CENTRO STUDI AVANZATI PER L'INTELLIGENZA ARTIFICIALE

CSAIA.
CENTRO DI STUDI AVANZATI
PER L'INTELLIGENZA ARTIFICIALE

Manifesto CSAIA

AI al servizio del Rule of Law

Introduzione

Considerando la natura, gli scopi e gli obiettivi della CSAIA, già espressi nel suo statuto,

Considerando che l'IA abbraccia sia le sfide che le opportunità a livello globale,

Considerando il momento storico critico per le relazioni internazionali multilaterali,

Considerando la necessità di affrontare le sfide e le opportunità fornite dall'IA attraverso una cooperazione europea e internazionale che vada oltre le frontiere fisiche e intellettuali con una visione interdisciplinare e trasversale,

Considerando che il dialogo scientifico è uno strumento essenziale per garantire scelte informate e lungimiranti,

Considerando che la CSAIA è un forum di discussione scientifica internazionale per identificare e implementare le migliori pratiche e le migliori soluzioni nell'IA,

Convinti che lo sviluppo di sistemi di IA debba rimanere al servizio dell'umanità e dello Stato di Diritto e contribuire al mantenimento della Pace,

Nel rispetto dei più alti standard europei e internazionali e delle normative applicabili,

Rivolgendosi alle Istituzioni Nazionali, Europee e Internazionali per avviare un dialogo costruttivo e operativo basato su principi comuni

Con l'obiettivo di facilitare la risoluzione delle difficoltà attraverso la trattazione scientifica interdisciplinare, trasversale e internazionale delle complesse tematiche legate agli sviluppi e alle applicazioni dell'IA,

Con l'obiettivo di rafforzare la consapevolezza di tutti gli attori,

CSAIA propone la seguente dichiarazione di intenti per l'adesione delle istituzioni che desiderano partecipare a questo dialogo e approccio costruttivo.

Punti di programma

1. Progettazione, privacy e sicurezza by design, diritti di proprietà intellettuale:

- la progettazione di un sistema di IA può seguire, nel rispetto delle regole di privacy e sicurezza riconosciute a livello internazionale, strumenti e best practice riconosciuti a livello globale utilizzando una metodologia strutturata denominata pipeline, che prevede le fasi di raccolta dei dati, elaborazione dei dati, affinamento dei dati, modellazione, valutazione del modello e pubblicazione del modello secondo gli standard della metodologia CRISP-ML (Q). Per ogni caso d'uso individuato e sviluppato si può prevedere un'architettura di tipo pipeline, garantendo l'utilizzo delle tecnologie più adatte (IA generativa, IA simbolica e informatica tradizionale, o un ibrido delle tre).

- per quanto riguarda gli aspetti di copyright, un modello di IA sposta il valore dell'informazione dai dati al modello in grado di estrarre elementi informativi tangibili dai dati, di interpretarli, elaborarli, sintetizzarli, prevederli e costruire nuovi concetti che diventano di proprietà dell'istituzione che ha addestrato il modello e restano sotto la sua responsabilità

- IA per la Sicurezza. Fatto salvo l'obiettivo di tutelare la Sicurezza Fisica, Logica e di Rete del paese, l'IA può davvero costituire un valido supporto per chi si occupa di proteggere e garantire l "Always on" delle infrastrutture critiche Industriali, di quelle della Pubblica Amministrazione, della Sanità, delle Telco e dell'Informatica, nonché quelle legate alla sicurezza Nazionale e allo Spazio. L'incremento di attori malevoli, di gruppi sponsorizzati da stati avversi, di terrorismo misto a criminalità ci impongono di guardare all' IA come una possibile e realistica risorsa tecnologica di difesa e contrasto anche in presenza dell'utilizzo di una IA malevola utilizzata per fomentare uno scontro tecnologico ibrido. Scontro Ibrido che si manifesta attraverso la propagazione di fake news, deep fake, falsi profili social, Honey pots, fake blogs, ramsomware, DDOS, trolls e influenza sul pensiero dominante di un paese. Ecco perché CSAIA ha inteso implementare anche questa linea di speculazione e di studio al servizio della Comunità.

2. Human in the loop and Ethical Compass:

- la progettazione di sistemi di IA deve sempre ispirarsi al concetto di coabitazione della tecnologia con le capacità umane, con l'obiettivo di estendere le sue capacità, ma mai di sostituirla

- la messa a punto dei modelli generativi di IA deve sempre avvenire sotto la stretta supervisione del coinvolgimento umano e non deve mai essere consentito di prendere automaticamente decisioni che potrebbero avere conseguenze potenzialmente gravi e impattare negativamente sull'individuo

- il sistema di IA deve sempre essere progettato per essere modificabile attraverso il coinvolgimento umano, al fine di rispettare il diritto all'oblio (alias diritto all'oblio)

- i sistemi basati sull'IA devono rispettare i principi etici riconosciuti a livello globale di trasparenza, privacy, equità, spiegabilità, sicurezza, robustezza, responsabilità e controllo umano

3. Governance e segregazione:

- sviluppo del modello, la preparazione dei dati, la gestione e il deployment e il monitoraggio delle prestazioni sono attività che una piattaforma di governance deve garantire attraverso operazioni di machine learning e strumenti di governance per le fasi di sviluppo e deployment

- in caso di utilizzo esclusivo di qualsiasi modello di IA (inclusa l'IA generativa), l'adattamento del modello al contesto dovrebbe essere operato con l'apprendimento automatico utilizzando dati di input leciti e imparziali, evitando contaminazioni esterne e garantendo il coinvolgimento di dati sintetici o altri concetti di ingegneria della privacy al fine di ridurre al minimo il fenomeno delle allucinazioni, dell'overfitting e dei bias cognitivi

- è necessaria una collaborazione internazionale per iniziative volte a promuovere la collaborazione tra studiosi di diritto, ricercatori di IA, policy maker e professionisti per creare un quadro completo per affrontare le questioni legali dell'IA generativa

4. Machine Unlearning (MU):

- il sistema di IA dovrebbe essere sempre progettato per essere modificabile attraverso il coinvolgimento umano, al fine di rispettare il diritto all'oblio e altri diritti individuali rilevanti:

- la conformità alle legislazioni sulla privacy come il GDPR dell'UE (che include anche il diritto all'oblio) consentirebbe di rimuovere i dati personali dai modelli addestrati, salvaguardando la privacy delle persone

- una maggiore sicurezza contro gli attacchi di "avvelenamento dei dati", rimuovendo i cosiddetti dati tossici volti a manipolare il comportamento di un modello

- una migliore adattabilità dei modelli per rimanere pertinenti anche quando la distribuzione dei dati cambia nel tempo.

- conformità normativa alle leggi e ai regolamenti in evoluzione

5. Conformità normativa:

I sistemi basati sull'IA devono essere sicuri e conformi ai diritti e ai valori fondamentali esistenti, in linea con la privacy e con altre normative pertinenti.

Nell'UE, il GDPR definisce i requisiti per l'uso legittimo dei dati personali da parte dei modelli di IA. Gli sviluppatori di IA devono garantire, tramite la tecnologia incorporata, che i permessi e le autorizzazioni richiesti, come il consenso esplicito, siano forniti volontariamente e che siano specifici, informati e non ambigui. In alcuni casi, l'IA può trattare i dati personali sulla base di altre basi legali, come il "legittimo interesse".

6. Contaminazione:

È necessario promuovere la conoscenza dei progetti sviluppati e la loro riproposizione nei casi previsti, al fine di spiegare come la condivisione delle conoscenze possa avvenire nella pratica e quali benefici possa apportare.

7. Contesto pedagogico e IA: prospettive emendative ed emancipative nell'educazione

Nel contesto socio-educativo contemporaneo, l'intelligenza artificiale (IA) offre un'opportunità rilevante per promuovere un apprendimento critico, personalizzato e sostenibile. La pedagogia emendativa valorizza l'errore come opportunità di apprendimento, mentre la pedagogia emancipativa promuove l'autonomia e la responsabilizzazione. L'IA, vista come strumento al servizio dell'educazione, può aiutare a decostruire e ricostruire la conoscenza, promuovendo il pensiero critico nelle nuove generazioni. La pedagogia garantisce un approccio etico all'innovazione tecnologica, alimentando la fiducia nel potenziale educativo dell'IA e superando la "paura dell'educazione". In questa prospettiva, la pedagogia si pone come garante di un approccio etico all'innovazione tecnologica. Essa alimenta la fiducia nel potenziale educativo dell'IA e avverte gli educatori - e, con loro, i futuri giuristi - di superare la "paura di educare", riconoscendo l'educazione come estensione privilegiata dell'interrogazione critica. Una tale visione educativa può rendere possibile comprendere e utilizzare la tecnologia in modo coerente con i principi fondanti dello Stato di diritto: equità, giustizia e legalità.

8. Agentic IA:

L'autonomia dell'Agentic IA pone sfide sia etiche che legali: richiede una solida governance, meccanismi di controllo umano e un costante allineamento con le normative attuali e in evoluzione. La valutazione dei rischi etici (autonomia vs. controllo umano, parzialità e discriminazione, trasparenza e spiegabilità, privacy e sorveglianza, responsabilità) e legali (responsabilità, conformità normativa, protezione della proprietà intellettuale, criminalità informatica) deve essere costantemente presente.

9. L'intelligenza artificiale nella finanza:

L'intelligenza artificiale (AI) rappresenta un'opportunità strategica per promuovere una finanza più inclusiva, consapevole e strettamente legata agli impatti sistemici a livello socio-economico e ambientale. L'integrazione dell'IA e, più in generale, delle tecnologie digitali nei processi finanziari consente di sviluppare servizi e prodotti orientati all'empowerment degli attori economici, rafforzando la diffusione della conoscenza finanziaria e l'accesso anche agli strumenti più avanzati di gestione del capitale, favorendo l'inclusività e contribuendo alla sostenibilità degli ecosistemi

La conoscenza delle dinamiche di mercato e l'adozione di pratiche finanziarie consapevoli sono leve fondamentali per l'empowerment dei cittadini e la costruzione di una società economicamente più matura e democratica. In questo contesto, il settore fintech emerge come motore di crescita in grado di aumentare l'efficienza, migliorare la sicurezza nella gestione del rischio e promuovere una maggiore trasparenza dei processi.

Tuttavia, lo sviluppo dell'IA nella finanza richiede un approccio normativo dinamico e co-evolutivo, concepito secondo una logica a "doppia elica", in cui innovazione tecnologica e regolamentazione procedono in sinergia. Strumenti come le "sandbox regolamentari" forniscono ambienti controllati in cui sperimentare soluzioni innovative nel rispetto dei principi etici, della trasparenza decisionale e della spiegabilità degli algoritmi. La finanza sostenuta dall'IA, così strutturata, può contribuire in modo sostanziale alla comprensione critica e all'uso responsabile delle tecnologie emergenti, in linea con i principi fondanti dell'economia civica, della coesione sociale e della tutela ambientale.

Contributor:

- Marco Braccioli - Consigliere Fondazione ICSA, Membro del Consiglio Direttivo Afcea International-Rome Chapter, Membro del Comitato Tecnico CESMA Centro Studi Militari Aeronautici, Roma

- Ulrika Dellrud – Global chief privacy and data ethics officer, Board Member and NED, Certification Director ISACA Belgium, AI Governance Advisory Board Member, IAPP, Advisor to European Data Protection Board, Member of the NY Bar

- Agostino Ghiglia - GPDP College Member, Data Protection Authority, Rome

- Simon Greenman, Chief AI Officer Best Practice AI, past World Economic Forum Global AI Council, Harvard Business School Alumni Angels UK, London

- Konstantinos Kenanidis - Secretary General Office of the Orthodox Church in EU, Bruxelles

- Suany Mazzitelli - Éditeur LexisNexis, membre Observatoire de légistique, Paris

- Francesca Medda – Director at Institute of Finance and Technology, UCL, London

- Pier Paolo Maria Menchetti - EU Commission/EMA Chairman Thematic Panel, co-Director International School on AI Technology and Law - Ettore Majorana Foundation Erice, President CSAIA

- Saverio Occhipinti - General Counsel Club Deal Investors University LUISS Guido Carli ("LUISS Alumni 4 Growth"), LUISS University, Rome

- Anna Pouliou - Maastricht University, Faculty of Law, European Center of Privacy and Cybersecurity; Data Protection Commission, CERN, Genève

- Daniele Pulcini - University for Peace - UN Mandated, Rome

- Domenico Franco Sivilli - General Director for Information Resources and Statistics, Council of State, Rome

- Cinzia Turli - Pedagogy. San Raffaele Telematic University, Rome; G. D'Annunzio University, Chieti Pescara

(Ultimo aggiornamento: 18/09/2025)

Associazione CSAIA
ETS rep. n. 135104

Viale Città d'Europa, 681
00144 ROMA.
Tel. +39 335 8005403
www.centrostudicsaia.it
info@centrostudicsaia.it