffico urbano del Comune di Milano.....	12
Il Gemello Digitale del Comune di Bologna	12
L'AI nel welfare	12
RiskER: Stratificazione del Rischio nella Regione Emilia-Romagna	12
Sanità digitale per l'ASST di Vimercate.....	12
INPS.....	13
Approfondimento: Il Caso INPS nell'Implementazione dell'Intelligenza Artificiale	13
5. Progetti di Intelligenza artificiale in Italia tra sviluppo e ricerca	14
iGenius e il Modello Italia.....	14
Human Hall for Artificial Intelligence: un progetto interdisciplinare per lo sviluppo etico dell'AI	15



1. Iniziative e Norme Europee sull'Intelligenza Artificiale

Premessa

Questa sezione non vuole essere totalmente esaustiva rispetto alle iniziative portate avanti dall'Unione Europea e dai suoi organi rispetto alla regolamentazione sull'intelligenza artificiale (AI), ma si limita a commentarne molto sinteticamente gli sviluppi più recenti. Tuttavia, è importante notare quanto la UE abbia preso in carico la questione in modo diretto e specifico almeno dal 2018 con il *Libro Bianco sull'Intelligenza Artificiale al servizio del cittadino*, la *Dichiarazione di Cooperazione sull'Intelligenza Artificiale* e in seguito con il *Piano Coordinato sull'Intelligenza Artificiale*.

L'AI Act

Approvato dal Parlamento Europeo il 13 marzo 2024, rappresenta la prima norma trasversale, che dunque intende regolare diversi settori e ambiti, al mondo sull'AI, con l'obiettivo di creare un quadro normativo armonizzato e proporzionato per queste tecnologie nell'Unione Europea. La proposta di regolamento si basa sul principio che l'AI debba essere sviluppata e utilizzata in modo sicuro, etico e rispettoso dei diritti fondamentali e dei valori europei. Per quel che riguarda un primo metodo di classificazione, i sistemi di AI¹ vengono classificati in base al loro livello di rischio: inaccettabile, alto, limitato e minimo. Maggiore è il rischio, maggiori le responsabilità e i paletti per chi sviluppa o adopera sistemi di intelligenza artificiale. Così, i sistemi con rischio inaccettabile, come quelli che possono manipolare il comportamento umano o quelli che utilizzano sistemi di scoring sociale, vengono vietati nei confini della Comunità. I sistemi ad alto rischio, come quelli utilizzabili per la selezione del personale o la sorveglianza biometrica, devono soddisfare rigorosi obblighi di trasparenza, sicurezza e controllo umano.

Per quanto riguarda i sistemi per "uso generale" (parzialmente classificabili come a basso rischio, secondo la nomenclatura dell'AI Act) viene richiesto agli sviluppatori di marcare i contenuti generati da AI in modo leggibile dalle macchine e riconoscibili dagli utenti umani, inclusi i deepfake, che devono essere etichettati tramite watermarking. Gli utenti devono essere informati quando interagiscono con un chatbot guidato da intelligenza artificiale. In questa categoria il regolamento identifica inoltre sistemi ad alto impatto², come GPT-4 di OpenAI e Gemini di Alphabet (Google), definendo come gli sviluppatori di questi sistemi debbano conformarsi a regole di sicurezza informatica, trasparenza nei processi di addestramento e condivisione della documentazione tecnica prima della commercializzazione.

Il regolamento che disciplinano anche l'immissione sul mercato, la messa in servizio e l'uso dei sistemi di AI nell'UE, si applica principalmente ai fornitori di sistemi di intelligenza artificiale, siano essi pubblici o privati. Le regole riguardano anche i fornitori situati in paesi terzi, nel caso in cui l'output dei loro sistemi venga utilizzato nel territorio della Comunità. Questa disposizione è simile a quella del GDPR, che ha esteso l'applicazione globale delle normative sulla circolazione e protezione dei dati.

L'AI Act mira a creare un mercato unico per l'intelligenza artificiale, in modo da facilitare la libera circolazione e il riconoscimento dei sistemi conformi alle norme UE. Intende promuovere una maggior fiducia nei sistemi

¹ Nel documento l'intelligenza artificiale viene definita come "un sistema automatizzato progettato per funzionare con diversi livelli di autonomia e che può mostrare capacità di adattamento dopo l'installazione e che, per obiettivi espliciti o impliciti, deduce, dagli input che riceve, come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali".

² I sistemi di intelligenza artificiale ad alto impatto vengono indicati come quelli che superano una soglia di capacità computazionale pari a 10^{25} FLOPs (Floating Point Operations Per Second). Questa soglia, che riflette il potere di calcolo utilizzato durante l'addestramento dei modelli di AI e include attività e metodi come il pre-addestramento, la generazione di dati sintetici e il perfezionamento.



di AI, prevenendo e mitigando i rischi a questi associati. Vuole supportare l'innovazione e l'eccellenza nel campo dell'AI, promuovendo la cooperazione tra gli Stati membri e ponendo le basi per la promozione di incentivi e finanziamenti per lo sviluppo di sistemi di intelligenza artificiale che possano rifarsi, nei termini di sicurezza ed eticità, alla formulazione del documento.

L'AI Act, inoltre, istituisce un comitato europeo per l'Intelligenza Artificiale e prevede la creazione di centri di competenza, di test e di innovazione per supportare gli sviluppatori e gli utilizzatori di sistemi di AI.

AI Pact: Iniziativa Volontaria per l'Implementazione Anticipata dell'AI Act

Promosso dalla Commissione Europea, rappresenta un'iniziativa volontaria che incoraggia le organizzazioni a prepararsi anticipatamente per l'attuazione delle misure previste dall'AI Act. La proposta invita le aziende a iniziare a conformarsi ai requisiti dell'AI Act prima che le normative diventino obbligatorie, supportando così una transizione più fluida. L'AI Pact si articola su due pilastri: il primo mira a creare una rete collaborativa per condividere *best practices* e informazioni pratiche sull'implementazione dell'AI Act; il secondo incoraggia i fornitori e gli utilizzatori di sistemi di AI ad intraprendere azioni preventive per conformarsi alle nuove norme. Le organizzazioni partecipanti dovranno comunicare regolarmente i loro progressi, promuovendo così una maggior trasparenza e incentivando la fiducia nelle tecnologie sviluppate. Questo approccio proattivo, simile a quanto avvenuto con il GDPR, mira a evitare problematiche di conformità una volta che il regolamento sarà pienamente operativo.

Convenzione del Consiglio d'Europa sull'Intelligenza Artificiale: Un Approccio Basato sui Diritti Umani

A maggio il Consiglio d'Europa ha adottato la prima convenzione sull'intelligenza artificiale, volto a garantire il rispetto delle norme giuridiche in materia di diritti umani, democrazia e Stato di diritto nell'utilizzo dei sistemi di AI.

Questo trattato, sviluppato in due anni dal Comitato sull'Intelligenza Artificiale (Committee on Artificial Intelligence - CAI), coinvolge 46 Stati membri, l'Unione Europea e 11 Stati non membri, oltre a rappresentanti del settore privato, della società civile e dell'accademia. La convenzione mira a garantire che l'uso dell'AI rispetti i diritti umani, la democrazia e lo Stato di diritto. Essa disciplina l'uso dell'AI nei settori pubblico e privato, imponendo obblighi di trasparenza, supervisione, responsabilità e protezione della dignità umana e dell'autonomia individuale. Prevede misure per prevenire e mitigare i rischi associati all'utilizzo di sistemi di AI, comprese potenziali moratorie o divieti sull'uso di sistemi di intelligenza artificiale in contesti specifici. Il trattato sarà aperto alla firma a Vilnius, in Lituania, il 5 settembre 2024.

Il trattato adotta un approccio basato sul rischio, simile a quello dell'AI Act, che prevede la valutazione delle potenziali conseguenze negative dell'uso dell'AI per sfruttarne i vantaggi in modo sicuro e responsabile. L'approccio basato sul rischio del trattato prevede che ogni Stato adotti misure graduate e differenziate per gestire i rischi associati ai sistemi di AI, considerando il contesto, l'uso previsto, la gravità e la probabilità degli impatti potenziali. Le misure includono l'identificazione, la valutazione, la prevenzione e la mitigazione dei rischi, la documentazione dei rischi e dei relativi approcci di gestione, e la necessità di testare i sistemi di AI prima del loro utilizzo e quando vengono significativamente modificati. Tale approccio è iterativo e coinvolge diverse parti interessate, garantendo che le misure adottate siano appropriate per il contesto specifico e i rischi identificati.



La convenzione prevede per le parti due modalità per conformarsi ai suoi principi e obblighi nel quadro della regolamentazione del settore privato: le parti possono scegliere di essere sottoposte direttamente alle disposizioni applicabili della convenzione o, in alternativa, prendere altre misure per conformarsi alle disposizioni del trattato rispettando appieno i loro obblighi internazionali in materia di diritti umani, democrazia e Stato di diritto. Dunque le modalità di assunzione della convenzione possono passare da:

1. **Legislative Measures:** adeguare le leggi nazionali per riflettere i principi e gli obblighi della convenzione. Ciò include l'adozione di leggi che regolamentano l'uso dell'AI in modo conforme ai diritti umani, alla democrazia e allo Stato di diritto.
2. **Administrative Measures:** implementare regolamenti e politiche amministrative che garantiscano che le attività relative all'AI siano conformi ai requisiti della convenzione. Questo può includere la creazione di agenzie di supervisione e monitoraggio specifiche per l'AI.
3. **Other Measures:** gli Stati possono anche adottare altre misure appropriate, come programmi di formazione e sensibilizzazione, per promuovere la conformità e la comprensione delle norme stabilite dalla convenzione.

Inoltre gli Stati firmatari devono istituire meccanismi di vigilanza indipendenti per monitorare la conformità alle disposizioni del trattato e devono fornire rapporti periodici alla Conferenza, che supervisiona l'implementazione della convenzione; devono coinvolgere il pubblico e le parti interessate in consultazioni e dibattiti per garantire che le politiche e le misure adottate riflettano una comprensione ampia e inclusiva delle implicazioni dell'AI. Gli Stati sono chiamati a promuovere l'alfabetizzazione digitale e le competenze necessarie per affrontare i rischi legati all'AI.

2. Iniziative e Politiche italiane sull'Intelligenza Artificiale

Posizione Italiana sul Regolamento sull'Intelligenza Artificiale nell'Unione Europea

All'interno delle negoziazioni che hanno portato alla definizione dell'AI Act, la posizione italiana si è distinta per il suo sostegno a un quadro normativo comune che tuteli i diritti fondamentali, stabilisca sanzioni proporzionate ai rischi e favorisca un avanzamento tecnologico competitivo rispetto ad altri attori globali come Stati Uniti e Cina.

L'Italia ha promosso nel corso delle trattative una visione "umano-centrica" delle norme di indirizzo, mirando alla semplificazione e chiarezza delle definizioni relative all'AI, ai sistemi generativi e ai modelli fondativi. Ha proposto la compartimentazione di settori esclusi dal regolamento originario chiedendo definizioni chiare nella coordinazione con le normative di settore, specialmente bancarie e assicurative.

Ha sostenuto un sistema di autovalutazione per le aziende riguardo ai sistemi di AI, auspicando linee guida e un archivio di esempi. Per quanto riguarda gli obblighi, l'Italia supporta la definizione di oneri proporzionati lungo la catena del valore dei sistemi di AI, evitando oneri eccessivi soprattutto per le PMI.

Nel contesto delle negoziazioni, l'Italia ha collaborato strettamente con Francia e Germania per influenzare il regolamento, cercando di garantire un approccio meno restrittivo verso i modelli di AI generativa e promuovendo norme basate sul rischio che non ostacolino l'innovazione. Si è inoltre impegnata a promuovere una politica industriale europea competitiva attraverso la cooperazione internazionale e il sostegno a start-up



e PMI nel settore dell'AI, con un focus sulla semplificazione amministrativa e la protezione dei dati sensibili nell'economia europea dei dati.

Il Piano triennale per l'informatica nella Pubblica Amministrazione 2024-2026

Il Piano Triennale per l'Informatica nella Pubblica Amministrazione 2024-2026, redatto AgID (l'Agenzia per l'Italia Digitale), pubblicato il 12 febbraio 2024, rappresenta un punto significativo per l'implementazione delle tecnologie di intelligenza artificiale nella pubblica amministrazione italiana. Il piano è frutto di una concertazione tra varie amministrazioni e soggetti istituzionali.

Il documento adotta la definizione di sistema di intelligenza artificiale fornita dall'OCSE, utilizzata anche nel Regolamento europeo per l'intelligenza artificiale (il precedentemente trattato AI Act).

Il piano individua tre aree prioritarie di intervento per l'AI nella pubblica amministrazione: l'automatizzazione di compiti ripetitivi, il miglioramento delle capacità predittive per decisioni basate sui dati e il supporto alla personalizzazione dei servizi pubblici. L'automatizzazione vuole essere mirata a liberare risorse umane da attività routinarie, permettendo loro di concentrarsi su compiti di maggiore valore aggiunto. Il miglioramento delle capacità predittive dovrebbe tradursi in decisioni più informate e tempestive, basate su analisi dati approfondite. Infine, la personalizzazione dei servizi pubblici punta a offrire soluzioni su misura per i cittadini, migliorando la qualità e l'efficacia dell'erogazione dei servizi.

Per supportare l'adozione dell'AI, il piano prevede la redazione di tre linee guida specifiche entro il 2024: una per promuovere l'adozione dell'intelligenza artificiale nella pubblica amministrazione, una per il *procurement* e una per lo sviluppo di applicazioni di AI. Le linee guida per l'adozione dell'AI forniranno strumenti di *assessment* per consentire alle pubbliche amministrazioni di identificare sia le esigenze che le motivano a utilizzare sistemi di AI, sia lo stato delle infrastrutture digitali a loro disposizione. L'obiettivo è avviare almeno 150 progetti di innovazione mediante AI entro il 2025 e 400 entro il 2026. Le linee guida per il *procurement* orienteranno le PA nell'approvvigionamento di tecnologie AI, con l'obiettivo di avviare almeno 100 iniziative di acquisizione di servizi di intelligenza artificiale nel 2025 e 300 nel 2026. Le linee guida per lo sviluppo di applicazioni AI mirano a fornire alle PA gli strumenti necessari per affrontare progetti di sviluppo di soluzioni ad hoc conformi alla normativa vigente, con almeno 50 progetti di sviluppo di soluzioni AI nel 2025 e 100 nel 2026.

Un altro aspetto cruciale del piano è la valorizzazione del patrimonio informativo pubblico, considerato un obiettivo strategico essenziale. AgID dovrà completare entro il 2024 una ricognizione delle basi di dati strategiche per identificare il corpus di dati necessario per addestrare i Large Language Models correlati all'AI generativa e ottenere scelte pertinenti e accuratamente allineate con le specificità dei vari domini e settori interni alla PA. Entro il 2025, l'agenzia dovrà digitalizzare queste basi di dati strategiche e, a partire dal 2026, promuoverle per l'addestramento dei sistemi di AI.

Parallelamente, AgID prevede di realizzare applicazioni di intelligenza artificiale a valenza nazionale attraverso un processo che inizia con l'identificazione delle soluzioni già esistenti a livello nazionale basate sull'AI. Il passo successivo sarà lo sviluppo ulteriore nel 2025 di queste soluzioni, al fine di trasformarle da prototipi a servizi estendibili su scala nazionale, per poi procedere nel 2026 con il dispiegamento nei territori delle soluzioni realizzate. Il piano sottolinea inoltre l'importanza della standardizzazione, della sostenibilità e dell'adozione di modelli di AI ad alto impatto, come i *foundation models*, che richiedono dataset di elevata qualità rappresentativi della realtà amministrativa.

Il piano triennale include anche un decalogo sui principi generali per l'utilizzo dell'AI nella pubblica amministrazione, elaborato con il contributo di vari stakeholders tra cui il mondo della ricerca, dell'università e dell'impresa, e pubbliche amministrazioni come INPS, Consip, INAIL e i comuni italiani. I principi guida comprendono il miglioramento dei servizi e la riduzione dei costi, l'analisi del rischio associato all'uso di AI, la trasparenza e la responsabilità nelle decisioni adottate, l'inclusività e l'accessibilità, la privacy e la sicurezza, la



formazione e lo sviluppo delle competenze, la standardizzazione, la sostenibilità, la governance dei *foundation models* e la gestione dei dati.

La Strategia Italiana per l'Intelligenza Artificiale 2024-2026

La Strategia Italiana per l'Intelligenza Artificiale 2024-2026 delinea un piano dettagliato per integrare l'AI in vari settori chiave del Paese. La strategia è primariamente articolata su quattro direttrici: Pubblica Amministrazione, Imprese, Formazione e Ricerca Scientifica.

Per la Pubblica Amministrazione, il documento sottolinea l'importanza di adottare tecnologie di AI per migliorare l'efficienza operativa e la qualità dei servizi. Prevede la formulazione di linee guida specifiche per promuovere l'adozione di sistemi di intelligenza artificiale, orientare il *procurement* e sviluppare applicazioni AI. Le linee guida sono fondamentali per garantire la trasparenza, la sicurezza e l'interoperabilità dei sistemi implementati. Inoltre, la strategia ha come obiettivo la possibilità di indirizzare linee di finanziamento verso progetti pilota su scala nazionale in grado di testare e perfezionare queste tecnologie, assicurando che le soluzioni sviluppate siano scalabili ai vari livelli e replicabili su tutto il territorio nazionale.

Le imprese sono incoraggiate a integrare l'AI nei loro processi produttivi e gestionali per migliorare la competitività e stimolare l'innovazione. La strategia promuove partenariati pubblico-privati per facilitare la collaborazione tra imprese, università e centri di ricerca. Viene posto l'accento sulla creazione di un ecosistema favorevole per le start-up e le PMI, con l'obiettivo di attrarre investimenti e sviluppare nuove soluzioni tecnologiche. In questo contesto, la strategia prevede misure di supporto per le imprese ICT italiane, promuovendone il ruolo abilitante nello sviluppo di applicazioni di AI e migliorando l'accesso ai finanziamenti.

In ambito formativo, la strategia prevede l'introduzione di percorsi educativi che abbiano al centro l'intelligenza artificiale a tutti i livelli, dalle scuole superiori alle università, per preparare le nuove generazioni all'uso consapevole di queste tecnologie. Viene promossa la formazione specialistica attraverso dottorati e programmi di *reskilling* e *upskilling* per aggiornare le competenze dei lavoratori, programmi essenziali per allineare le competenze dei lavoratori italiani alle nuove esigenze dettate dall'evoluzione tecnologica. La strategia sottolinea inoltre l'importanza di promuovere la parità di genere nelle discipline STEM e di sensibilizzare la società tanto sui benefici, quanto sui rischi associati all'utilizzo delle tecnologie dell'intelligenza artificiale.

La ricerca scientifica è un pilastro fondamentale della strategia, prevedendo investimenti significativi nella ricerca fondazionale³ e applicata. La strategia mira a sviluppare nuovi algoritmi e tecnologie, promuovendo la collaborazione internazionale per mantenere l'Italia competitiva a livello globale. Particolare attenzione è rivolta alla sostenibilità, alla privacy e alle implicazioni etiche dell'AI. Viene enfatizzata la necessità di consolidare l'ecosistema italiano della ricerca, favorendo la mobilità e l'attrazione di talenti, nonché la creazione di infrastrutture adeguate per supportare lo sviluppo tecnologico.

³ La ricerca fondazionale si concentra sullo sviluppo di nuovi algoritmi, modelli, sistemi di spiegabilità e robustezza, paradigmi di privacy e riservatezza dei dati e architetture di calcolo. È orientata a sviluppare le basi teoriche e metodologiche che consentono l'evoluzione delle applicazioni pratiche di AI. La strategia italiana enfatizza la necessità di un approccio fondazionale che sia in linea con i principi di affidabilità e responsabilità propri dei paradigmi europei. Punta a promuovere la collaborazione tra diverse competenze e settori disciplinari, volendo evitare di focalizzarsi su singole tecnologie specifiche e aprendo il campo all'esplorazione di nuove strade il cui potenziale potrebbe emergere nel medio periodo.



Parallelamente, la strategia prevede la realizzazione di applicazioni di intelligenza artificiale a valenza nazionale. Questo processo inizia con l'identificazione delle soluzioni già esistenti e prosegue con il loro sviluppo e dispiegamento su larga scala. La creazione di una Fondazione per l'AI è un elemento chiave, con il compito di coordinare le azioni strategiche, gestire i fondi dedicati e monitorare l'implementazione della strategia per garantire un miglioramento continuo.

La posizione del Garante Privacy nella Regolamentazione dell'Intelligenza Artificiale in Italia

La recente approvazione dell'AI Act da parte del Parlamento europeo ha sollevato questioni tra il governo italiano e il Garante per la Protezione dei Dati Personali per la protezione dei dati personali. Il dibattito riguarda chi debba avere l'autorità di vigilanza sull'intelligenza artificiale. Il governo affida tali funzioni all'AgiD e all'Agenzia per la cybersicurezza nazionale (ACN), mentre il Garante, sostenendo l'importanza di un organismo indipendente, rende noto quanto sarebbe auspicabile che il controllo rimanga presso le Autorità di protezione dei dati personali, in quanto dotate delle competenze necessarie per gestire l'impatto dell'AI sui diritti fondamentali e garantire un livello elevato di tutela dei cittadini.

Il Garante ha delineato una serie di indicazioni dettagliate riguardanti l'uso dell'intelligenza artificiale, nel solco di un documento indirizzato alla "regolamentazione" dei servizi sanitari nazionali. Queste linee guida mirano a garantire che l'implementazione dell'AI sia conforme ai diritti fondamentali e alle libertà degli interessati e che i trattamenti dei dati siano condotti in modo trasparente, etico e sicuro.

Una delle prime indicazioni riguarda la base giuridica per il trattamento dei dati. Il trattamento dei dati sanitari, che rientrano nelle categorie speciali di dati personali, deve essere giustificato da un quadro normativo specifico che rispetti il diritto dell'Unione o degli Stati membri. Questo quadro normativo deve prevedere misure adeguate per proteggere i diritti e gli interessi degli individui coinvolti. La decisione automatizzata e la profilazione basate sull'intelligenza artificiale devono essere consentite solo se esplicitamente previste dalla legge e se sono in atto misure adeguate di tutela.

Il Garante enfatizza l'importanza dei principi di *accountability* e di *privacy by design* e *by default*. Le organizzazioni devono dimostrare di aver adottato misure tecniche e organizzative adeguate per proteggere i dati personali fin dalla fase di progettazione dei sistemi di AI. Queste misure includono la valutazione delle scelte tecnologiche e organizzative, garantendo che ogni trattamento di dati sia necessario e proporzionato rispetto agli obiettivi perseguiti. I principi di conoscibilità, non esclusività e non discriminazione algoritmica sono fondamentali per l'uso etico degli algoritmi di intelligenza artificiale. Gli utenti devono essere informati sull'esistenza di processi decisionali automatizzati e avere accesso a informazioni significative sulla logica utilizzata. Inoltre, deve esserci sempre un intervento umano in grado di supervisionare e, se necessario, correggere le decisioni automatizzate, garantendo che i sistemi di AI non creino discriminazioni.

La valutazione d'impatto sulla protezione dei dati (DPAI) è un requisito cruciale per tutti i trattamenti di dati che comportano un rischio elevato per i diritti e le libertà degli individui, inclusi quelli effettuati tramite intelligenza artificiale. Una DPAI deve essere condotta prima dell'implementazione di tali sistemi per identificare e mitigare i potenziali rischi. La qualità dei dati è un altro aspetto critico. I dati utilizzati nei sistemi di AI devono essere accurati e aggiornati per evitare risultati errati o discriminatori. La garanzia dell'integrità e della riservatezza dei dati è essenziale, con misure adeguate per proteggere i dati da accessi non autorizzati, perdite o danneggiamenti.



La trasparenza nei processi decisionali automatizzati è fondamentale per mantenere la fiducia del pubblico. Le organizzazioni devono rendere pubbliche le valutazioni d'impatto e le informazioni chiave sui sistemi di AI utilizzati. Inoltre, la supervisione umana deve essere mantenuta in tutte le fasi, soprattutto durante l'addestramento degli algoritmi, per garantire che le decisioni basate sulle tecnologie di intelligenza artificiale siano accurate e non discriminatorie. L'uso dell'AI deve essere accompagnato da un'attenzione costante ai profili etici del trattamento dei dati personali. La validazione degli algoritmi deve assicurare che l'introduzione di sistemi di AI migliori la qualità delle prestazioni senza ripercussioni negative in termini sociali, deontologici o etici.

Disegno di Legge Italiano sull'Intelligenza Artificiale: Principi e Obiettivi

Il recente disegno di legge italiano sull'intelligenza artificiale (Disposizioni e delega al Governo in materia di intelligenza artificiale), discusso e approvato dal Consiglio dei Ministri, mira a regolamentare l'uso dell'AI nel rispetto dei diritti fondamentali e delle libertà costituzionali. Questo disegno di legge si articola su vari fronti, dalla ricerca e sviluppo all'applicazione pratica dell'AI in ambiti come la sanità, il lavoro, la pubblica amministrazione e la giustizia.

Il documento sottolinea l'importanza di utilizzare l'AI in modo corretto, trasparente e responsabile, promuovendo un approccio antropocentrico che garantisca la tutela dei diritti umani e la prevenzione dei danni. La legge prevede che l'adozione dei sistemi di intelligenza artificiale avvenga nel rispetto della trasparenza, della sicurezza, della protezione dei dati personali e della non discriminazione, con un'attenzione particolare alla parità di genere e alla sostenibilità.

In ambito sanitario, il disegno di legge prevede l'uso dell'AI per migliorare la prevenzione, la diagnosi e la cura delle malattie, garantendo al contempo che l'accesso alle prestazioni sanitarie non sia discriminatorio. Gli utenti devono essere informati sull'uso delle tecnologie di intelligenza artificiale e sui benefici diagnostici e terapeutici derivanti dal loro utilizzo. Inoltre, la legge promuove lo sviluppo di sistemi di AI in grado di migliorare le condizioni di vita delle persone con disabilità, agevolandone l'accessibilità, l'autonomia e l'inclusione sociale.

Per quanto riguarda il lavoro, l'AI deve essere utilizzata per migliorare le condizioni lavorative, tutelare l'integrità psico-fisica dei lavoratori e aumentare la qualità delle prestazioni lavorative. L'utilizzo delle tecnologie di intelligenza artificiale deve rispettare la dignità umana e la riservatezza dei dati personali, garantendo l'assenza di discriminazioni basate su sesso, età, origine etnica, credo religioso, orientamento sessuale, opinioni politiche e condizioni personali, sociali ed economiche.

Nel settore della pubblica amministrazione, il disegno di legge promuove l'adozione di tecnologie AI per incrementare l'efficienza delle attività, ridurre i tempi dei procedimenti e migliorare la qualità dei servizi erogati ai cittadini. Tuttavia, la responsabilità finale delle decisioni deve rimanere in capo alle persone, garantendo che l'AI sia utilizzata come strumento di supporto e non come sostituto del giudizio umano. Il DDL stabilisce che ogni decisione automatizzata debba essere trasparente e comprensibile, con meccanismi di controllo che assicurino la conformità ai principi etici e legali. Prevede inoltre lo sviluppo di sistemi di AI che possano assistere nella gestione dei processi amministrativi, migliorando l'accessibilità e l'efficacia dei servizi offerti.

Il disegno di legge prevede anche la creazione di un Osservatorio sull'adozione dei sistemi di intelligenza artificiale nel mondo del lavoro, con il compito di "definire una strategia sull'utilizzo dell'intelligenza artificiale



in ambito lavorativo, monitorare l'impatto sul mercato del lavoro, identificare i settori lavorativi maggiormente interessati dall'avvento dell'intelligenza."

La legge introduce anche disposizioni specifiche per la protezione del diritto d'autore delle opere create con l'ausilio di sistemi di intelligenza artificiale, richiedendo che il contributo umano sia rilevante e dimostrabile, in modo da riconosce la creatività e l'originalità delle persone.

3. Regolamentazione delle AI da Parte delle Amministrazioni Locali statunitensi

Le esperienze delle città di Seattle e Boston offrono esempi significativi di come le amministrazioni locali possano sviluppare politiche per l'AI generativa.

Politica di Seattle sull'AI Generativa

La città di Seattle ha introdotto una politica completa per l'utilizzo dell'AI generativa, con l'obiettivo di garantire che queste tecnologie siano utilizzate in modo sicuro e trasparente. La politica di Seattle stabilisce una serie di principi guida, tra cui innovazione e sostenibilità, trasparenza e responsabilità, validità e affidabilità, riduzione dei *bias* e dei danni, protezione della privacy, spiegabilità e interpretabilità, sicurezza e resilienza. Questi principi sono progettati per garantire che l'IA sia utilizzata per migliorare i servizi pubblici, rispettando al contempo i diritti e la privacy dei cittadini.

L'acquisizione e l'uso di tecnologie di IA generativa a Seattle sono regolamentati attraverso processi rigorosi. I dipartimenti della città devono ottenere l'autorizzazione per utilizzare strumenti di IA generativa, e qualsiasi output generato deve essere revisionato da umani prima di essere utilizzato ufficialmente. Questo processo, noto come "*Human in the Loop*", garantisce che le decisioni finali siano sempre verificate da un operatore umano, riducendo il rischio di errori o *bias* indesiderati.

La politica di Seattle prevede anche misure specifiche per garantire la trasparenza e l'attribuzione dei contenuti generati dall'AI. Ogni immagine, video o testo prodotto da un sistema di intelligenza artificiale deve essere chiaramente attribuito al sistema che l'ha generato e i processi di revisione devono essere documentati accuratamente. Questa trasparenza aiuta a costruire fiducia nei sistemi di IA e a garantire che i cittadini siano informati sull'origine dei contenuti.

Per affrontare i potenziali i e danni, la politica di Seattle richiede l'uso di strumenti come il *Racial Equity Toolkit*⁴ per valutare gli impatti delle tecnologie di AI attraverso una lente di equità. Questo approccio proattivo mira a identificare e mitigare i *bias* prima che possano causare danni significativi.

Linee Guida di Boston sull'AI Generativa

Analogamente a Seattle la città di Boston ha sviluppato linee guida per l'uso dell'AI generativa, mirate a promuovere un'adozione responsabile e sperimentale di queste tecnologie. Le linee guida di Boston sono incentrate su principi di empowerment, inclusione e rispetto, trasparenza e responsabilità, innovazione e gestione del rischio, privacy e sicurezza.

Boston incoraggia l'uso dei sistemi di intelligenza artificiale per supportare il lavoro della forza lavoro pubblica e migliorare i servizi per i residenti, sottolineando l'importanza di agire con trasparenza per costruire fiducia e favorire l'apprendimento collettivo. La città promuove una cultura dell'esperimentazione responsabile,

⁴ Uno strumento sviluppato dalla città di Seattle per valutare le politiche, le iniziative, i programmi e le questioni di bilancio in termini di impatto sull'equità razziale. Il toolkit guida le organizzazioni attraverso un processo strutturato per identificare e mitigare le disuguaglianze razziali, coinvolgendo stakeholder, analizzando dati, determinando i benefici e gli oneri, avanzando opportunità o minimizzando i danni, e valutando i risultati per garantire che le decisioni prese non perpetuino le disparità esistenti, ma promuovano l'equità e la giustizia sociale



mantenendo un controllo rigoroso sull'uso delle nuove tecnologie e assicurando che l'AI sia utilizzata per scopi pubblici e benefici.

Un elemento chiave delle linee guida di Boston è la responsabilità nella verifica dei contenuti generati dall'AI. I dipendenti pubblici sono responsabili di verificare l'accuratezza e l'affidabilità delle informazioni prodotte dall'IA, specialmente quando queste vengono utilizzate per comunicazioni pubbliche o decisioni. Questa verifica è fondamentale per prevenire la diffusione di informazioni errate o soggette a *bias*.

Le linee guida di Boston sottolineano anche l'importanza di non condividere informazioni sensibili o private nei *prompt* utilizzati per generare contenuti con le applicazioni di intelligenza artificiale. Proteggere la privacy dei dati è essenziale per mantenere la fiducia dei cittadini e garantire che le tecnologie di AI siano utilizzate in modo sicuro e responsabile.

4. Intelligenza Artificiale nei Servizi Pubblici italiani: alcuni esempi

Questa sezione presenta alcuni esempi significativi di come l'intelligenza artificiale venga utilizzata nei servizi pubblici e nel benessere sociale in Italia. Non intende essere esaustiva, ma vuole offrire uno spaccato di alcuni ambiti applicativi in cui l'AI è già presente e delle questioni che il suo utilizzo può sollevare.

L'intelligenza artificiale nella gestione urbana

La Smart Control Room del Comune di Venezia

Descrizione: centrale operativa situata presso l'Isola Nuova del Tronchetto. Inaugurata nel settembre 2020, è un progetto realizzato in collaborazione tra il Comune di Venezia, la Polizia Locale, TIM e Venis Spa. La Smart Control Room integra diverse tecnologie avanzate, tra cui intelligenza artificiale, *internet delle cose*, cloud computing e connettività 5G, per monitorare e gestire in tempo reale la mobilità, la sicurezza e i flussi turistici della città. La piattaforma tecnologica utilizzata, Mindcity, consente di raccogliere ed elaborare dati provenienti da sensori distribuiti sul territorio, videocamere intelligenti e la rete telefonica TIM, fornendo agli operatori informazioni in tempo reale e previsioni utili per la gestione urbana.

Link: [La stanza dove si vede tutto quello che succede a Venezia](#)

[Venezia, contributo di accesso e Smart Control Room - Smart Controlled - Ep1](#) (video)

La Smart City Control Room del Comune di Firenze

Descrizione: Progettata da Snap4City, è un hub tecnologico avanzato che sfrutta tecnologie come l'Internet of Things (IoT), l'analisi dei dati e l'intelligenza artificiale per monitorare e gestire in tempo reale diversi aspetti della vita urbana, tra cui traffico, sicurezza, illuminazione e servizi pubblici. Snap4City, sviluppato dal DISIT Lab dell'Università di Firenze, è una piattaforma open-source basata su FIWARE (una piattaforma open-source che fornisce API e strumenti per lo sviluppo di applicazioni smart) che supporta l'integrazione e l'elaborazione di dati provenienti da molteplici fonti, consentendo l'implementazione rapida di sistemi di supporto decisionale e applicazioni smart. Utilizza tecnologie di intelligenza artificiale avanzate, come il machine learning, l'analisi predittiva, il riconoscimento di anomalie, natural language processing e l'Explainable artificial intelligence, per gestire e ottimizzare in tempo reale vari aspetti della vita urbana. Queste tecnologie vengono applicate a situazioni come il monitoraggio del traffico per ridurre la congestione, la gestione energetica per migliorare l'efficienza dei consumi, la sicurezza pubblica per rilevare comportamenti anomali e prevenire incidenti e la risposta alle emergenze per coordinare interventi rapidi e mirati. Inoltre, strumenti di visualizzazione dei dati aiutano gli operatori a interpretare rapidamente le informazioni critiche. Il progetto è stato co-finanziato dal PON Città Metropolitane e vede la collaborazione di vari partner tecnologici, tra cui Alia (gestione dei rifiuti), Silfi (illuminazione pubblica), Sas (analisi avanzata dei dati) e Global Service (manutenzione e gestione dei servizi). Snap4City ha già dimostrato la sua efficacia in diverse città europee, tra cui Firenze, Pisa, Livorno e Modena, offrendo soluzioni smart per migliorare la qualità della vita urbana.



Link: [Scenario: Firenze Smart City Control Room](#)
[Scheda progetto](#)

Il monitoraggio del traffico urbano del Comune di Milano

Descrizione: Telecamera dotata di intelligenza artificiale per monitorare il traffico urbano. Riconosce diverse tipologie di utenti della strada, inclusi pedoni e ciclisti, fornendo dati precisi per migliorare la pianificazione urbana e promuovere la mobilità sostenibile. Il progetto è promosso dall'Agenzia Mobilità Ambiente e Territorio (AMAT) e ha visto la collaborazione dell'ente di certificazione RINA, che ha rilasciato il certificato di conformità "Trustable AI Management System" per garantire la qualità e la sicurezza del sistema.

Link: [AMAT ha ottenuto la certificazione di un sistema avanzato di Intelligenza Artificiale per il monitoraggio della mobilità urbana milanese](#)

[Milano: la telecamera con intelligenza artificiale e i dubbi sulla privacy](#)

Il Gemello Digitale del Comune di Bologna

Descrizione: progetto volto a creare una rappresentazione digitale dinamica della città, progettato per supportare i processi decisionali attraverso l'uso di tecnologie avanzate. Il sistema utilizza machine learning, big data e internet delle cose per raccogliere e analizzare dati in tempo reale da varie fonti, come sensori urbani, dati satellitari e sistemi informativi esistenti. Tra i campi di applicazione: la mobilità urbana, l'energia, i cambiamenti climatici e la pianificazione urbana. Potrà essere utilizzato per monitorare il traffico e ottimizzare i percorsi dei veicoli, gestire il consumo energetico degli edifici pubblici, prevedere l'impatto dei cambiamenti climatici sulla città e supportare la pianificazione di nuovi quartieri. Avviato nel 2023, è finanziato dal Comune di Bologna (fondi europei PON Metro) con il supporto della Fondazione Innovazione Urbana. Tra i partner principali ci sono la Fondazione Bruno Kessler, l'Università di Bologna, il consorzio interuniversitario Cineca. Il progetto prevede una fase iniziale di sviluppo e implementazione seguita da un progressivo ampliamento delle funzionalità e dei campi di applicazione, con l'obiettivo di diventare uno strumento integrato per la gestione urbana e il coinvolgimento dei cittadini.

Link: [Presentazione gemello digitale](#)

[Bologna avrà un Gemello digitale](#)

L'AI nel welfare

RiskER: Stratificazione del Rischio nella Regione Emilia-Romagna

Descrizione: Progetto sviluppato dall'Agenzia Sanitaria e Sociale Regionale (ASSR) della Regione Emilia-Romagna, volto alla stratificazione del rischio della popolazione assistita utilizzando tecniche di intelligenza artificiale e modelli statistici avanzati. Il progetto impiega una regressione logistica multivariata (machine learning supervisionato) per analizzare dati amministrativi sanitari, come ricoveri ospedalieri, accessi al pronto soccorso e prescrizioni farmaceutiche, al fine di calcolare la probabilità di eventi avversi come l'ospedalizzazione evitabile o il decesso.

Link: [Guida alla stratificazione del rischio della popolazione residente in Emilia-Romagna con dati amministrativi: l'algoritmo riskER](#)

[Complessità assistenziale](#)

Sanità digitale per l'ASST di Vimercate

Descrizione: Sviluppato da Almaviva per l'ASST di Vimercate. Utilizza tecniche di machine learning e analisi predittiva e prescrittiva per gestire i processi sociosanitari. Monitorando i pazienti lungo il loro percorso clinico, il sistema analizza dati provenienti da esami ematici, esami fisici, comorbidità, ricoveri, accessi al pronto soccorso e farmacoterapia. In pratica, il sistema prevede l'insorgenza di complicanze, il momento della dialisi per i pazienti con insufficienza renale cronica, il re-ricovero per scompenso cardiaco, e la tossicità oncologica, supportando così le decisioni mediche del personale sanitario.

Link: [Asst Vimercate: l'intelligenza artificiale in corsia](#)



INPS

Descrizione: Sistema di classificazione delle email automatizza la gestione delle oltre 4 milioni di email ricevute annualmente, utilizzando modelli di Deep Learning come BERT per analizzare il contenuto e indirizzare le comunicazioni agli uffici competenti. Sviluppato in collaborazione con Accenture, mira a ridurre i tempi di risposta e migliorare il servizio ai cittadini. Sistema Informativo Integrato del Sistema Lavoro (SIISL), utilizza algoritmi di machine learning per l'analisi dei dati e la previsione delle esigenze del mercato del lavoro. Facilita l'incrocio tra domanda e offerta di lavoro, migliorando l'allocazione delle risorse umane. Inoltre, l'INPS utilizza modelli predittivi e analisi dei Big Data per monitorare i servizi erogati, identificando aree di miglioramento e ottimizzando le risorse, in collaborazione con enti di ricerca e università.

Link: [Intelligenza Artificiale per potenziare la Pubblica Amministrazione](#)

Approfondimento: Il Caso INPS nell'Implementazione dell'Intelligenza Artificiale

L'implementazione di sistemi di intelligenza artificiale all'interno dell'INPS rappresenta un caso significativo di utilizzo di queste tecnologie mirate all'innovazione e l'efficientamento dei servizi pubblici. Le "Linee guida sull'implementazione di sistemi di Intelligenza Artificiale in INPS", delineate in una Direttiva del Direttore Generale (N. 8 del 8 aprile 2024), mirano a massimizzare il valore per i cittadini attraverso l'uso responsabile e trasparente delle tecnologie AI. L'intelligenza artificiale è individuata come una tecnologia rivoluzionaria nel percorso di ottimizzazione, gestione ed erogazione dei servizi pubblici, in grado di migliorare l'efficienza operativa e la qualità dei servizi offerti.

Nel contesto del Piano Strategico Digitale ICT 2022-2024, l'INPS ha avviato un percorso di evoluzione che prevede l'adozione di sistemi di AI per sfruttare i dati raccolti dall'Istituto e ottimizzare i servizi offerti. Questa strategia può consentire di rispondere in modo più preciso e tempestivo alle necessità dei cittadini, migliorando l'esperienza degli utenti e l'efficacia dell'erogazione dei servizi. Tra le principali potenzialità dei sistemi di AI, nel piano vengono evidenziate la capacità di automatizzare attività ripetitive, la personalizzare dei servizi per i cittadini e il supporto al processo decisionale attraverso l'analisi di grandi quantità di dati, la possibilità di migliorare l'efficienza operativa e incrementare la sicurezza e garantire la *compliance* normativa.

L'INPS ha già concluso con successo numerose iniziative progettuali basate su tecnologie AI, tra cui l'adozione di chatbot per migliorare l'interazione con gli utenti, la gestione delle richieste al customer service e lo smistamento automatico della posta elettronica certificata (PEC). Questi progetti hanno permesso di risparmiare tempo e risorse, migliorando al contempo la qualità dei servizi offerti.

Le linee guida per l'implementazione dell'intelligenza artificiale nell'INPS includono la promozione di un centro di competenza per la diffusione delle conoscenze sull'AI, la prioritizzazione delle iniziative progettuali in base al loro impatto e valore aggiunto, il monitoraggio degli impatti positivi per l'utenza e la gestione dei rischi associati all'utilizzo di queste tecnologie in conformità con il regolamento europeo sull'Intelligenza Artificiale. Inoltre, l'INPS ha adottato il framework INPS AI@SCALE, un modello strutturato per garantire un'implementazione industrializzata e standardizzata delle soluzioni di AI. Questo framework si articola in tre aree principali: *AI Value Management*, *AI & Data Scope / Risk Definition* e *AI Governance*. L'*AI Value Management* stabilisce i casi d'uso, pianifica le priorità e monitora i progressi dei progetti di intelligenza artificiale; l'*AI & Data Scope / Risk Definition* definisce le classi di algoritmi, i dati necessari e le tecniche di anonimizzazione per garantire la conformità normativa e la sicurezza; l'*AI Governance* coordina la gestione complessiva, assicurando che i progetti siano allineati con gli obiettivi strategici dell'istituto.

L'INPS ha adottato diverse scelte tecnologiche per l'implementazione delle soluzioni AI, valutando i costi e i benefici di ciascuna opzione, dalle soluzioni SaaS alle infrastrutture on-premise. La strategia adottata mira a



sostenere lo sviluppo di sistemi AI autonomi italiani, promuovendo la crescita del sistema Paese e garantendo la sostenibilità a lungo termine delle soluzioni adottate. Le linee guida sottolineano l'importanza di una formazione continua per il personale, lo sviluppo di competenze specialistiche in *data science* e *AI engineering*, oltre alla necessaria integrazione di risorse umane già specializzate, e la creazione di partnership con università e aziende leader nel settore che possano offrire al personale dell'INPS opportunità di partecipazione a progetti di ricerca congiunti. Questo approccio mira a costruire un ecosistema di competenze e innovazioni che supporti la trasformazione digitale dell'INPS, rendendo l'istituto più efficiente, sicuro e in grado di rispondere efficacemente alle esigenze degli utenti.

5. Progetti di Intelligenza artificiale in Italia tra sviluppo e ricerca

Questa sezione illustra due progetti avanzati legati all'intelligenza artificiale in Italia. Non è una rassegna completa, ma offre uno sguardo su due "casi studio" esemplari delle connessioni che nel campo legano l'AI al settore pubblico e a quello privato.

iGenius e il Modello Italia

Il "Modello Italia" è un progetto di intelligenza artificiale sviluppato da iGenius⁵ in collaborazione con Cineca⁶, mirato a creare un Large Language Model italiano e open-source, progettato per rispondere alle specifiche esigenze linguistiche e culturali del contesto italiano. Il modello è stato addestrato con trilioni di token (attraverso un *tokenizer*⁷ proprietario di iGenius) per garantire una comprensione approfondita del linguaggio naturale italiano, riducendo i pregiudizi e assicurando la sicurezza dei dati in conformità con le normative GDPR.

Il "Modello Italia" vuole essere integrabile in vari settori strategici, tra cui la sanità, la finanza e la sicurezza nazionale, supportando le aziende e le pubbliche amministrazioni nell'innovazione e nell'efficienza operativa. Ospitato al Tecnopolo di Bologna, sede del supercomputer Leonardo, il progetto beneficia anche di significativi investimenti da parte di CDP Venture Capital, che ha stanziato un miliardo di euro per lo sviluppo di tecnologie avanzate come l'AI.

L'Università La Sapienza di Roma ha recentemente condotto un'analisi approfondita sul modello, mettendolo a confronto con altri modelli simili come Anita 8B, Maestrone 7B, Qwen2-7B, Italia 9B e Occiglot 7B (di questi, solo Qwen e Italia sono addestrati da zero sulla lingua italiana). L'analisi ha utilizzato sei benchmark standard per valutare le capacità dei modelli, con i risultati che indicano prestazioni inferiori del Modello Italia rispetto ai concorrenti. Nicolas Pantaleo di iGenius ha risposto ai risultati evidenziando quanto la conversione non ufficiale utilizzata per i benchmark possa aver influenzato negativamente le performance e ha promesso che una versione ufficiale con dettagli tecnici sarà rilasciata su Hugging Face⁸. L'analisi promossa dalla Sapienza

⁵ Startup tecnologica fondata da Uljan Sharka, con sede in Italia, specializzata nello sviluppo di soluzioni di intelligenza artificiale per l'analisi dei dati aziendali.

⁶ Consorzio interuniversitario senza scopo di lucro, fondato nel 1967, che comprende 70 università italiane, due ministeri (Ministero dell'Istruzione e del Merito, Ministero dell'Università e della Ricerca), e numerose istituzioni pubbliche nazionali.

⁷ Uno strumento che suddivide il testo in unità più piccole chiamate "token". Questi token possono essere parole, subparole o anche caratteri singoli. Nei modelli di linguaggio naturale, i tokenizer trasformano il testo in una sequenza numerica comprensibile per il modello, facilitando così l'elaborazione e l'analisi del linguaggio

⁸ Hugging Face è una piattaforma online e azienda tra le più importanti nell'intelligenza artificiale e nel machine learning, nota per la sua libreria open-source "Transformers" che offre modelli pre-addestrati per varie applicazioni di



mira a garantire trasparenza e riproducibilità, con l'intenzione di rilasciare a breve una suite di valutazione open-source. Questa iniziativa rientra in un contesto più ampio di ricerca sui LLM in Italia, con progetti come LLaMAntino dell'Università degli Studi di Bari e Cerbero dell'Università di Pisa.

Human Hall for Artificial Intelligence: un progetto interdisciplinare per lo sviluppo etico dell'AI

Il progetto "Human Hall for Artificial Intelligence" (HH4AI) dell'Università degli Studi di Milano (nel contesto dell'hub scientifico Human Hall) nasce dalla collaborazione di due gruppi di ricerca interdisciplinari formato da costituzionalisti e giuristi da un lato e informatici dall'altro. Questi studiano da anni l'impatto degli algoritmi sui diritti umani.

HH4AI si vuole concentrare sullo sviluppo di tecnologie AI che siano etiche e rispettose dei diritti umani, in risposta alla crescente pervasività delle decisioni algoritmiche nella vita delle persone. Il progetto mira a creare modelli affidabili e conformi ai valori etici, tramite la stretta collaborazione tra tecnologi e giuristi. Questo approccio interdisciplinare dovrebbe consentire di sviluppare un'AI migliore e di fornire alle aziende un vantaggio competitivo, aiutandole a rispettare le normative e a minimizzare i rischi sanzionatori e reputazionali.

Il progetto è attualmente unico nel suo genere per la stretta collaborazione tra le discipline che porta in campo, che vogliono dialogare e sviluppare tanto aspetti teorici quanto strumenti pratici disponibili.

elaborazione del linguaggio naturale. La piattaforma include anche il "Model Hub", un repository per condividere e scaricare modelli e una libreria di dataset per il *training* e il *testing* dei modelli.



Agenzia formativa
accreditata dalla Regione
Toscana n° OF0299
Fornitore
ACQUISTINRETEPA

ISCRITTA ALL'ANAGRAFE NAZIONALE DELLE RICERCHE DEL MIUR
ACCREDITATA SNA PER LA FORMAZIONE OIV
Viale G. Luporini 37/57 | 55100 Lucca
Tel. +39 0583 582783 | 583342 Fax +39 0583 1900211
info@promopa.it | promopa.it
N°80 Registro Persone Giuridiche Prefettura di Lucca | CF / PI 01922510464

