

I

CONSIGLIO NAZIONALE
DEGLI INGEGNERI



Costruiamoci il futuro Intelligenza Artificiale: un approccio etico

Roberto Magnani

19 Febbraio 2024



Struttura dei Seminari

COSTRUIAMOCI IL FUTURO

Utilizzi pratici delle Intelligenze Artificiali

22 Gennaio

intelligenza artificiale
nell'ingegneria e
nell'industria

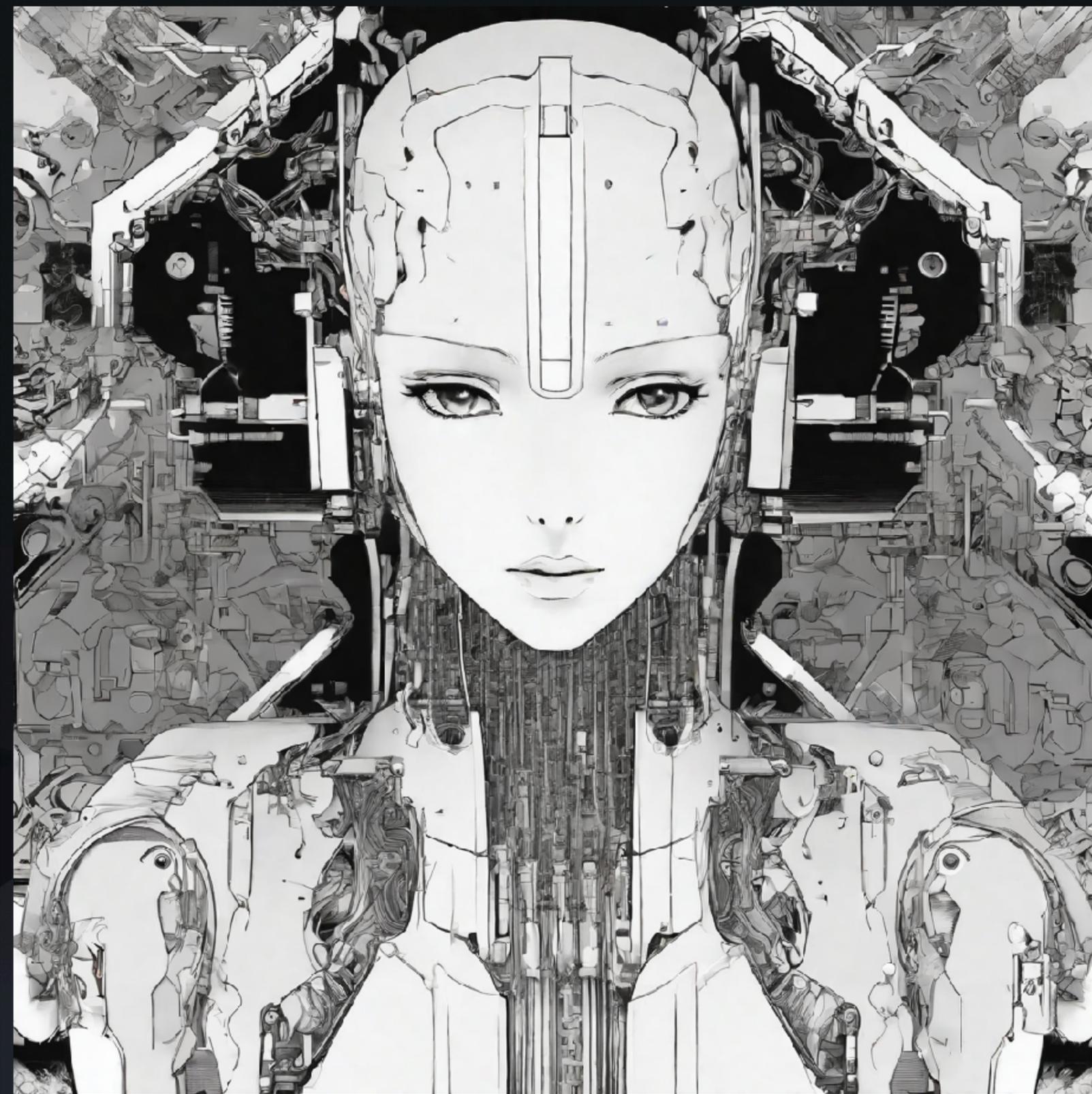
19 Febbraio

Uso etico
dell'intelligenza
artificiale



Cosa vedremo oggi

- Intelligenza artificiale tra scienza e tecnologia:
- Sfide etiche, Equità, Trasparenza Responsabilità e Sicurezza
- Opacità da interdipendenza tecno-sociale
- Come rilevare i pregiudizi nei processi con intelligenza artificiale
- Approccio ex ante / intra/ex post in fase di progettazione.
- Intelligenza artificiale responsabile
- Il Regolamento UE: AI ACT, DMA , DSA
- Il ruolo del Digital Ethics Officer nelle organizzazioni industriali



La situazione sociale e qualche mito da sfatare

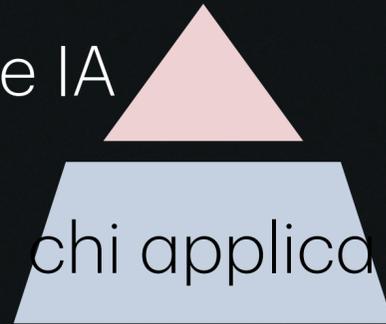
La situazione attuale

Tecnologia in mano a pochi

Pochi specialisti

Molti utilizzatori non perfettamente coscienti dell'uso

chi detiene IA



chi applica



Sistemi con Intelligenza Artificiale



chi usa non coscientemente

chi non usa ancora

Dove dovremmo arrivare

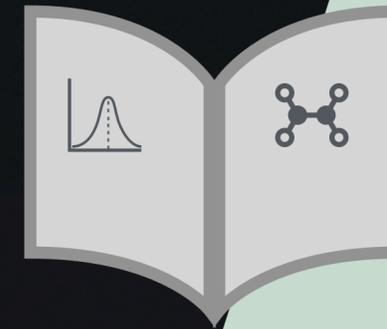
IA sicura

Uso consapevole

Supporto alle professioni

Benefici sociali per tutti

Creare
Consapevolezza



chi detiene IA

chi applica IA

chi usa
coscientemente

Sistemi con Intelligenza Artificiale

chi usa non scientemente

o non usa ancora



Tra scienza e tecnologia

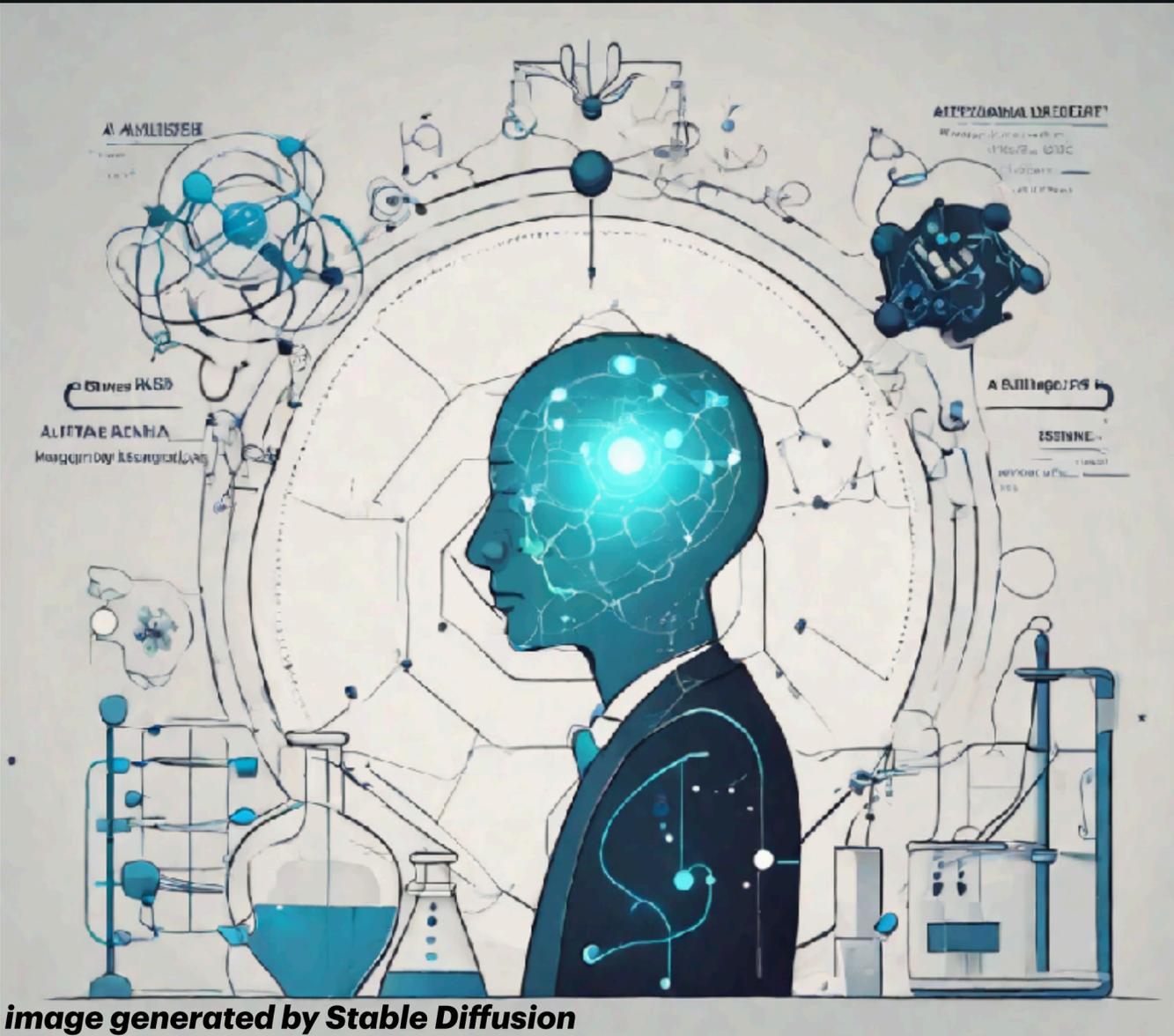


image generated by Stable Diffusion

- Il pubblico deve essere consapevole dei potenziali rischi e deve essere in grado di prendere decisioni informate su come agire
- Con la IA è in atto un cambio di paradigma che dovrebbe vedere tutti coinvolti nell'indirizzare gli sviluppi appropriati delle scienze
- L'IA e i sistemi automatizzati possono migliorare situazioni critiche ma anche approfondire ingiustizia sociale

Il mito del “Tech Solutioning”

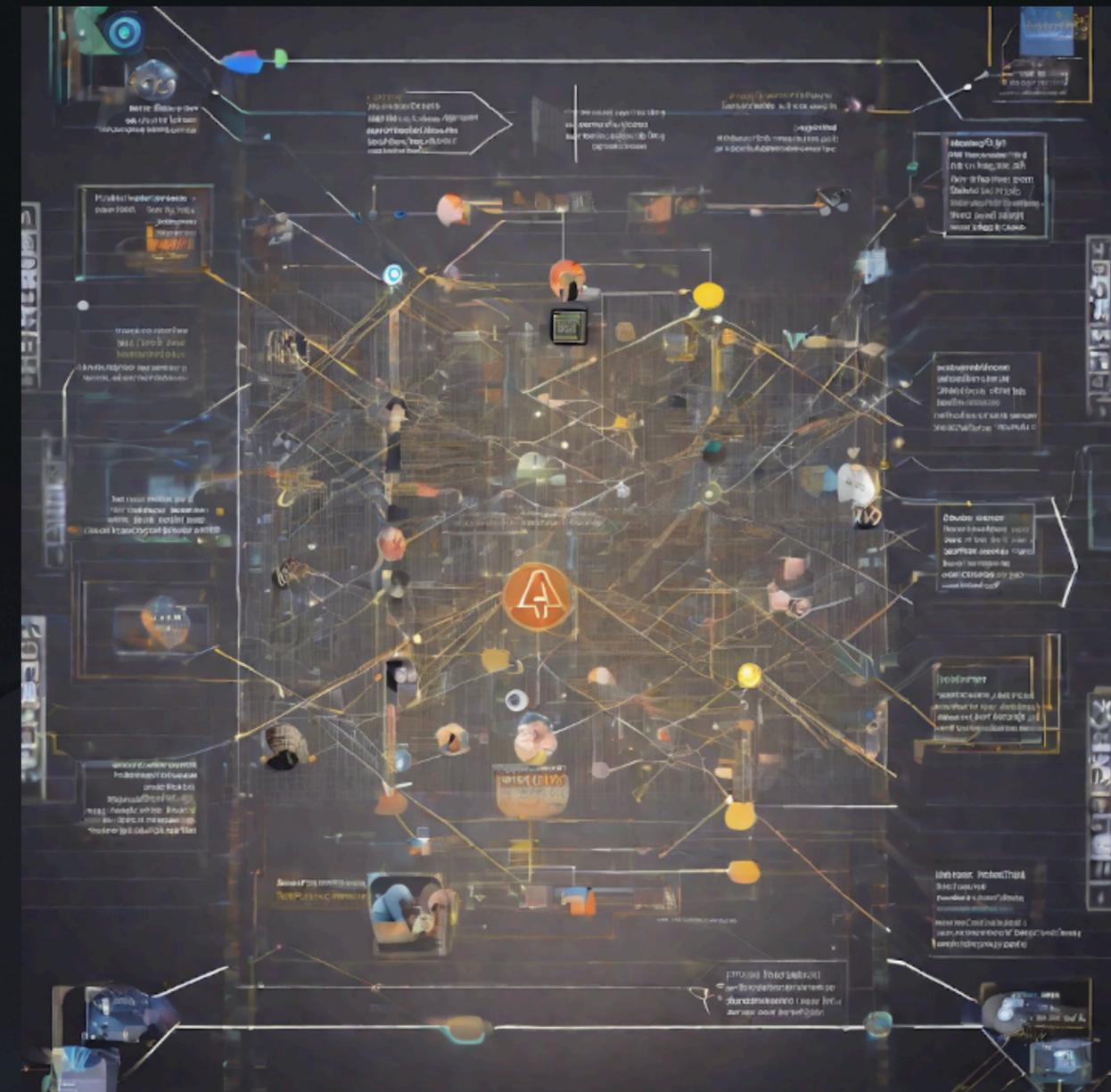
La presunzione che la tecnologia possa risolvere qualsiasi problema

- **La tecnologia necessita di un indirizzo preventivo**

- Esempio. È reale il rischio di prevaricazione sui Paesi poveri per il reperimento dei materiali vitali per la costruzione dei mezzi “hardware”

- È necessario un **ethos** visionario per intuire l'evoluzione dei progetti tecnologici

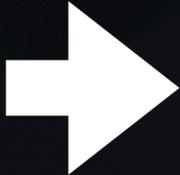
- **La responsabilità e trasparenza di sviluppatori e utenti di IA sono fondamentali per promuovere inclusione e non discriminazione**



Principi Etici per l'Intelligenza Artificiale rilasciati dall'ONU

- 1. Umanità:** L'IA dovrebbe essere sviluppata e utilizzata per migliorare il benessere umano, rispettando la dignità, i diritti, la libertà e la diversità di tutte le persone.
- 2. Giustizia:** Le tecnologie IA dovrebbero essere equamente distribuite, garantendo che tutti possano accedervi e beneficiarne senza discriminazioni o ingiustizie.
- 3. Trasparenza:** Le decisioni e le azioni dell'IA devono essere comprensibili e tracciabili, consentendo alle persone di comprendere le logiche e le implicazioni delle decisioni prese dall'IA.
- 4. Responsabilità:** Gli sviluppatori e gli utilizzatori di IA devono assumersi la responsabilità delle azioni e delle conseguenze delle tecnologie AI, garantendo che siano conformi agli standard etici e legali.
- 5. Non danneggiare:** L'IA non dovrebbe essere utilizzata per causare danni fisici o psicologici alle persone, né per violare i diritti umani o indebolire la sicurezza individuale o collettiva.
- 6. Privacy e Dati:** L'IA dovrebbe rispettare e proteggere la privacy delle persone e garantire la sicurezza e l'integrità dei dati personali.
- 7. Ambiente e Sviluppo Sostenibile:** Lo sviluppo e l'uso dell'IA dovrebbero contribuire alla protezione dell'ambiente e alla promozione dello sviluppo sostenibile.

La situazione e l'utilità delle norme

- Senza tecnologie proprietarie, il **53%** delle società si affida a prodotti di **IA di terze parti** ("Building Robust RIA Programs as Third-Party IA Tools Proliferate," 2023 - del MIT Sloan Management Review /con Boston Consulting Group),
 - su 1.240 rappresentanti di rilievo di società con un fatturato annuo di almeno 100 milioni di dollari, sparse su 59 settori in 87 Paesi,
 - aumento nella gestione della cosiddetta **IA responsabile dal 16% al 29%**.
 - Il **71%** delle organizzazioni **non supervisiona** i propri processi di attuazione
 - **3/4 si deve affidare a terze parti** per l'uso della tecnologia per mancanza di competenze interne,
 - rischi di reputazione
 - perdita di fiducia dei clienti
 - perdite finanziarie
 - problemi legali
-  **Alla ricerca di linee guida**

Quando l'etica supera la norma

- È in atto la ricerca di un'eticità di comportamento, nata da una "motivazione" che non sia unicamente economica
- La tendenza legislativa in Europa si posiziona in un atteggiamento meno coercitivo di altre norme tecniche e risulta più centrata sul cittadino, e non solo sul "consumatore"
- La nuova legislazione comunitaria sui servizi digitali è incentrata sulla trasparenza ed equità degli algoritmi, cercando di proteggere i minori e gli utenti meno attrezzati

procedere eticamente in un mondo di tecnologia che evolve freneticamente

la norma non basta, va accompagnata da una cultura che travalichi il tecnicismo e renda sin dall'inizio una visione dei possibili effetti

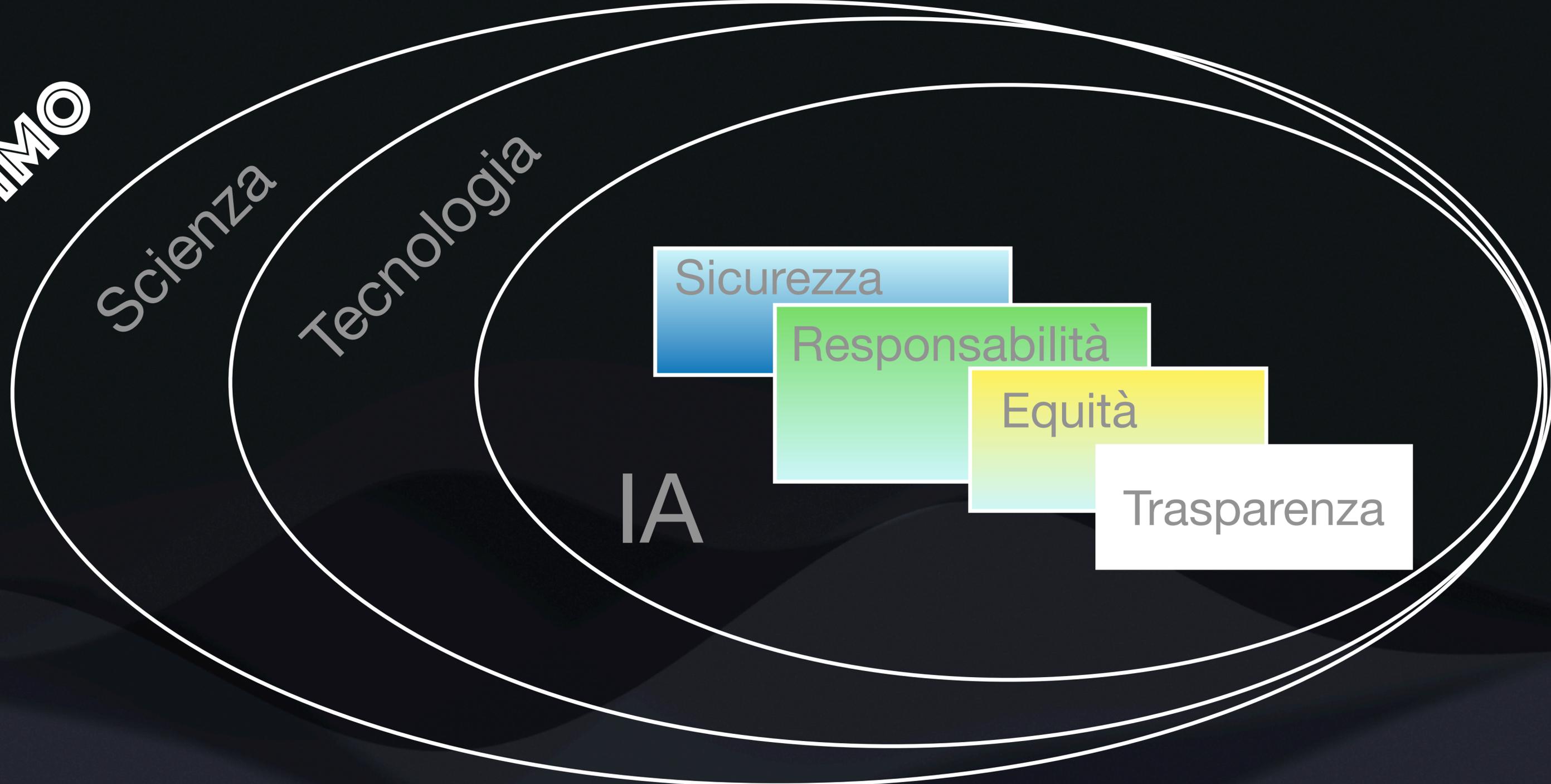


**ETICA
ALGORETICA**

Il concetto di etica evolve

Le relazioni

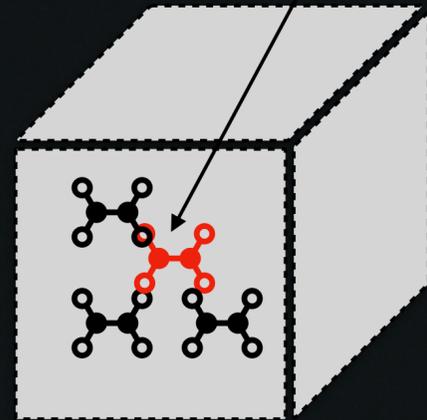
UMANESIMO



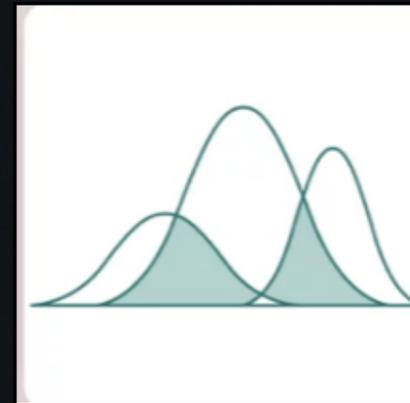
La qualità dell'Intelligenza Artificiale



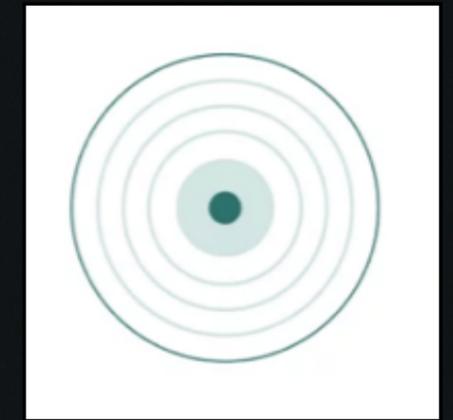
Accuratezza



Spiegabilità



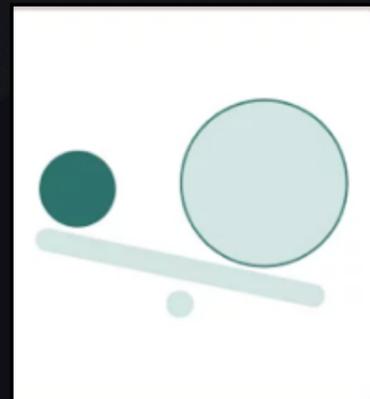
Stabilità



**Solidità
concettuale**



Affidibilità



**Equità
Riservatezza**

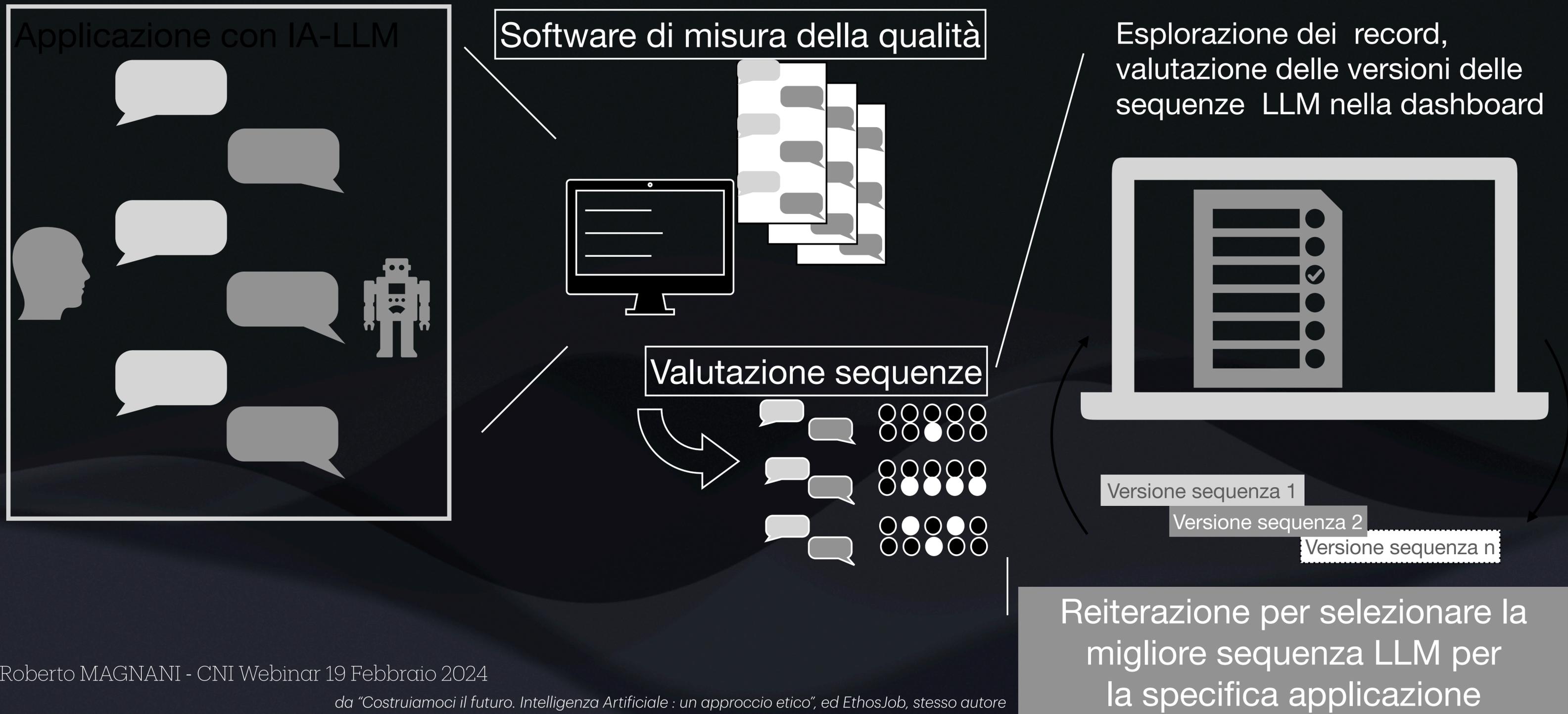


**Qualità dei
Dati**



ed altro

Esempio di spiegabilità Schema di analisi qualità di una soluzione con applicazione software dedicata



Il contorno

Potenziali effetti negativi su tutta la catena di attuazione di un progetto con IA

	Recupero di risorse	Rielaborazione	Attuazione
Individuale	Sfruttamento del lavoro Appropriazione dei dati	Sfruttamento del lavoro Rielaborazione dati non consensuale	Manipolazione dei comportamenti Danno fisico Sorveglianza scorretta
Collettivo	Sfruttamento di gruppi marginalizzati/indigeni Danno alla riservatezza di gruppo dallo sfruttamento di dati collettivi Sfruttamento di risorse collettive	Sfruttamento di gruppi marginalizzati/indigeni Danno alla riservatezza di gruppo dallo elaborazione di dati collettivi	Sorveglianza di gruppo scorretta Sfruttamento dei lavoratori occasionali Discriminazione di gruppo attraverso gli algoritmi
Sociale	Sfruttamento del lavoro istituzionalizzato Esaurimento delle risorse naturali Perdita di biodiversità Danno agli animali Emissione / Estrazioni	Sfruttamento del lavoro istituzionalizzato Contaminazione di risorse naturali Emissioni in fase di addestramento del modello Scarti chimici Utilizzo di acqua socilai/politiche Emissione / Estrazioni	Sfruttamento del lavoro istituzionalizzato Allargamento delle ineguaglianze strutturali Emissioni del modelli in uso Aumento della lacune digitale Erosione della fiducia nelle istituzioni socilai/politiche Emissione / Estrazioni

Le prestazioni “etiche”

Verifica	Misura	Azione in caso di problemi
Eguaglianza	Verificare che il sistema non abbia bias o discriminazioni nei confronti di gruppi di persone vulnerabili.	Rivedere i dati di addestramento e il modello per identificare e rimuovere i bias.
Privacy	Verificare che il sistema non violi la privacy degli utenti.	Informare gli utenti delle informazioni che vengono raccolte e su come vengono utilizzate.
Sicurezza	Verificare che il sistema sia sicuro e protetto da attacchi.	Implementare misure di sicurezza per proteggere il sistema e i dati.
Accountability	Verificare che il sistema uomo macchina sia responsabile delle sue azioni.	Sviluppare un processo per spiegare e giustificare le decisioni del sistema.
Trasparenza	Verificare che il sistema sia trasparente e spiegabile.	Fornire agli utenti informazioni sul funzionamento del sistema.
Benefici sociali	Verificare che il sistema abbia un impatto positivo sulla società.	Monitorare l'impatto del sistema sulla società e apportare modifiche se necessario.

Equità

Verifiche

<i>Tipo di casistica</i>	<i>Verifica di equità procedurale</i>	<i>Verifica di equità effettiva</i>
Accesso	<i>Tutti gli utenti hanno uguale accesso al progetto?</i>	<i>Tutti gli utenti possono utilizzare il progetto in modo efficace?</i>
Partecipazione	<i>Tutti gli utenti hanno uguali opportunità di partecipare al progetto?</i>	<i>Tutti gli utenti sentono di essere rappresentati nel progetto?</i>
Benefici	<i>Tutti gli utenti ricevono gli stessi benefici dal progetto?</i>	<i>I benefici del progetto sono distribuiti in modo equo?</i>
Effetti	<i>Il progetto ha un impatto positivo su tutti gli utenti?</i>	<i>Il progetto non ha effetti negativi su nessuno?</i>

Ovviare ai pregiudizi

Mitigazione dei pregiudizi

<i>Tecnica</i>	<i>Descrizione</i>	<i>Benefici</i>	<i>Svantaggi</i>
Pulizia dei dati	Rimozione di dati errati, mancanti o duplicati	Migliora l'accuratezza e la robustezza dell'algoritmo	Richiede tempo e risorse
Equilibratura dei dati	Riduzione del disallineamento tra le classi di dati	Migliora l'accuratezza dell'algoritmo su tutte le classi	Può ridurre la generalizzazione dell'algoritmo
Ridimensionamento dei dati	Normalizzazione dei dati su un intervallo comune	Migliora la convergenza dell'algoritmo	Può ridurre la capacità di apprendimento dell'algoritmo
Analisi del bias	Identificazione e mitigazione dei pregiudizi nei dati	Migliora l'accuratezza e la fairness dell'algoritmo	Può essere difficile e dispendioso in termini di tempo
Diversificazione dei dati	Aggiunta di dati da nuove fonti o con nuove caratteristiche	Migliora la generalizzazione dell'algoritmo	Può essere difficile e dispendioso in termini di tempo

Cosa cambia nella progettazione

Ex Ante

La progettazione ex ante è importante per garantire che i sistemi di Intelligenza artificiale siano progettati e sviluppati in modo etico.

Nella tabella alcuni temi e azioni da considerare in fase di progettazione di un sistema

Tema	Azione
Scopo e obiettivi del sistema	Definire chiaramente lo scopo e gli obiettivi del sistema, in modo da garantire che siano in linea con i valori etici di riferimento.
Utenti e "stakeholder"	Identificare gli utenti e gli stakeholder coinvolti nel sistema, in modo da comprendere le loro esigenze e aspettative.
Rischi e impatti potenziali	Valutare i rischi e gli impatti potenziali del sistema, sia positivi che negativi, in modo da mitigare quelli negativi e massimizzare quelli positivi.
Principi etici	Definire i principi etici che guideranno la progettazione e lo sviluppo del sistema.
Metodi di progettazione etica	Adottare metodi di progettazione etica, come la valutazione dell'impatto etico, la trasparenza e l'accountability.

INTRA

Progettazione etica

Tabella non esaustiva

È importante adattare questi aspetti alle specifiche esigenze del progetto e considerare anche altri aspetti etici pertinenti.

Tema	Descrizione	Azioni INTRA (Input, Norme, Trasparenza, Responsabilità, Auditabilità)
Fairness	Assicurarsi che il sistema fornisca risultati imparziali e non discriminatori.	- Implementare algoritmi di apprendimento che riducano i bias - Analizzare i dati di addestramento per identificare e correggere eventuali disparità - Monitorare continuamente l'impatto del sistema su gruppi diversi
Privacy	Proteggere la privacy delle informazioni personali e sensibili degli utenti.	- Implementare tecniche di anonimizzazione - Consentire agli utenti di gestire le proprie preferenze sulla privacy - Crittografare i dati sensibili durante la trasmissione e lo storage - Definire e rispettare rigorose politiche di accesso ai dati
Trasparenza	Rendere il funzionamento del sistema comprensibile agli utenti.	- Fornire spiegazioni chiare sulle decisioni prese dal sistema - Implementare funzionalità di interpretazione del modello - Comunicare in modo trasparente sui criteri utilizzati per addestrare il modello
Responsabilità	Definire chiaramente le responsabilità per il funzionamento del sistema.	- Assegnare responsabilità specifiche a persone o team - Creare procedure per affrontare le conseguenze negative delle decisioni del sistema - Monitorare e aggiornare regolarmente il sistema per affrontare nuove sfide e rischi
Auditabilità	Consentire la registrazione e l'analisi delle decisioni del sistema.	- Registrazione dettagliata delle interazioni del sistema - Creazione di log per le decisioni prese dal modello - Implementazione di strumenti di audit per monitorare l'aderenza alle norme etiche

Ex Post

Progettazione etica

verifica ex-post dell'etica di un sistema di IA

Verifica	Misura	Azione in caso di problemi
Eguaglianza	Verificare che il sistema non abbia bias o discriminazioni nei confronti di gruppi di persone vulnerabili.	Rivedere i dati di addestramento e il modello per identificare e rimuovere i bias.
Privacy	Verificare che il sistema non violi la privacy degli utenti.	Informare gli utenti delle informazioni che vengono raccolte e su come vengono utilizzate.
Sicurezza	Verificare che il sistema sia sicuro e protetto da attacchi.	Attuare misure di sicurezza per proteggere il sistema e i dati.
Accountability	Verificare che il sistema sia responsabile o abbia responsabile delle azioni.	Sviluppare un processo per spiegare e giustificare le decisioni del sistema.
Trasparenza	Verificare che il sistema sia trasparente e spiegabile.	Fornire agli utenti informazioni sul funzionamento del sistema.
Benefici sociali	Verificare che il sistema abbia un impatto positivo sulla società.	Monitorare l'impatto del sistema sulla società e apportare modifiche se necessario.

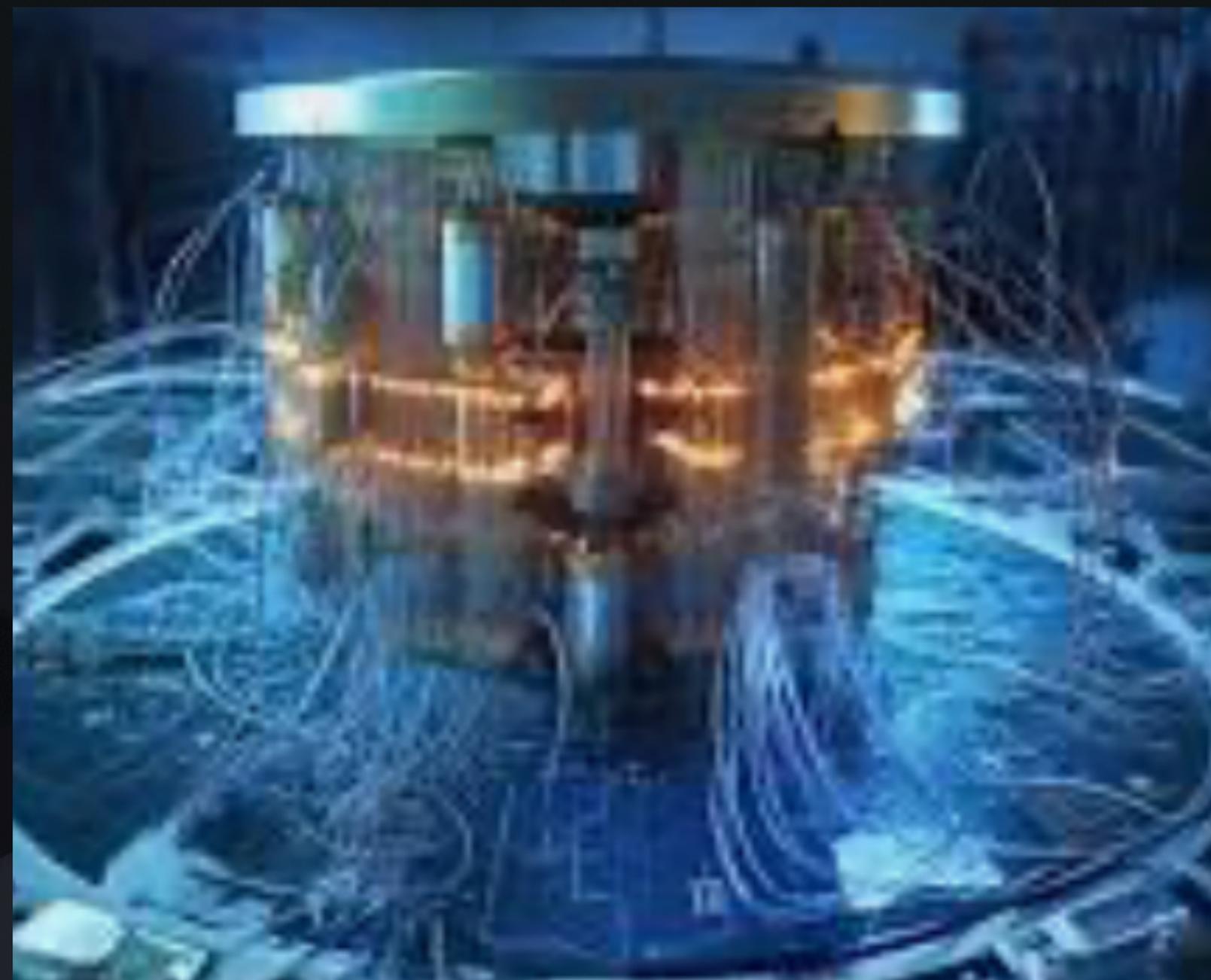
Intelligenza Artificiale e Quantum Computing: un connubio rivoluzionario

PRO

- Velocità di elaborazione esponenzialmente maggiore
- Capacità di risolvere problemi complessi in tempi ridotti
- Sviluppo di nuove applicazioni in campi come:
 - Medicina
 - Finanza
 - Materiali avanzati
 - Cybersecurity

RISCHI

- Minaccia alla crittografia esistente
- Possibilità di deepfake e disinformazione avanzate
- Problemi di sicurezza e controllo
- Disuguaglianza tecnologica e sociale



AI Act

l'AI Act

REGOLAMENTO DEL PARLAMENTO EUROPEO E DEL CONSIGLIO CHE STABILISCE REGOLE ARMONIZZATE SULL'INTELLIGENZA ARTIFICIALE

Protezione dei diritti: Mette al centro la tutela dei diritti e delle libertà individuali.

Valutazione del rischio: Regola l'IA in base al suo potenziale impatto sulla società. Più alto è il rischio, più rigorose sono le regole.

Vietati sistemi ad alto rischio: Proibisce l'uso di sistemi di IA che presentano rischi inaccettabili per la sicurezza delle persone.

5. **Governance:** Introduce un sistema di governance con poteri di applicazione a livello europeo.

Approccio U.E. basato sul rischio



Aspetti salienti di governance dell'AI Act

Nuova architettura di governance:

- Istituito un **ufficio per l'IA** all'interno della Commissione per supervisionare i modelli di IA per finalità generali.
- **Un gruppo scientifico** di esperti indipendenti fornisce consulenza all'ufficio per l'IA.
- Il comitato per l'IA rimane una **piattaforma di coordinamento e un organo consultivo** della Commissione.
- Viene istituito un forum consultivo per i portatori di interessi.

Trasparenza e protezione dei diritti fondamentali:

- È prevista una **valutazione d'impatto** sui diritti fondamentali per i sistemi di **IA ad alto rischio**.
- È prevista una **maggiore trasparenza** per l'uso dei sistemi di IA ad alto rischio.
- Gli utenti di un **sistema di riconoscimento delle emozioni devono informare le persone** fisiche quando sono esposte a tale sistema.

Sanzioni:

- **sanzioni pecuniarie** fissate in **% del fatturato annuo** globale.
- **Max proporzionati** per PMI start up
- **Modalità di reclamo** per persone fisiche giuridiche

Misure a sostegno dell'Innovazione

- **Spazi di sperimentazione normativa:**

- Creare un ambiente controllato per lo sviluppo e la convalida di sistemi di IA innovativi.
- Permettere di testare sistemi di IA innovativi in condizioni reali.

- **Test in condizioni reali:**

- Consentire di testare i sistemi di IA in condizioni reali, con specifiche condizioni e tutele.

- **Sostegno alle piccole imprese:**

- Alleggerire gli oneri amministrativi.
- Prevedere deroghe limitate e chiaramente specificate.

- Creare un quadro giuridico favorevole all'innovazione.
- Promuovere un apprendimento normativo basato su dati concreti.

I tempi di attuazione

Processo complesso

Tabella di marcia:

- 2021: Proposta del Regolamento sull'IA da parte della Commissione Europea.
- 2022-2023: Negoziations del Regolamento sull'IA tra il Parlamento europeo e il Consiglio dell'UE.
- 2024: Adozione del Regolamento sull'IA.
 - Comporta l'adozione di norme applicative normalizzate
- entro 2026: Implementazione ed applicazione del Regolamento sull'IA negli Stati membri dell'UE.

Il processo di normazione per l'AI Act

1. La Commissione europea crea una richiesta di armonizzazione.
2. Le organizzazioni europee per la standardizzazione (ESO), come CEN, CENELEC ed ETSI, sviluppano gli standard armonizzandoli
3. La Commissione europea approva gli standard.

articolo 40 Idel'AI Act europeo stabilisce che i sistemi di intelligenza artificiale (IA) ad alto rischio che sono **conformi agli standard armonizzati sono presuntivamente conformi ai requisiti essenziali dell'AI Act.**, a meno che non vi siano prove concrete che dimostrino il contrario.

I sistemi ad alto rischio

alcuni esempi

I **sistemi di intelligenza artificiale ad alto rischio** sono quelli che potrebbero avere un impatto significativo sulla sicurezza delle persone o dei loro diritti.

Sistemi di sorveglianza: Tecnologie di riconoscimento facciale o di monitoraggio che possono essere utilizzate per tracciare le persone senza il loro consenso.

Sistemi di valutazione creditizia: Algoritmi che determinano l'accesso al credito o alle risorse finanziarie.

Sistemi diagnostici medici: Algoritmi che influenzano decisioni importanti sulla salute, come diagnosi o trattamenti.

4. **Veicoli autonomi:** Tecnologie che guidano veicoli senza intervento umano.

5. **Sistemi di selezione del personale:** Algoritmi utilizzati per selezionare candidati per un lavoro o un'opportunità.

Sistemi di IA ad alto rischio nell'ingegneria civile

Controllo strutturale automatizzato: Algoritmi che valutano la stabilità e l'integrità di strutture come ponti, grattacieli o dighe. Un errore in questo sistema potrebbe mettere a rischio la vita delle persone.

Sistemi di pianificazione del cantiere: Algoritmi che ottimizzano la pianificazione delle attività di costruzione, gestendo risorse, tempi e logistica. Se mal progettati, potrebbero causare ritardi, sprechi o incidenti sul cantiere.

Riconoscimento automatico dei difetti: Algoritmi che analizzano immagini o dati per individuare difetti o anomalie nelle strutture. Ad esempio, rilevare crepe o corrosione in un ponte. Un falso negativo potrebbe avere conseguenze gravi.

Sistemi di monitoraggio delle condizioni: Tecnologie che raccolgono dati in tempo reale su edifici o infrastrutture. Se un sistema di monitoraggio non funziona correttamente, potrebbe non segnalare problemi critici.

5. **Simulazioni di carico e stress:** Algoritmi che simulano il comportamento di una struttura sotto carichi o condizioni estreme. Errori in queste simulazioni potrebbero portare a progettazioni errate o cedimenti strutturali

I rischi etici nell'uso dell'IA nell'ing. civile

Bias e Discriminazione: -Ppregiudizi presenti nei dati di addestramento, portando a decisioni discriminatorie. Ad esempio, se i dati storici mostrano discriminazione di genere o razza, l'IA potrebbe perpetuarla.

Privacy e Sorveglianza: L'uso di sensori e dati per monitorare infrastrutture può sollevare preoccupazioni sulla privacy. Come vengono gestiti e protetti i dati raccolti?

Responsabilità Legale: Chi è responsabile in caso di errori o danni causati dall'IA? La definizione di responsabilità diventa complessa quando l'IA è coinvolta.

Sicurezza Cibernetica: L'IA può essere vulnerabile agli attacchi informatici. Proteggere i sistemi IA è cruciale per garantire la sicurezza delle infrastrutture.

5. **Trasparenza e Interpretabilità:** Le decisioni prese dall'IA spesso mancano di spiegazioni chiare. Come garantire che le scelte siano comprensibili e giustificabili?

Esempio: Le sfide critiche nei veicoli autonomi connessi (CAV)

Aspetti critici da un punto di vista etico

- **Etica:** Riconoscimento, decisioni di salvataggio della vita
- **Responsabilità:** Chi è responsabile in caso di incidente?
- **Privacy e Sicurezza informatica:** Rischi di intrusione e raccolta dati
- **Giurisprudenza:** Adattamento alle regole stradali in varie giurisdizioni

Livelli di Autonomia Automobilistica

- Dal livello 0 al livello 5: Differenze e responsabilità associate

Soluzioni Potenziali

- Standard come IEEE 2846™-2022 per affrontare le sfide
- Librerie software contenenti le leggi stradali per aiutare i CAV a operare legalmente
- In fase di rilascio regolamentazione europea valida per tutti i 27 Paesi

- Livello 0** (controllo completo del conducente,
- Livello 1** (automazione minore, ad esempio controllo della velocità di crociera, intervento del conducente richiesto)
- Livello 2** (automazione parziale, ADAS, controllo dello sterzo e dell'accelerazione, intervento del conducente richiesto)
- Livello 3** (rilevamento ambientale, il veicolo può eseguire la maggior parte delle attività del conducente, intervento del conducente richiesto)
- Livello 4** (automazione estesa, intervento del conducente facoltativo)
- Livello 5** (capacità di guida completa, nessun intervento o presenza del conducente richiesto)

I CAV hanno un grande futuro ma richiedono attenzione alle questioni di sicurezza, etica, cibersicurezza e conformità.

Il Digital Service Package dell'U.E.

1. Digital Markets Act (DMA):

- **Scopo:** Il DMA mira a regolare i **grandi mercati digitali** per promuovere l'innovazione, la crescita e la competitività sia nell'UE che a livello globale.
 - Creare **parità di condizioni** per le imprese nel mercato digitale.
 - Regolare le **grandi società tecnologiche** che agiscono come "gatekeeper" (controllori dell'accesso) per servizi cruciali.
 - Contrastare gli **abusi di posizione dominante** da parte delle piattaforme online.
- Entrata in Vigore: Alcune disposizioni sono già in vigore dal **1° novembre 2022**. Altre disposizioni saranno applicate a partire dal **2 maggio 2023** e dal **25 giugno 2023**.

2. Digital Services Act (DSA):

- **Scopo:** Il DSA mira a disciplinare i **servizi digitali e le piattaforme online** nell'UE per rafforzare la tutela dei diritti degli utenti online.
 - Migliorare la sicurezza digitale e la protezione degli utenti.
 - Regolare le piattaforme di condivisione contenuti, i servizi di intermediazione, i servizi di cloud e hosting web e i motori di ricerca online.
- Entrata in Vigore: Alcune disposizioni sono già in vigore dal **16 novembre 2022** per le piattaforme online di grandi dimensioni ("VLOP") e i motori di ricerca online di grandi dimensioni ("VLOSE"). L'applicazione totale inizia dal **17 febbraio 2024**.

DATA ACT



- Il **Data Act** mira a promuovere l'innovazione e a stimolare lo sviluppo di servizi basati sui dati.
- Si concentra sulla possibilità di accedere a un **mercato di dati competitivi** e sulla protezione dei diritti di proprietà intellettuale e dei segreti commerciali.
- 2. **Accesso ai Dati: meccanismi per consentire agli enti pubblici di accedere ai dati del settore privato**, specialmente in situazioni di emergenza come terremoti o incendi.
- 3. **Equa Compensazione:** stabilisce **linee guida chiare per l'equa compensazione** quando i dati vengono messi a disposizione.
- 4. **Prevenzione degli Abusi: affronta gli squilibri contrattuali** e prevede meccanismi di risoluzione delle controversie. Protegge **i diritti di proprietà intellettuale e i segreti commerciali**.
- 5. **Controllo dei Dati:** offre ai consumatori **il controllo sui dati** condivisi sui propri dispositivi.

Nuove norme USA

Recente ordine esecutivo del Presidente USA



• L'ordine esecutivo stabilisce nuovi standard per la sicurezza e la privacy dell'AI, promuove l'equità e i diritti civili, tutela i consumatori e i lavoratori, promuove l'innovazione e la concorrenza, e



- Richiede ai produttori dei sistemi AI più potenti di condividere i risultati dei test di sicurezza con il governo degli Stati Uniti e di sviluppare standard, strumenti e test per garantire che i sistemi AI siano sicuri e affidabili.
- Protegge contro i rischi dell'uso dell'AI per ingegnerizzare materiali biologici pericolosi e sviluppa nuovi standard per lo screening della sintesi biologica.
- L'ordine esecutivo è parte della strategia globale dell'amministrazione Biden-Harris per l'innovazione responsabile.

Vista x Geo

<i>Tema</i>	<i>Europa</i>	<i>Stati Uniti</i>	<i>Cina</i>
Approccio alla regolamentazione	Regolamentazione forte e centralizzata	Regolamentazione decentralizzata e basata sul consenso ?	Regolamentazione guidata dal mercato - limitata alle IA generative
Principi etici fondamentali	Diritti umani, dignità umana, trasparenza, non discriminazione, responsabilità	Diritti umani, sicurezza nazionale, concorrenza	Sviluppo economico, sicurezza nazionale, stabilità sociale
Ambito di applicazione	Tutti gli aspetti dell'IA, inclusi sviluppo, utilizzo e impatto sociale	Utilizzo dell'IA in determinate aree critiche, come la sicurezza nazionale e la giustizia penale	Utilizzo dell'IA in settori chiave, come l'intelligenza artificiale per la creazione di armi, la sorveglianza di massa e la censura
Strumenti di regolamentazione	Regolamento, linee guida, obblighi di trasparenza, controlli di sicurezza	Regolamento, incentivi fiscali, investimenti in ricerca e sviluppo	Incentivi economici, sostegno alla ricerca e sviluppo, investimenti in infrastrutture
Autorità di regolamentazione	Commissione Europea, Agenzia Europea per la Sicurezza delle Reti e dell'Informazione (ENISA), singoli Stati membri	Dipartimento di Giustizia degli Stati Uniti, Ufficio del Direttore dell'Intelligence Nazionale (ODNI), singoli Stati	Commissione Nazionale per lo Sviluppo e la Riforma (NDRC), Ministero dell'Industria e dell'Informazione Tecnologica (MIIT)

Norme ISO relative all'uso di intelligenza artificiale

sviluppate dal Comitato Tecnico ISO/IEC JTC1/SC 42, che si occupa di intelligenza artificiale

Numero norma	Area di intervento	Punti principali	Misure
ISO/IEC 24028:2020	Principi generali	Trasparenza, spiegabilità, affidabilità, equità, sicurezza, privacy, sostenibilità	- Definizione dei requisiti e delle aspettative - Valutazione dei rischi - Misure di mitigazione - Informazione e coinvolgimento delle parti interessate
ISO/IEC 42001:2023	Sistema di gestione dell'IA	Leadership - Pianificazione - Approvvigionamento - Produzione e fornitura - Misurazione, analisi e miglioramento	- Definizione della politica e degli obiettivi - Rischio e opportunità - Sviluppo e gestione dei processi - Misurazione, monitoraggio e analisi - Miglioramento continuo
ISO/IEC 20546:2019	Big data	Definizione e vocabolario	- Definizione di termini e concetti - Schema di classificazione - Principi di progettazione e sviluppo
ISO/IEC TR 20547-2:2018	Architettura di riferimento per i big data	Usi e requisiti derivati	- Analisi dei casi d'uso - Definizione dei requisiti - Architettura di riferimento
ISO/IEC TR 20547-3:2020	Architettura di riferimento per i big data	Parte 3: Architettura di riferimento	- Principi di progettazione e sviluppo - Struttura dell'architettura - Elementi dell'architettura
ISO/IEC TR 20547-5:2018	Roadmap delle norme	Roadmap delle norme sull'intelligenza artificiale	- Sviluppo delle norme - Priorità - Collaborazione internazionale

La norma ISO/IEC 42001 per la gestione dell'Intelligenza Artificiale

Punti

- Norma internazionale per la gestione responsabile dell'IA.
- Applicabile a tutti i tipi di aziende e settori.
- Promuove lo sviluppo e l'utilizzo di sistemi di IA affidabili, trasparenti e responsabili.
- Si basa sul modello HLS (High Level Structure) per l'integrazione con altri sistemi di gestione.

Obiettivi

- Promuovere lo sviluppo e l'utilizzo di sistemi di IA affidabili, trasparenti e responsabili.
- Evidenziare i principi e i valori etici nell'impiego dei sistemi di IA.
- Aiutare le organizzazioni a identificare e mitigare i rischi legati all'implementazione dell'IA.
- Incoraggiare le organizzazioni a dare priorità al benessere umano, alla sicurezza e all'esperienza dell'utente.
- Assistere le organizzazioni nel rispetto delle leggi e dei regolamenti in materia di protezione dei dati.

L'approccio IEEE

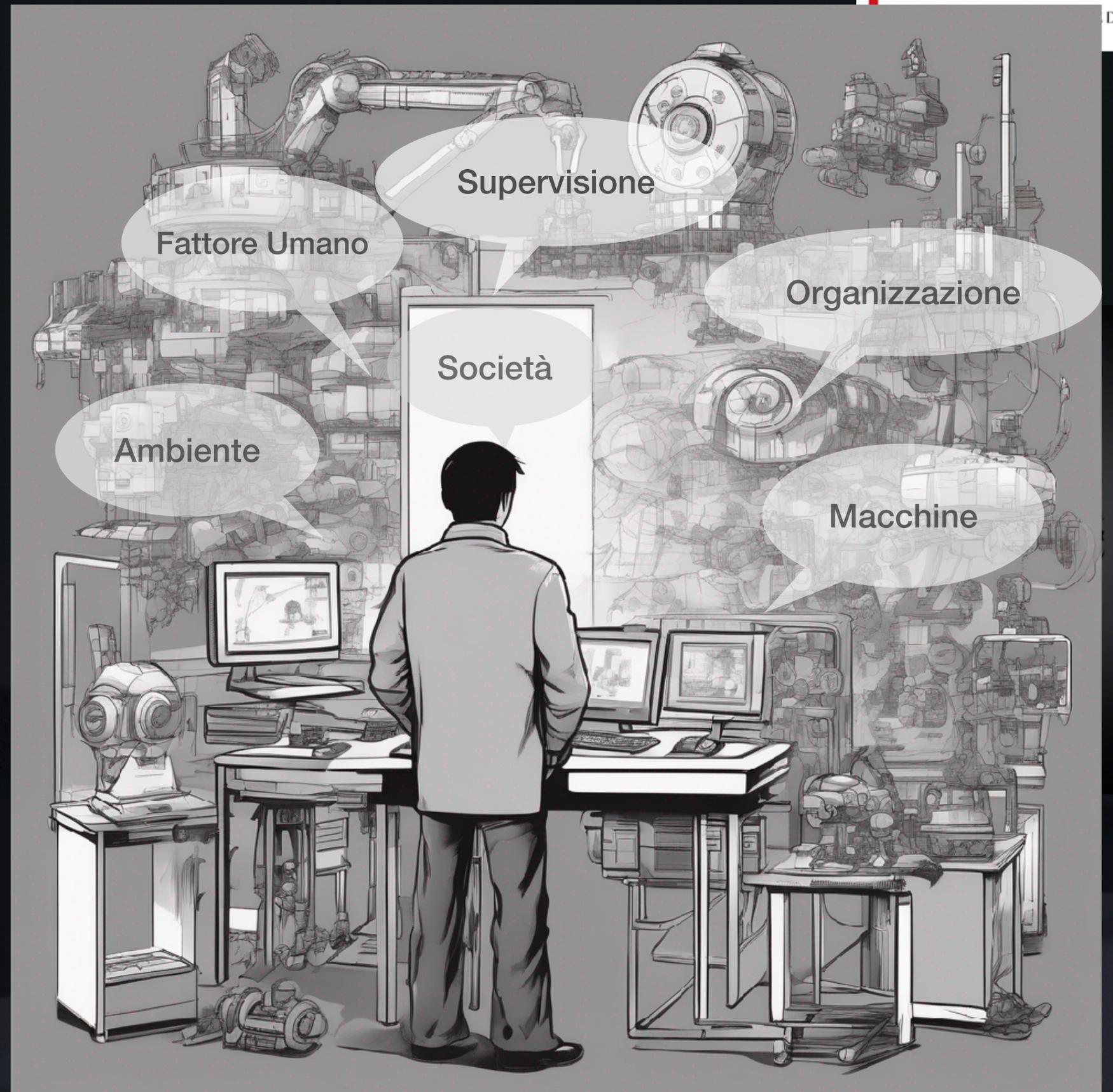
In corso il rilascio di una certificazione specifica d'industria

Titolo	Anno di Pubblicazione	Descrizione
IEEE P7000 - Standard per i Sistemi Autonomi Affidabili	2021	Fornisce una serie di requisiti e raccomandazioni per lo sviluppo e la gestione di sistemi autonomi affidabili.
IEEE P7001 - Standard per la Trasparenza dell'Intelligenza Artificiale	2022	Definisce i requisiti per la trasparenza dei sistemi di intelligenza artificiale, in modo da permettere agli utenti di comprendere come funzionano e quali sono le loro potenziali
IEEE P7002 - Standard per l'Etica dell'Intelligenza Artificiale	In corso di sviluppo	Stabilisce i principi etici per lo sviluppo e l'utilizzo dell'intelligenza artificiale, garantendo che sia utilizzata in modo responsabile e non discriminatorio.
IEEE P7003 - Standard per la Sicurezza dell'Intelligenza Artificiale	In corso di sviluppo	Definisce i requisiti per la sicurezza dei sistemi di intelligenza artificiale, in modo da ridurre il rischio di incidenti e danni.
IEEE P7004 - Standard per la Robustezza dell'Intelligenza Artificiale	In corso di sviluppo	Stabilisce i requisiti per la robustezza dei sistemi di intelligenza artificiale, in modo da garantirne il corretto funzionamento in condizioni diverse.
IEEE P7005 - Standard per la Privacy dell'Intelligenza Artificiale	In corso di sviluppo	Definisce i requisiti per la protezione della privacy dei dati personali utilizzati nei sistemi di intelligenza artificiale.

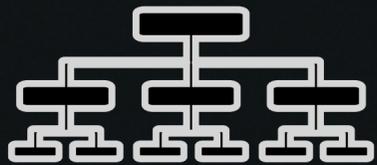
Digital Ethics Officer

Un “executive” con un ruolo alle dirette dipendenze dell’amministratore delegato

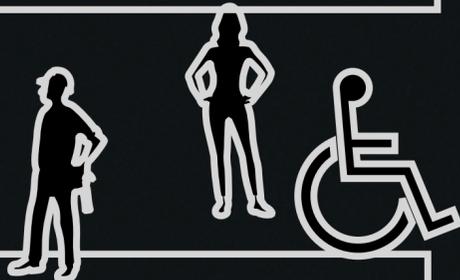
Le aree di coinvolgimento



Le attività



Sensibilizza e consiglia le differenti entità organizzative



Sostiene la trasversalità dei riferimenti e delle persone coinvolte



Funge da riferimento



Organizza e guida la gestione operativa dell'etica



Individua esposizioni e rischi etici

Competenza In IA



Crea e condivide una strategia etica



Sviluppa sistemi di misura e controllo dell'etica



Elimina o minimizza i rischi etici



Integra le funzioni e migliora le sinergie



Organizza momenti interni di analisi e riflessione per valutare le reazioni interne ed esterne



G7: I 4 pilastri del protocollo di Hiroshima

passo importante per lo sviluppo e l'utilizzo responsabile dell'IA

Il successo dipender dalla collaborazione internazionale e dall'impegno di tutti gli attori coinvolti nello sviluppo e nell'utilizzo

Pilastro	Obiettivo
Analisi dei rischi prioritari, delle sfide e delle opportunità dell'IA generativa	Identificare e valutare i potenziali rischi e benefici dell'IA generativa.
Principi guida internazionali del processo di Hiroshima per tutti gli attori dell'IA nell'ecosistema dell'IA	Stabilire un insieme di principi guida universali per lo sviluppo e l'utilizzo responsabile dell'IA.
Codice di condotta internazionale di Hiroshima per le organizzazioni che sviluppino sistemi avanzati di IA	Fornire un codice di condotta volontario per le organizzazioni che sviluppino sistemi di IA avanzati.
Cooperazione basata su progetti a sostegno dello sviluppo di strumenti di IA responsabili e migliori pratiche	Promuovere la collaborazione internazionale per lo sviluppo di strumenti e migliori pratiche per l'IA responsabile.

Conclusioni

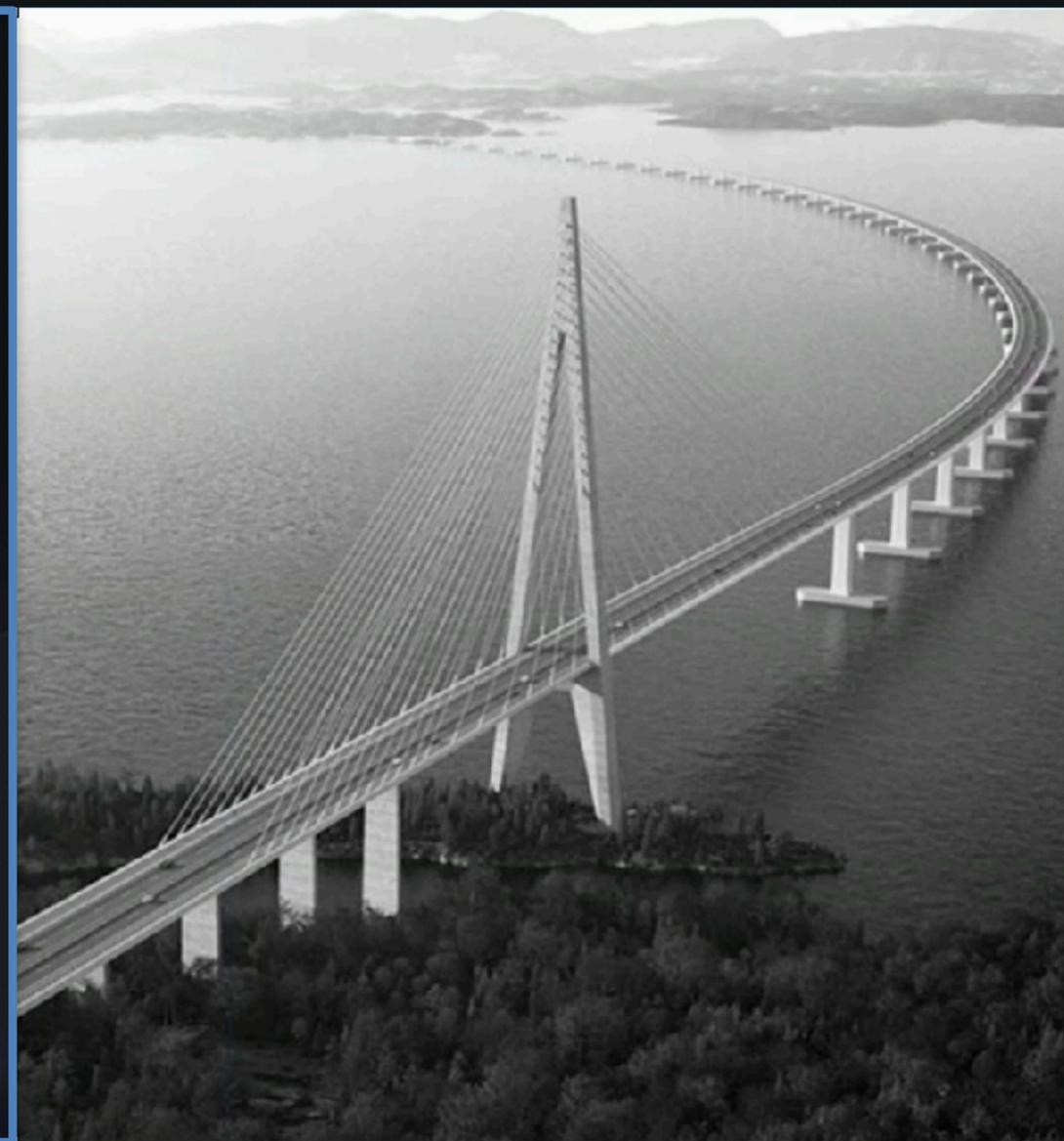
Siamo all'inizio di un viaggio, con tutte le opportunità ma anche i rischi tipici di un viaggio emozionante

Non è troppo presto per considerare l'uso nel breve termine ma anche gli effetti complessivi a lungo termine dell'IA nel campo dell'ingegneria, specialmente in aree strategiche

Diverse organizzazioni iniziano a nominare specialisti per promuovere gli "approcci etici" tra le parti interessate e i pilastri interni per garantire uno sviluppo sicuro, giusto e privo di rischi legali

Necessità di influenzare gli organismi nelle geografie per armonizzare approcci, attualmente molto diversi

Quello europeo sembra il più avanzato in base alla valutazione del rischio



GRAZIE

Struttura del libro



Intelligenza artificiale
tra scienza e
tecnologia

Intelligenza Artificiale
Etica

Per una buona
Intelligenza Artificiale

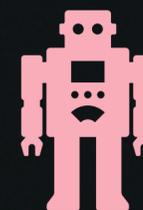
Apprendimento o combinazioni?
Dove si usa l'intelligenza artificiale?

Giustizia tecnologica

Il superamento del "tech solutionism"

Costruire il futuro della tecnologia

L'utilità delle norme



Intelligenza artificiale responsabile

Il concetto di "Good Artificial Intelligence", ovvero "IA come strumento per il bene"

La proposta di Regolamento UE: trasparenza ed equità

Il ruolo del Digital Ethics Officer
Per un nuovo umanesimo tecnologico

Conclusioni future: nelle mani di una umanità consapevole

Chances e sfide etiche oltre le norme e gli algoritmi

Equità | Fairness

Trasparenza | Transparency

Democrazia | Democracy

Opacità da interdipendenza tecno-sociale Bias: i pregiudizi

Questioni di reputazione e di conformità alle norme

Come rilevare i pregiudizi nei processi con intelligenza artificiale

Tecnologie per la prevenzione dei bias l'effetto dell'intelligenza artificiale sulle altre scienze e sulla società

Criticità etiche connesse all'intelligenza artificiale Approccio ex ante / intra/ex post Interpretabilità nell'elaborazione del linguaggio naturale

Benefici di una valutazione d'impatto etico

Computer quantistico e crittografia

Prefazione "Il prezzo della consapevolezza" Di Giovanni Caprara

L'autore



Roberto Magnani

Ingegnere elettronico, sviluppa la sua carriera nel settore IT nei laboratori di multinazionali in Italia, Francia, USA, Svizzera, Spagna e Irlanda. Nell'ultimo decennio attua per l'Europa i servizi Web per il Public Cloud di una grande multinazionale in un Campus tecnologico di Dublino per poi assumere la responsabilità di progetti digitali per Healthcare /Life Science in EMEA sempre via Web, con l'utilizzo di intelligenza artificiale. Dal 2022 è consulente indipendente; consigliere di AEIT- Associazione italiana elettronica elettrotecnica informatica e telecomunicazioni, focalizzandosi sugli aspetti etici e normativi dell'intelligenza artificiale e l'introduzione del Quantum computing nell'industria. È autore di articoli e interventi in Italia e all'estero sugli stessi argomenti e recentemente di un'istant book "Intelligenza artificiale per le professioni" edito da EBS e di "Creiamoci il futuro. Intelligenza artificiale, un approccio etico" edito da EthosJob

<https://www.linkedin.com/in/robertomagnani/>