# Manifesto per un’Intelligenza Artificiale Etica e Human-Centric

L’intelligenza artificiale (IA) rappresenta una delle più grandi opportunità del nostro tempo, ma anche una delle più grandi sfide. Se sviluppata e utilizzata in modo responsabile, l’IA può migliorare la qualità della vita, promuovere l’innovazione e risolvere problemi complessi. Tuttavia, senza un quadro etico solido e un controllo umano adeguato, rischia di amplificare disuguaglianze, minacciare i diritti fondamentali e compromettere la sicurezza delle nostre società.

Questo manifesto nasce con l’obiettivo di unire autorità accademiche, politiche e istituzionali in un impegno comune: **garantire che l’IA sia sviluppata e implementata in modo etico, trasparente e human-centric**. Chiediamo all’Unione Europea di adottare misure concrete che vadano oltre il rispetto del GDPR, includendo controlli rigorosi sulla sicurezza etica e decisionale dei sistemi di IA.

The Manifesto for an Ethical and Human-Centric Artificial Intelligence

Copyright (C) Alessandro Verna and Alvise Gottieri May 9 2025.

Permission is granted to copy, distribute and/or modify this document

under the terms of the Gnu Free Documentation License, Version 1.3

https://www.gnu.org/licenses/fdl-1.3.html#license-text or any later

version published by the Free Software Foundation;

with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.

Sommario

[Manifesto per un’Intelligenza Artificiale Etica e Human-Centric 1](#_Toc193188601)

[1. Perché un Manifesto per un’IA Etica e Human-Centric? 5](#_Toc193188602)

[1.1 L’Impatto Trasformativo dell’IA 5](#_Toc193188603)

[1.1.1 Opportunità dell’IA 5](#_Toc193188604)

[1.1.2 Rischi dell’IA 5](#_Toc193188605)

[1.2 Le Sfide Etiche dell’IA 5](#_Toc193188606)

[1.2.1 Discriminazione e Bias 6](#_Toc193188607)

[1.2.2 Privacy e Sorveglianza 6](#_Toc193188608)

[1.2.3 Trasparenza e Responsabilità 6](#_Toc193188609)

[1.2.4 Impatto Occupazionale 6](#_Toc193188610)

[1.3 La Necessità di un Approccio Human-Centric 6](#_Toc193188611)

[1.3.1 Supervisione Umana 6](#_Toc193188612)

[1.3.2 Allineamento ai Valori Umani 6](#_Toc193188613)

[1.3.3 Benessere Sociale e Ambientale 6](#_Toc193188614)

[1.4 Il Nuovo Significato di Formazione delle Nuove Generazioni 7](#_Toc193188615)

[1.4.1 Educazione all’IA 7](#_Toc193188616)

[1.4.2 Adattamento Cognitivo 7](#_Toc193188617)

[1.4.3 Formazione Continua 7](#_Toc193188618)

[1.5 Il Pericolo dell’Utilizzo dell’IA per Scopi Militari 8](#_Toc193188619)

[1.5.1 Armi Autonome e Decisioni Letali 8](#_Toc193188620)

[1.5.2 Violazione del Diritto Internazionale 8](#_Toc193188621)

[1.5.3 Erosione del Controllo Umano 8](#_Toc193188622)

[1.5.4 Necessità di un Controllo Umano Finale 8](#_Toc193188623)

[1.5.5 Implicazioni per la Sicurezza Globale 9](#_Toc193188624)

[2. I Principi Fondamentali del Manifesto 10](#_Toc193188625)

[2.1 Supervisione Umana 10](#_Toc193188626)

[2.1.1 Livelli di Supervisione 10](#_Toc193188627)

[2.1.2 Implicazioni Pratiche 10](#_Toc193188628)

[2.1.3 Azioni Necessarie 10](#_Toc193188629)

[2.2 Sicurezza Etica e Decisionale 11](#_Toc193188630)

[2.2.1 Meccanismi di Controllo 11](#_Toc193188631)

[2.2.2 Implicazioni Pratiche 11](#_Toc193188632)

[2.2.3 Azioni Necessarie 11](#_Toc193188633)

[2.3 Trasparenza e Spiegabilità 11](#_Toc193188634)

[2.3.1 Modelli Interpretabili 11](#_Toc193188635)

[2.3.2 Implicazioni Pratiche 11](#_Toc193188636)

[2.3.3 Azioni Necessarie 12](#_Toc193188637)

[2.4 Equità e Non Discriminazione 12](#_Toc193188638)

[2.4.1 Identificazione e Correzione dei Bias 12](#_Toc193188639)

[2.4.2 Implicazioni Pratiche 12](#_Toc193188640)

[2.4.3 Azioni Necessarie 12](#_Toc193188641)

[2.5 Privacy e Protezione dei Dati 12](#_Toc193188642)

[2.5.1 Minimizzazione dei Dati 12](#_Toc193188643)

[2.5.2 Implicazioni Pratiche 13](#_Toc193188644)

[2.5.3 Azioni Necessarie 13](#_Toc193188645)

[2.6 Benessere Sociale e Ambientale 13](#_Toc193188646)

[2.6.1 Impatto Sociale 13](#_Toc193188647)

[2.6.2 Impatto Ambientale 13](#_Toc193188648)

[2.6.3 Azioni Necessarie 13](#_Toc193188649)

[2.7 Responsabilità e Tracciabilità 13](#_Toc193188650)

[2.7.1 Meccanismi di Responsabilità 13](#_Toc193188651)

[2.7.2 Implicazioni Pratiche 14](#_Toc193188652)

[2.7.3 Azioni Necessarie 14](#_Toc193188653)

[3. Verifica del Codice: Garantire Sicurezza Etica e Decisionale 15](#_Toc193188654)

[3.1 Elementi di Sicurezza nel Codice 15](#_Toc193188655)

[3.1.1 Robustezza Tecnica 15](#_Toc193188656)

[3.1.2 Controlli Etici Integrati 15](#_Toc193188657)

[3.1.3 Tracciabilità delle Decisioni 16](#_Toc193188658)

[3.1.4 Spiegabilità 16](#_Toc193188659)

[3.2 Processi di Verifica Indipendenti 16](#_Toc193188660)

[3.2.1 Audit Regolari 17](#_Toc193188661)

[3.2.2 Certificazioni Etiche 17](#_Toc193188662)

[3.2.3 Linee Guida per gli Sviluppatori 18](#_Toc193188663)

[3.2.4 Coinvolgimento delle Parti Interessate 18](#_Toc193188664)

[3.2.5 Collaborazione Internazionale 19](#_Toc193188665)

[3.3 La Minaccia dell’Accelerazionismo 19](#_Toc193188666)

[3.3.1 Rischi dell’Accelerazionismo 19](#_Toc193188667)

[3.3.2 Concorrenza Sleale e Mancanza di Standard 20](#_Toc193188668)

[3.3.3 Impatto sulla Sicurezza Etica e Decisionale 20](#_Toc193188669)

[3.3.4 Azioni Necessarie per Contrastare l’Accelerazionismo 21](#_Toc193188670)

[4. Le Azioni Concrete per un’IA Etica e Human-Centric 22](#_Toc193188671)

[4.1 Un Quadro Normativo Completo 22](#_Toc193188672)

[4.1.1 Regolamenti Vincolanti 22](#_Toc193188673)

[4.1.2 Standard Tecnici 22](#_Toc193188674)

[4.1.3 Certificazioni Etiche 23](#_Toc193188675)

[4.2 Collaborazione Globale 23](#_Toc193188676)

[4.2.1 Forum Internazionali 23](#_Toc193188677)

[4.2.2 Standard Globali 23](#_Toc193188678)

[4.2.3 Condivisione di Risorse 24](#_Toc193188679)

[4.3 Educazione e Consapevolezza 24](#_Toc193188680)

[4.3.1 Programmi di Formazione 24](#_Toc193188681)

[4.3.2 Campagne di Sensibilizzazione 24](#_Toc193188682)

[4.3.3 Iniziative per il Pubblico 25](#_Toc193188683)

[4.4 Partecipazione Pubblica 25](#_Toc193188684)

[4.4.1 Panel Civici 25](#_Toc193188685)

[4.4.2 Consultazioni Pubbliche 25](#_Toc193188686)

[4.4.3 Meccanismi di Feedback 26](#_Toc193188687)

[5. Un Appello all’Azione 27](#_Toc193188688)

[6. Conclusione 28](#_Toc193188689)

## 1. Perché un Manifesto per un’IA Etica e Human-Centric?

L’intelligenza artificiale (IA) è una delle tecnologie più transformative del nostro tempo, con il potenziale di rivoluzionare settori come la sanità, l’istruzione, l’economia e la governance. Tuttavia, il suo sviluppo solleva importanti questioni etiche che devono essere affrontate per garantire che l’IA sia utilizzata in modo responsabile e sostenibile. Questo manifesto nasce dalla necessità di unire autorità accademiche, politiche e istituzionali in un impegno comune: **garantire che l’IA sia sviluppata e implementata in modo etico, trasparente e human-centric**. Di seguito, approfondiamo le ragioni che rendono questo manifesto non solo necessario, ma urgente.

### 1.1 L’Impatto Trasformativo dell’IA

L’IA sta trasformando ogni aspetto della nostra società, offrendo opportunità senza precedenti per migliorare la qualità della vita, risolvere problemi complessi e promuovere l’innovazione. Tuttavia, questa trasformazione non è priva di rischi.

#### 1.1.1 Opportunità dell’IA

* **Miglioramento della Qualità della Vita**: L’IA può rivoluzionare settori come la sanità, l’istruzione e i trasporti, offrendo soluzioni personalizzate e efficienti.
* **Soluzione di Problemi Globali**: L’IA può contribuire a risolvere problemi complessi, come il cambiamento climatico, la povertà e le disuguaglianze sociali.
* **Innovazione Economica**: L’IA può stimolare la crescita economica, creando nuovi posti di lavoro e opportunità di business.

#### 1.1.2 Rischi dell’IA

* **Amplificazione delle Disuguaglianze**: Senza controlli adeguati, l’IA può amplificare le disuguaglianze esistenti, creando nuove forme di esclusione sociale ed economica.
* **Violazione dei Diritti Umani**: L’IA può essere utilizzata per violare i diritti fondamentali, come la privacy, la libertà di espressione e la non discriminazione.
* **Perdita di Controllo**: Senza supervisione umana, l’IA può prendere decisioni dannose o imprevedibili, con conseguenze potenzialmente disastrose.

### 1.2 Le Sfide Etiche dell’IA

Lo sviluppo e l’uso dell’IA sollevano importanti questioni etiche che devono essere affrontate per garantire che questa tecnologia sia al servizio dell’umanità.

#### 1.2.1 Discriminazione e Bias

Gli algoritmi di IA possono perpetuare pregiudizi razziali, di genere o economici, amplificando disuguaglianze esistenti.

**Esempio**: Software di riconoscimento facciale con bias razziali, sistemi di reclutamento che discriminano le donne.

#### 1.2.2 Privacy e Sorveglianza

L’IA raccoglie e analizza enormi quantità di dati personali, aumentando i rischi di abuso e sorveglianza di massa.

**Esempio**: Sorveglianza di massa da parte di governi autoritari, uso improprio dei dati da parte delle aziende.

#### 1.2.3 Trasparenza e Responsabilità

Molti sistemi di IA sono “scatole nere”, rendendo difficile comprendere come prendono decisioni che influenzano la vita delle persone.

**Esempio**: Decisioni opache in contesti critici come la sanità o la giustizia.

#### 1.2.4 Impatto Occupazionale

L’automazione guidata dall’IA potrebbe eliminare milioni di posti di lavoro, creando nuove disuguaglianze sociali ed economiche.

**Esempio**: Perdita di posti di lavoro in settori come la manifattura, i trasporti e i servizi.

### 1.3 La Necessità di un Approccio Human-Centric

Un approccio human-centric all’IA significa mettere le persone al centro dello sviluppo e dell’uso di questa tecnologia, garantendo che sia al servizio dell’umanità e non il contrario.

#### 1.3.1 Supervisione Umana

Gli esseri umani devono mantenere il controllo sulle decisioni dell’IA, con la possibilità di intervenire, correggere o disattivare i sistemi in caso di necessità.

**Esempio**: Modelli di supervisione come “Human-in-the-loop” e “Human-on-the-loop”.

#### 1.3.2 Allineamento ai Valori Umani

L’IA deve essere progettata per rispettare i diritti fondamentali e i valori etici, come la privacy, l’equità e la non discriminazione.

**Esempio**: Integrazione di principi etici nel codice e nei processi decisionali.

#### 1.3.3 Benessere Sociale e Ambientale

L’IA deve contribuire al benessere della società e alla sostenibilità ambientale, promuovendo equità e riducendo gli impatti negativi.

**Esempio**: Applicazioni dell’IA per la riduzione delle emissioni di carbonio e la promozione dell’inclusione sociale.

### 1.4 Il Nuovo Significato di Formazione delle Nuove Generazioni

L’impatto cognitivo dell’IA sta trasformando il modo in cui le persone apprendono, pensano e interagiscono con il mondo. Questo richiede un nuovo approccio alla formazione delle nuove generazioni, che le prepari a vivere e lavorare in un mondo sempre più influenzato dall’intelligenza artificiale.

#### 1.4.1 Educazione all’IA

Le nuove generazioni devono essere educate sui principi etici e di sicurezza dell’IA, per comprendere sia le opportunità che i rischi di questa tecnologia.

**Programmi Scolastici**: Integrare l’educazione all’IA nei programmi scolastici, dalla scuola primaria all’università.

**Competenze Digitali**: Insegnare competenze digitali avanzate, come la programmazione e l’analisi dei dati.

**Pensiero Critico**: Promuovere il pensiero critico e la capacità di valutare l’impatto sociale dell’IA.

#### 1.4.2 Adattamento Cognitivo

L’IA sta cambiando il modo in cui le persone processano le informazioni e prendono decisioni. Questo richiede un adattamento cognitivo per evitare dipendenze tecnologiche e perdita di autonomia.

**Bilanciamento Uomo-Tecnologia**: Insegnare alle nuove generazioni a bilanciare l’uso della tecnologia con il pensiero autonomo e la creatività.

**Consapevolezza Digitale**: Promuovere la consapevolezza dei rischi cognitivi dell’IA, come la dipendenza da algoritmi e la perdita di privacy.

#### 1.4.3 Formazione Continua

In un mondo in rapida evoluzione, la formazione non può limitarsi ai primi anni di vita. È essenziale promuovere la formazione continua per adattarsi alle nuove sfide e opportunità dell’IA.

**Programmi di Aggiornamento**: Creare programmi di formazione continua per lavoratori e cittadini, focalizzati sulle nuove tecnologie e competenze.

**Apprendimento Permanente**: Promuovere una cultura dell’apprendimento permanente, che valorizzi l’aggiornamento costante delle competenze.

### 1.5 Il Pericolo dell’Utilizzo dell’IA per Scopi Militari

L’uso dell’IA in contesti militari rappresenta una delle applicazioni più controverse e pericolose di questa tecnologia. Senza un controllo umano adeguato, l’IA può essere utilizzata per prendere decisioni autonome su target umani e civili, con conseguenze potenzialmente catastrofiche.

#### 1.5.1 Armi Autonome e Decisioni Letali

Le armi autonome, conosciute anche come “robot killer”, sono sistemi di IA in grado di identificare, selezionare e attaccare target senza intervento umano. Questo solleva gravi preoccupazioni etiche e legali.

**Errori Irreversibili**: Un sistema di IA potrebbe identificare erroneamente un civile come un bersaglio militare, causando vittime innocenti.

**Escalation dei Conflitti**: L’uso di armi autonome potrebbe accelerare i conflitti, rendendo più difficile il controllo e la de-escalation.

**Responsabilità Opaca**: Senza un controllo umano diretto, è difficile attribuire la responsabilità per le decisioni letali prese dall’IA.

#### 1.5.2 Violazione del Diritto Internazionale

L’uso di armi autonome potrebbe violare i principi del diritto internazionale umanitario, come la distinzione tra combattenti e civili e il principio di proporzionalità.

**Indiscriminazione**: Un sistema di IA potrebbe non essere in grado di distinguere adeguatamente tra obiettivi militari e civili, portando a violazioni dei diritti umani.

**Mancanza di Empatia**: L’IA non ha la capacità di comprendere il contesto umano o di mostrare compassione, rendendola inadatta a prendere decisioni che coinvolgono la vita umana.

#### 1.5.3 Erosione del Controllo Umano

L’uso autonomo dell’IA in contesti militari erode il controllo umano sulle decisioni critiche, trasferendo la responsabilità a macchine che non possono essere ritenute moralmente o legalmente responsabili.

**Perdita di Autonomia**: La delega di decisioni letali all’IA riduce il ruolo degli esseri umani nel processo decisionale, minando i principi etici e democratici.

**Rischio di Abuso**: Governi o gruppi armati potrebbero utilizzare armi autonome per scopi repressivi o per commettere crimini di guerra senza accountability.

#### 1.5.4 Necessità di un Controllo Umano Finale

È essenziale che l’uomo mantenga l’ultima parola sulle decisioni che coinvolgono target umani e civili. Questo principio, noto come “meaningful human control”, garantisce che le decisioni letali siano sempre soggette a valutazione e approvazione umana.

**Divieto di Armi Autonome**: Promuovere un divieto globale sullo sviluppo e l’uso di armi autonome che possano prendere decisioni letali senza controllo umano.

**Regolamentazione Internazionale**: Sviluppare trattati e accordi internazionali che limitino l’uso dell’IA in contesti militari e garantiscano il rispetto del diritto internazionale umanitario.

**Supervisione Umana Obbligatoria**: Implementare meccanismi che garantiscano che ogni decisione che coinvolga target umani e civili sia soggetta a supervisione e approvazione umana.

#### 1.5.5 Implicazioni per la Sicurezza Globale

L’uso incontrollato dell’IA in contesti militari minaccia la sicurezza globale, aumentando il rischio di conflitti incontrollabili e di una corsa agli armamenti basata sull’IA.

**Corsa agli Armamenti**: I paesi potrebbero competere per sviluppare armi autonome sempre più avanzate, creando un’instabilità globale.

**Proliferazione delle Armi**: Le armi autonome potrebbero cadere nelle mani di gruppi armati non statali, aumentando il rischio di terrorismo e conflitti asimmetrici.

**Diplomazia Preventiva**: Promuovere il dialogo tra le nazioni per prevenire una corsa agli armamenti basata sull’IA.

**Controlli sulle Esportazioni**: Implementare controlli rigorosi sull’esportazione di tecnologie militari basate sull’IA.

## 2. I Principi Fondamentali del Manifesto

Proponiamo un quadro etico basato su sette principi fondamentali, che devono guidare lo sviluppo e l’implementazione dell’IA. Questi principi sono essenziali per garantire che l’IA sia al servizio dell’umanità, rispettando i diritti fondamentali e promuovendo il benessere sociale. Di seguito, approfondiamo ciascun principio in dettaglio.

### 2.1 Supervisione Umana

La supervisione umana è il principio cardine per garantire che l’IA rimanga uno strumento al servizio delle persone, e non il contrario. Questo principio richiede che gli esseri umani mantengano il controllo sulle decisioni dell’IA, con la possibilità di intervenire, correggere o disattivare i sistemi in caso di necessità.

#### 2.1.1 Livelli di Supervisione

* **Uomo nel loop (Human-in-the-loop)**: In questo modello, un operatore umano supervisiona attivamente le decisioni dell’IA e può correggerle in tempo reale. Questo approccio è particolarmente importante in contesti critici, come la sanità o la giustizia.
* **Uomo sopra il loop (Human-on-the-loop)**: L’IA opera in modo autonomo, ma un operatore umano può intervenire se il sistema commette errori o produce risultati inaccettabili. Questo modello è adatto per applicazioni come i veicoli autonomi.
* **Uomo fuori dal loop (Human-out-of-the-loop)**: L’IA agisce in modo completamente autonomo, senza intervento umano. Questo approccio è sconsigliato per applicazioni ad alto rischio, poiché riduce il controllo umano.

#### 2.1.2 Implicazioni Pratiche

* **Contesti Critici**: In settori come la sanità, la sicurezza e la giustizia, la supervisione umana è essenziale per evitare errori potenzialmente letali o dannosi.
* **Responsabilità**: La supervisione umana garantisce che ci sia sempre una persona responsabile delle decisioni dell’IA, facilitando la tracciabilità e la correzione degli errori.

#### 2.1.3 Azioni Necessarie

* **Progettazione di Interfacce Umane**: I sistemi di IA devono includere interfacce intuitive che permettano agli operatori umani di comprendere e intervenire facilmente.
* **Formazione degli Operatori**: Gli operatori umani devono essere adeguatamente formati per supervisionare e correggere i sistemi di IA.

### 2.2 Sicurezza Etica e Decisionale

La sicurezza etica e decisionale riguarda la capacità dell’IA di prendere decisioni allineate ai valori umani e ai diritti fondamentali, minimizzando i rischi di danni o violazioni.

#### 2.2.1 Meccanismi di Controllo

* **Limitazione delle Decisioni**: I sistemi di IA devono essere configurati per evitare decisioni estreme o pericolose, come quelle che potrebbero violare i diritti umani.
* **Allineamento ai Valori Umani**: Il codice deve incorporare principi etici, come il rispetto della privacy, l’equità e la non discriminazione, in ogni fase del processo decisionale.

#### 2.2.2 Implicazioni Pratiche

* **Prevenzione dei Danni**: Senza controlli etici, i sistemi di IA possono prendere decisioni dannose, come discriminazioni o violazioni della privacy.
* **Fiducia del Pubblico**: La sicurezza etica è essenziale per costruire fiducia nell’IA, favorendone l’adozione e l’integrazione nella società.

#### 2.2.3 Azioni Necessarie

* **Integrazione di Controlli Etici**: Gli sviluppatori devono integrare meccanismi di controllo etico nel codice, come strumenti per rilevare e correggere bias.
* **Valutazioni di Rischio**: Ogni sistema di IA deve essere sottoposto a valutazioni di rischio per identificare potenziali violazioni etiche.

### 2.3 Trasparenza e Spiegabilità

La trasparenza e la spiegabilità sono essenziali per garantire che i processi decisionali dell’IA siano comprensibili e accessibili, anche per utenti non tecnici.

#### 2.3.1 Modelli Interpretabili

* **Algoritmi Trasparenti**: I sistemi di IA devono utilizzare algoritmi che producono risultati comprensibili, evitando “scatole nere” inaccessibili.
* **Spiegazioni Chiare**: Le decisioni dell’IA devono essere accompagnate da spiegazioni intuitive, come visualizzazioni o linguaggio naturale.

#### 2.3.2 Implicazioni Pratiche

* **Fiducia e Accettazione**: La trasparenza è essenziale per costruire fiducia nell’IA, favorendone l’adozione da parte del pubblico.
* **Responsabilità**: La spiegabilità facilita la tracciabilità delle decisioni, permettendo di identificare e correggere errori.

#### 2.3.3 Azioni Necessarie

* **Standard di Trasparenza**: Devono essere definiti standard chiari per la trasparenza e la spiegabilità dei sistemi di IA.
* **Strumenti di Visualizzazione**: Gli sviluppatori devono creare strumenti che rendano comprensibili i processi decisionali dell’IA.

### 2.4 Equità e Non Discriminazione

L’equità e la non discriminazione sono principi fondamentali per garantire che l’IA non perpetui o amplifichi disuguaglianze esistenti.

#### 2.4.1 Identificazione e Correzione dei Bias

* **Dati di Addestramento**: I dati utilizzati per addestrare i sistemi di IA devono essere rappresentativi e privi di bias.
* **Algoritmi Imparziali**: Gli algoritmi devono essere progettati per evitare discriminazioni basate su genere, etnia, età o altre caratteristiche personali.

#### 2.4.2 Implicazioni Pratiche

* **Giustizia Sociale**: Senza equità, l’IA può amplificare disuguaglianze, creando ingiustizie sociali ed economiche.
* **Conformità Legale**: La non discriminazione è essenziale per garantire il rispetto delle leggi e dei regolamenti.

#### 2.4.3 Azioni Necessarie

* **Audit dei Dati**: I dati di addestramento devono essere sottoposti a audit regolari per identificare e correggere bias.
* **Test di Imparzialità**: I sistemi di IA devono essere testati per garantire che non producano decisioni discriminatorie.

### 2.5 Privacy e Protezione dei Dati

La privacy e la protezione dei dati sono essenziali per garantire che i dati personali siano gestiti in modo sicuro e trasparente.

#### 2.5.1 Minimizzazione dei Dati

* **Raccolta Limitata**: I dati personali devono essere raccolti solo quando strettamente necessario e per scopi specifici.
* **Consenso Informato**: Gli utenti devono essere informati su come i loro dati verranno utilizzati e devono dare il loro consenso esplicito.

#### 2.5.2 Implicazioni Pratiche

* **Protezione dei Diritti**: La privacy è un diritto fondamentale che deve essere rispettato in ogni fase del ciclo di vita dell’IA.
* **Fiducia del Pubblico**: La protezione dei dati è essenziale per costruire fiducia nell’IA.

#### 2.5.3 Azioni Necessarie

* **Conformità al GDPR**: I sistemi di IA devono rispettare il Regolamento Generale sulla Protezione dei Dati (GDPR) e altre normative sulla privacy.
* **Sicurezza dei Dati**: I dati personali devono essere protetti da accessi non autorizzati e violazioni.

### 2.6 Benessere Sociale e Ambientale

L’IA deve contribuire al benessere della società e alla sostenibilità ambientale, promuovendo equità e riducendo gli impatti negativi.

#### 2.6.1 Impatto Sociale

* **Equità e Inclusione**: L’IA deve essere utilizzata per promuovere equità e inclusione, migliorando la qualità della vita per tutti.
* **Riduzione delle Disuguaglianze**: L’IA può essere uno strumento per ridurre le disuguaglianze sociali ed economiche.

#### 2.6.2 Impatto Ambientale

* **Sostenibilità**: L’IA deve essere progettata per minimizzare il consumo energetico e l’impatto ambientale, ad esempio ottimizzando l’efficienza dei data center.

#### 2.6.3 Azioni Necessarie

* **Valutazioni di Impatto**: Ogni sistema di IA deve essere sottoposto a valutazioni di impatto sociale e ambientale.
* **Incentivi per la Sostenibilità**: Devono essere creati incentivi per promuovere lo sviluppo di IA sostenibile.

### 2.7 Responsabilità e Tracciabilità

La responsabilità e la tracciabilità sono essenziali per garantire che gli sviluppatori e gli operatori dell’IA siano responsabili delle loro creazioni.

#### 2.7.1 Meccanismi di Responsabilità

* **Tracciabilità delle Decisioni**: Le decisioni dell’IA devono essere registrate e tracciabili, permettendo di identificare e correggere errori.
* **Accountability**: Devono esistere meccanismi chiari per garantire che gli sviluppatori e gli operatori siano responsabili delle loro azioni.

#### 2.7.2 Implicazioni Pratiche

* **Correzione degli Errori**: La tracciabilità facilita la correzione degli errori e il miglioramento continuo dei sistemi di IA.
* **Conformità Legale**: La responsabilità è essenziale per garantire il rispetto delle leggi e dei regolamenti.

#### 2.7.3 Azioni Necessarie

* **Registri Dettagliati**: I sistemi di IA devono generare log chiari e completi che documentino le decisioni e i processi.
* **Audit Indipendenti**: Devono essere condotti audit regolari per verificare il rispetto dei principi etici.

## 3. Verifica del Codice: Garantire Sicurezza Etica e Decisionale

Uno degli aspetti più critici per garantire un’IA etica e human-centric è la **verifica del codice**. Il codice che alimenta i sistemi di IA deve essere progettato e testato per garantire che rispetti i principi etici e affronti le sfide pratiche descritte in questo manifesto.

### 3.1 Elementi di Sicurezza nel Codice

Il codice dei sistemi di IA deve includere meccanismi specifici per garantire che le decisioni siano sicure, trasparenti e allineate ai valori umani. Di seguito, approfondiamo i quattro elementi chiave che devono essere integrati nel codice.

#### 3.1.1 Robustezza Tecnica

La robustezza tecnica è fondamentale per garantire che i sistemi di IA siano resistenti a errori, manipolazioni e attacchi informatici. Questo elemento include:

* **Resistenza agli Errori**: Il codice deve essere progettato per gestire situazioni impreviste, come input anomali o dati corrotti, senza produrre risultati errati o dannosi.
* **Protezione dagli Attacchi**: I sistemi di IA devono essere protetti da attacchi informatici, come il *data poisoning* (manipolazione dei dati di addestramento) o gli *adversarial attacks* (input progettati per ingannare l’IA).
* **Continuità del Servizio**: Il codice deve garantire che il sistema continui a funzionare correttamente anche in condizioni di stress o sotto carico elevato.

Senza robustezza tecnica, i sistemi di IA possono fallire in situazioni critiche, come quelle mediche o di sicurezza, con conseguenze potenzialmente disastrose.

#### 3.1.2 Controlli Etici Integrati

I controlli etici integrati sono meccanismi progettati per prevenire decisioni discriminatorie, dannose o non allineate ai valori umani. Questo include:

* **Rilevazione dei Bias**: Il codice deve includere strumenti per identificare e correggere bias nei dati di addestramento e nei processi decisionali.
* **Limitazione delle Decisioni**: I sistemi di IA devono essere configurati per evitare decisioni estreme o pericolose, come quelle che potrebbero violare i diritti umani.
* **Allineamento ai Valori Umani**: Il codice deve incorporare principi etici, come il rispetto della privacy, l’equità e la non discriminazione, in ogni fase del processo decisionale.

Senza controlli etici, i sistemi di IA possono perpetuare discriminazioni o prendere decisioni dannose, minando la fiducia del pubblico e causando danni sociali.

#### 3.1.3 Tracciabilità delle Decisioni

La tracciabilità è essenziale per garantire che le decisioni dell’IA possano essere ricostruite e verificate. Questo elemento include:

* **Registri Dettagliati**: Il codice deve generare log chiari e completi che documentino come e perché una decisione è stata presa.
* **Identificazione delle Fonti**: Le decisioni devono essere collegabili ai dati e ai processi che le hanno generate, permettendo di risalire all’origine di eventuali errori o bias.
* **Auditabilità**: Il codice deve essere progettato per facilitare audit esterni, permettendo a terze parti di verificare il rispetto dei principi etici.

Senza tracciabilità, è impossibile identificare e correggere errori o bias, rendendo i sistemi di IA opachi e difficili da regolamentare.

#### 3.1.4 Spiegabilità

La spiegabilità è la capacità di rendere comprensibili i processi decisionali dell’IA, anche per utenti non tecnici. Questo elemento include:

* **Modelli Interpretabili**: Il codice deve utilizzare algoritmi che producono risultati comprensibili, evitando “scatole nere” inaccessibili.
* **Spiegazioni Chiare**: I sistemi di IA devono fornire spiegazioni intuitive delle loro decisioni, ad esempio attraverso visualizzazioni o linguaggio naturale.
* **Interazione Umana**: Il codice deve facilitare l’interazione tra l’IA e gli utenti, permettendo a questi ultimi di porre domande e ottenere risposte chiare.

Senza spiegabilità, gli utenti non possono fidarsi dei sistemi di IA, limitandone l’adozione e aumentando il rischio di decisioni errate o dannose.

### 3.2 Processi di Verifica Indipendenti

Per garantire che i sistemi di IA rispettino i principi etici e di sicurezza, è essenziale istituire **processi di verifica indipendenti**. Questi processi devono essere condotti da organismi terzi, privi di conflitti di interesse, e devono includere audit regolari, certificazioni e linee guida chiare per gli sviluppatori. Di seguito, approfondiamo i componenti chiave di questi processi.

#### 3.2.1 Audit Regolari

Gli audit regolari sono fondamentali per garantire che i sistemi di IA rispettino i principi etici e di sicurezza nel tempo. Questi audit devono essere condotti da esperti indipendenti e devono coprire tutti gli aspetti del sistema, dal codice ai dati di addestramento.

Gli audit regolari servono a identificare potenziali violazioni dei principi etici, come bias, discriminazioni o mancanza di trasparenza. Devono essere condotti periodicamente e in modo approfondito, utilizzando metodologie standardizzate.

* **Frequenza**: Gli audit devono essere condotti almeno una volta all’anno o dopo ogni aggiornamento significativo del sistema.
* **Metodologie Standardizzate**: Devono essere utilizzati protocolli riconosciuti a livello internazionale per garantire coerenza e affidabilità.
* **Copertura Completa**: Gli audit devono coprire tutti gli aspetti del sistema, inclusi dati, algoritmi, interfacce e processi decisionali.
* **Report Chiari**: I risultati degli audit devono essere documentati in report dettagliati, accessibili alle autorità competenti e al pubblico.

#### 3.2.2 Certificazioni Etiche

Le certificazioni etiche sono un modo per riconoscere pubblicamente i sistemi di IA che rispettano i principi etici e di sicurezza. Queste certificazioni devono essere rilasciate da organismi indipendenti e devono essere basate su criteri rigorosi.

Le certificazioni etiche servono a costruire fiducia nel pubblico e a incentivare le aziende a sviluppare sistemi di IA responsabili. Devono essere basate su standard chiari e devono essere rinnovate periodicamente.

* **Criteri Rigorosi**: Le certificazioni devono essere basate su criteri chiari e misurabili, come l’assenza di bias, la trasparenza e la sicurezza.
* **Organismi Indipendenti**: Le certificazioni devono essere rilasciate da organismi privi di conflitti di interesse.
* **Rinnovo Periodico**: Le certificazioni devono essere rinnovate periodicamente per garantire che i sistemi continuino a rispettare i principi etici.
* **Pubblicità**: Le aziende certificate devono poter utilizzare un logo o un marchio riconoscibile per comunicare il loro impegno etico.

#### 3.2.3 Linee Guida per gli Sviluppatori

Le linee guida per gli sviluppatori sono essenziali per garantire che i principi etici siano integrati fin dalle prime fasi dello sviluppo dell’IA. Queste linee guida devono essere chiare, pratiche e basate su esempi concreti.

Le linee guida servono a orientare gli sviluppatori nella progettazione e implementazione di sistemi di IA etici. Devono coprire tutti gli aspetti dello sviluppo, dalla raccolta dei dati alla progettazione degli algoritmi.

* **Esempi Pratici**: Le linee guida devono includere esempi concreti di come integrare i principi etici nel codice.
* **Strumenti e Risorse**: Devono essere forniti strumenti e risorse, come librerie software e checklist, per facilitare l’implementazione.
* **Formazione**: Gli sviluppatori devono essere formati sull’uso delle linee guida e sui principi etici dell’IA.
* **Aggiornamenti Regolari**: Le linee guida devono essere aggiornate regolarmente per riflettere le nuove sfide e tecnologie.

#### 3.2.4 Coinvolgimento delle Parti Interessate

Il coinvolgimento delle parti interessate, come cittadini, organizzazioni della società civile e autorità pubbliche, è essenziale per garantire che i processi di verifica riflettano le esigenze e i valori della società.

Il coinvolgimento delle parti interessate serve a garantire che i processi di verifica siano inclusivi e rappresentativi. Questo approccio promuove la trasparenza e la fiducia del pubblico.

* **Panel Civici**: Devono essere istituiti panel civici per coinvolgere i cittadini nella definizione delle regole e delle politiche sull’IA.
* **Consultazioni Pubbliche**: Devono essere organizzate consultazioni pubbliche per raccogliere feedback e suggerimenti.
* **Rappresentanza Equilibrata**: Le parti interessate devono includere rappresentanti di diverse comunità, inclusi gruppi vulnerabili.
* **Trasparenza**: I risultati delle consultazioni devono essere pubblicati e resi accessibili al pubblico.

#### 3.2.5 Collaborazione Internazionale

La collaborazione internazionale è essenziale per garantire che i processi di verifica siano armonizzati a livello globale e per promuovere standard etici comuni.

La collaborazione internazionale serve a evitare frammentazione e concorrenza sleale, promuovendo un approccio coordinato alla regolamentazione dell’IA. Questo include la condivisione di best practice e la creazione di standard globali.

* **Forum Internazionali**: Devono essere istituiti forum internazionali per discutere e armonizzare le regolamentazioni.
* **Standard Globali**: Devono essere sviluppati standard globali per la verifica e la certificazione dei sistemi di IA.
* **Condivisione di Risorse**: I paesi devono condividere risorse, come strumenti di audit e linee guida, per facilitare l’implementazione.
* **Accordi di Collaborazione**: Devono essere stipulati accordi di collaborazione tra paesi e organizzazioni internazionali.

### 3.3 La Minaccia dell’Accelerazionismo

L’accelerazionismo nello sviluppo dell’IA, alimentato dalla corsa tra superpotenze per il dominio tecnologico, rappresenta una delle maggiori minacce alla sicurezza etica e decisionale dei sistemi di IA. Questo approccio, focalizzato sulla velocità e sulla competizione, riduce drasticamente la possibilità che i processi di sicurezza e i principi etici siano adeguatamente sviluppati e implementati. Di seguito, approfondiamo i rischi e le azioni necessarie per contrastare questa tendenza.

#### 3.3.1 Rischi dell’Accelerazionismo

L’accelerazionismo porta a un rilascio prematuro di sistemi di IA, spesso senza adeguati controlli di sicurezza e etici. Questo approccio può avere conseguenze disastrose per la società e i diritti umani.

La pressione per essere i primi a mercato spinge le aziende e i governi a trascurare gli aspetti etici e di sicurezza, rilasciando sistemi di IA che non sono stati adeguatamente testati o validati. Questo aumenta il rischio di errori, violazioni della privacy e discriminazioni.

* **Rilascio Prematuro**: I sistemi di IA vengono rilasciati prima di essere completamente testati, aumentando il rischio di errori e decisioni dannose.Senza test approfonditi, i sistemi di IA possono produrre risultati imprevedibili o pericolosi, specialmente in contesti critici come la sanità o la sicurezza.
* **Mancanza di Controlli Etici**: Gli aspetti etici, come la trasparenza e la non discriminazione, vengono spesso trascurati per accelerare lo sviluppo. Questo porta a sistemi che perpetuano bias o violano i diritti fondamentali, minando la fiducia del pubblico.
* **Sicurezza Compromessa**: La fretta nello sviluppo può portare a vulnerabilità tecniche, esponendo i sistemi di IA ad attacchi informatici. Sistemi non sicuri possono essere manipolati da attori malevoli, con conseguenze potenzialmente disastrose.

#### 3.3.2 Concorrenza Sleale e Mancanza di Standard

La corsa tra superpotenze crea un ambiente in cui la concorrenza sleale prevale sulla collaborazione, rendendo difficile l’adozione di standard etici e di sicurezza comuni.

La competizione per il dominio tecnologico spinge i paesi e le aziende a ignorare gli standard etici e di sicurezza, privilegiando la velocità e l’efficienza. Questo crea un ambiente in cui chi rispetta le regole è svantaggiato.

* **Divergenza negli Standard**: I paesi adottano approcci diversi alla regolamentazione dell’IA, creando frammentazione e incoerenza, ma senza standard comuni, è difficile garantire che i sistemi di IA rispettino i principi etici a livello globale.
* **Concorrenza Basata sulla Velocità**: Le aziende sono incentivate a rilasciare prodotti rapidamente, spesso a scapito della qualità e della sicurezza. Questo approccio favorisce chi ignora i controlli etici, creando un mercato distorto.
* **Mancanza di Collaborazione**: La competizione tra superpotenze riduce la possibilità di collaborazione internazionale su questioni etiche e di sicurezza.Senza collaborazione, è difficile sviluppare soluzioni globali alle sfide dell’IA.

#### 3.3.3 Impatto sulla Sicurezza Etica e Decisionale

L’accelerazionismo compromette la sicurezza etica e decisionale dei sistemi di IA, aumentando il rischio di decisioni dannose o discriminatorie.

La fretta nello sviluppo riduce la possibilità di integrare controlli etici e di sicurezza nel codice, aumentando il rischio che i sistemi di IA prendano decisioni dannose o violino i diritti umani.

* **Decisioni Non Allineate ai Valori Umani**: I sistemi di IA possono prendere decisioni che non rispettano i diritti fondamentali o i valori etici.  
  Questo è particolarmente problematico in contesti come la giustizia o l’assistenza sanitaria, dove le decisioni hanno un impatto diretto sulle persone.
* **Aumento dei Bias e delle Discriminazioni**: La mancanza di controlli etici aumenta il rischio che i sistemi di IA perpetuino bias e discriminazioni, ciò può portare a ingiustizie sociali ed economiche, amplificando le disuguaglianze esistenti.
* **Riduzione della Trasparenza**: L’accelerazionismo spesso porta a sistemi di IA opachi, difficili da comprendere o verificare. Senza trasparenza, è impossibile garantire che le decisioni dell’IA siano giuste e responsabili.

#### 3.3.4 Azioni Necessarie per Contrastare l’Accelerazionismo

Per contrastare l’accelerazionismo, è essenziale adottare un approccio globale e coordinato, che promuova la collaborazione e la responsabilità.

Le azioni necessarie includono la creazione di standard internazionali, la promozione della collaborazione tra paesi e l’adozione di politiche che incentivino lo sviluppo etico dell’IA.

* **Standard Internazionali**: Devono essere sviluppati standard globali per la sicurezza e l’etica dell’IA, talistandard devono essere adottati da tutti i paesi e le aziende, garantendo coerenza e responsabilità.
* **Collaborazione Globale**: I paesi devono collaborare per affrontare le sfide dell’IA in modo coordinato, ciò include la condivisione di best practice, risorse e strumenti per lo sviluppo etico dell’IA.
* **Incentivi per lo Sviluppo Etico**: I governi devono creare incentivi per le aziende che sviluppano sistemi di IA etici e sicuri, includendo finanziamenti, agevolazioni fiscali o priorità negli appalti pubblici.
* **Educazione e Consapevolezza**: È essenziale promuovere l’educazione e la consapevolezza sui rischi dell’accelerazionismo con programmi di formazione per sviluppatori, decisori politici e cittadini

## 4. Le Azioni Concrete per un’IA Etica e Human-Centric

Per tradurre i principi etici in pratica e garantire che l’IA sia sviluppata e utilizzata in modo responsabile, è necessario adottare un piano d’azione concreto e strutturato. Questo piano si articola in quattro pilastri principali, ciascuno dei quali include azioni specifiche e responsabilità chiare. Di seguito, approfondiamo ciascun pilastro con descrizioni dettagliate e sottoparagrafi esplicativi.

### 4.1 Un Quadro Normativo Completo

Un quadro normativo completo è essenziale per garantire che i sistemi di IA rispettino i principi etici e di sicurezza. Questo quadro deve includere regolamenti vincolanti, standard tecnici e meccanismi di certificazione.

#### 4.1.1 Regolamenti Vincolanti

I regolamenti devono essere legalmente vincolanti e applicabili a tutti i sistemi di IA, indipendentemente dal settore o dal contesto di utilizzo. Devono coprire aspetti come la supervisione umana, la trasparenza, la non discriminazione e la protezione dei dati.

**Definizione Chiara dei Requisiti**: I regolamenti devono specificare in modo chiaro i requisiti etici e di sicurezza che i sistemi di IA devono rispettare.

**Meccanismi di Applicazione**: Devono essere istituiti organismi di controllo incaricati di far rispettare i regolamenti e di sanzionare le violazioni.

**Aggiornamenti Periodici**: I regolamenti devono essere aggiornati regolarmente per tenere conto delle nuove tecnologie e delle sfide emergenti.

#### 4.1.2 Standard Tecnici

Gli standard tecnici sono essenziali per garantire che i sistemi di IA siano progettati e implementati in modo coerente e sicuro. Questi standard devono coprire aspetti come la robustezza tecnica, la trasparenza degli algoritmi e la gestione dei dati.

**Standard Globali**: Gli standard devono essere sviluppati in collaborazione con organizzazioni internazionali per garantire coerenza a livello globale.

**Adozione Obbligatoria**: Gli standard devono essere adottati obbligatoriamente da tutti gli sviluppatori e operatori di IA.

**Strumenti di Supporto**: Devono essere forniti strumenti e risorse per aiutare gli sviluppatori a conformarsi agli standard.

#### 4.1.3 Certificazioni Etiche

Le certificazioni etiche sono un modo per riconoscere pubblicamente i sistemi di IA che rispettano i principi etici e di sicurezza. Queste certificazioni devono essere rilasciate da organismi indipendenti e basate su criteri rigorosi.

**Criteri Chiari**: Le certificazioni devono essere basate su criteri chiari e misurabili, come l’assenza di bias, la trasparenza e la sicurezza.

**Processo di Verifica**: Le certificazioni devono essere rilasciate solo dopo un processo di verifica indipendente e approfondito.

**Pubblicità**: Le aziende certificate devono poter utilizzare un logo o un marchio riconoscibile per comunicare il loro impegno etico.

### 4.2 Collaborazione Globale

La collaborazione globale è essenziale per garantire che i principi etici e di sicurezza siano rispettati a livello internazionale. Questo pilastro include la creazione di forum internazionali, lo sviluppo di standard globali e la condivisione di risorse.

#### 4.2.1 Forum Internazionali

I forum internazionali sono piattaforme per discutere e armonizzare le regolamentazioni sull’IA. Questi forum devono coinvolgere rappresentanti di governi, aziende, organizzazioni della società civile e esperti.

**Partecipazione Inclusiva**: I forum devono includere rappresentanti di tutti i paesi e le comunità, garantendo che le decisioni riflettano le esigenze globali.

**Agenda Chiara**: I forum devono avere un’agenda chiara e focalizzata su questioni specifiche, come la sicurezza, la trasparenza e la non discriminazione.

**Risultati Tangibili**: I forum devono produrre risultati tangibili, come accordi internazionali o linee guida comuni.

#### 4.2.2 Standard Globali

Gli standard globali sono essenziali per garantire coerenza e responsabilità nello sviluppo e nell’uso dell’IA. Questi standard devono coprire aspetti come la sicurezza, la trasparenza e la protezione dei dati.

**Sviluppo Collaborativo**: Gli standard devono essere sviluppati in collaborazione con organizzazioni internazionali, governi e aziende.

**Adozione Obbligatoria**: Gli standard devono essere adottati obbligatoriamente da tutti i paesi e le aziende.

**Monitoraggio e Aggiornamento**: Gli standard devono essere monitorati e aggiornati regolarmente per tenere conto delle nuove tecnologie e delle sfide emergenti.

#### 4.2.3 Condivisione di Risorse

La condivisione di risorse, come strumenti di audit, linee guida e best practice, è essenziale per facilitare l’implementazione dei principi etici e di sicurezza.

**Piattaforme di Condivisione**: Devono essere create piattaforme online per condividere risorse e strumenti tra paesi e organizzazioni.

**Formazione e Supporto**: Devono essere forniti programmi di formazione e supporto per aiutare gli sviluppatori e i decisori politici a utilizzare le risorse.

**Accesso Aperto**: Le risorse devono essere accessibili a tutti, garantendo che anche i paesi in via di sviluppo possano beneficiarne.

### 4.3 Educazione e Consapevolezza

L’educazione e la consapevolezza sono essenziali per garantire che tutti gli attori coinvolti nello sviluppo e nell’uso dell’IA comprendano i principi etici e di sicurezza. Questo pilastro include programmi di formazione, campagne di sensibilizzazione e iniziative per il pubblico.

#### 4.3.1 Programmi di Formazione

I programmi di formazione devono essere sviluppati per educare gli sviluppatori, i decisori politici e i cittadini sui principi etici e di sicurezza dell’IA.

**Contenuti Completi**: I programmi di formazione devono coprire tutti gli aspetti dell’IA, dalla progettazione all’implementazione.

**Accessibilità**: I programmi di formazione devono essere accessibili a tutti, inclusi i paesi in via di sviluppo.

**Aggiornamenti Regolari**: I programmi di formazione devono essere aggiornati regolarmente per riflettere le nuove tecnologie e le sfide emergenti.

#### 4.3.2 Campagne di Sensibilizzazione

Le campagne di sensibilizzazione devono essere lanciate per aumentare la consapevolezza del pubblico sui rischi e le opportunità dell’IA.

**Messaggi Chiari**: Le campagne devono utilizzare messaggi chiari e comprensibili per raggiungere un pubblico ampio.

**Coinvolgimento del Pubblico**: Le campagne devono coinvolgere il pubblico attraverso eventi, workshop e discussioni.

**Valutazione dell’Impatto**: L’efficacia delle campagne deve essere valutata regolarmente per garantire che raggiungano gli obiettivi prefissati.

#### 4.3.3 Iniziative per il Pubblico

Le iniziative per il pubblico, come panel civici e consultazioni, sono essenziali per garantire che le decisioni sull’IA riflettano i valori e le esigenze della società.

**Partecipazione Inclusiva**: Le iniziative devono coinvolgere rappresentanti di tutte le comunità, inclusi gruppi vulnerabili.

**Trasparenza**: I risultati delle iniziative devono essere pubblicati e resi accessibili al pubblico.

**Impatto sulle Politiche**: Le iniziative devono avere un impatto tangibile sulle politiche e le regolamentazioni sull’IA.

### 4.4 Partecipazione Pubblica

La partecipazione pubblica è essenziale per garantire che le decisioni sull’IA riflettano i valori e le esigenze della società. Questo pilastro include la creazione di panel civici, consultazioni pubbliche e meccanismi di feedback.

#### 4.4.1 Panel Civici

I panel civici sono gruppi di cittadini incaricati di discutere e fornire feedback sulle politiche e le regolamentazioni sull’IA.

**Rappresentanza Equilibrata**: I panel devono includere rappresentanti di tutte le comunità, garantendo che le decisioni riflettano le esigenze globali.

**Formazione Adeguata**: I partecipanti devono essere formati sui principi etici e di sicurezza dell’IA.

**Impatto sulle Decisioni**: I feedback dei panel devono avere un impatto tangibile sulle decisioni politiche.

#### 4.4.2 Consultazioni Pubbliche

Le consultazioni pubbliche sono processi per raccogliere feedback e suggerimenti dai cittadini sulle politiche e le regolamentazioni sull’IA.

**Accessibilità**: Le consultazioni devono essere accessibili a tutti, inclusi i cittadini con disabilità o con limitato accesso alla tecnologia.

**Trasparenza**: I risultati delle consultazioni devono essere pubblicati e resi accessibili al pubblico.

**Risposte Concrete**: I decisori politici devono rispondere ai feedback ricevuti, spiegando come sono stati integrati nelle decisioni.

#### 4.4.3 Meccanismi di Feedback

I meccanismi di feedback sono strumenti per raccogliere e analizzare le opinioni del pubblico sull’uso e l’impatto dell’IA.

**Facilità d’Uso**: I meccanismi di feedback devono essere facili da usare e accessibili a tutti.

**Analisi dei Dati**: I feedback raccolti devono essere analizzati per identificare tendenze e problemi comuni.

**Risposte Tempestive**: I decisori politici devono rispondere ai feedback in modo tempestivo, dimostrando che le opinioni del pubblico sono prese in considerazione.

## 5. Un Appello all’Azione

Il futuro dell’IA dipende dalle decisioni che prendiamo oggi. Senza un’azione tempestiva, rischiamo di creare sistemi che minacciano i diritti umani, amplificano disuguaglianze e compromettono la sicurezza delle nostre società.

Chiediamo alle autorità accademiche, politiche e istituzionali di unirsi a noi in questo impegno. Firmando questo manifesto, vi impegnate a:

1. Promuovere un’IA etica e human-centric.

2. Sostenere l’adozione di un quadro normativo completo che includa controlli sulla sicurezza etica e decisionale.

3. Collaborare a livello globale per garantire che l’IA sia sviluppata e utilizzata in modo responsabile.

L’Unione Europea ha l’opportunità di essere un leader globale nell’etica dell’IA. Agiamo ora per garantire che l’IA sia una forza positiva per l’umanità.

## 6. Conclusione

L’IA è una tecnologia potente, ma il suo impatto dipende da come la sviluppiamo e la utilizziamo. Questo manifesto rappresenta un appello all’azione per garantire che l’IA sia al servizio dell’umanità, rispettando i diritti fondamentali e promuovendo il benessere sociale.

Unitevi a noi in questo impegno. Insieme, possiamo costruire un futuro in cui l’IA sia etica, trasparente e human-centric.

**Firmato,**

[Le autorità accademiche, politiche e istituzionali che aderiscono al manifesto]