T

Adriano Fabris e Carlo Casonato

Quali criteri etici e giuridici per l'Intelligenza Artificiale? Il lavoro del TP 1 del progetto FAIR

1. *Presentazione generale del progetto FAIR*

La Fondazione FAIR – Future Artificial Intelligence Research è un ente senza fini di lucro istituito con l'obiettivo di attuare gli interventi previsti dal Piano Nazionale di Ripresa e Resilienza (PNRR) e da eventuali successivi programmi di finanziamento dedicati al settore dell'Intelligenza Artificiale. Essa rappresenta un'infrastruttura strategica nazionale per la promozione della ricerca scientifica e dell'innovazione tecnologica, il cui primo obiettivo è quello di coordinare il Partenariato Esteso FAIR, concepito per rafforzare la cooperazione tra mondo accademico, enti di ricerca e sistema produttivo.

L'organizzazione del partenariato si fonda sul modello Hub & Spoke, nel quale la Fondazione agisce come hub, ovvero soggetto attuatore e referente unico del progetto, responsabile della gestione, del coordinamento e del monitoraggio complessivo delle attività. Gli spoke, invece, rappresentano i soggetti esecutori, incaricati di sviluppare e implementare le linee di ricerca in collaborazione con lo hub, in un'ottica di integrazione e sinergia tra competenze differenti.

La Fondazione riunisce alcune tra le più autorevoli istituzioni di ricerca e università italiane, affiancate da importanti realtà industriali. Ne fanno parte quattro enti di ricerca pubblici – il Consiglio Nazionale delle Ricerche, la Fondazione Bruno Kessler, l'Istituto Nazionale di Fisica Nucleare e l'Istituto Italiano di Tecnologia – e tredici università italiane di eccellenza. A questi si aggiungono cinque aziende private – Bracco, Expert.ai, Intesa Sanpaolo, Leonardo e Lutech – e il Consorzio Interuniversitario Nazionale per l'Informatica.

 DOI: 10.4454/aphm7q94

In coerenza con la Strategia Nazionale Italiana per l'Intelligenza Artificiale, la Fondazione FAIR mira a promuovere una ricerca di frontiera capace di coniugare l'eccellenza scientifica con l'innovazione tecnologica e con un approccio etico e sostenibile. La sua azione intende superare la frammentazione della ricerca italiana nel campo dell'IA, favorendo la creazione di una rete unitaria e interdisciplinare che valorizzi le competenze diffuse sul territorio nazionale e ne incrementi la visibilità internazionale. Allo stesso tempo, la Fondazione si propone di sviluppare un'intelligenza artificiale antropocentrica, affidabile e sostenibile, promuovendo la formazione e l'attrazione di talenti e contribuendo alla costruzione di un ecosistema di ricerca e innovazione capace di garantire continuità e sostenibilità nel lungo periodo.

Il Progetto FAIR si fonda su una visione olistica e multidisciplinare dell'intelligenza artificiale, intesa non soltanto come tecnologia, ma come campo di conoscenza destinato a incidere profondamente sulla società. Il progetto affronta in modo integrato le dimensioni teoriche, modellistiche e ingegneristiche dell'IA, includendo anche la riflessione sulle implicazioni etiche, giuridiche e ambientali connesse al suo sviluppo. L'obiettivo è la creazione di sistemi intelligenti capaci di interagire e collaborare con gli esseri umani, di adattarsi a contesti in continuo mutamento, di operare in modo consapevole dei propri limiti e di rispettare i principi di sicurezza, trasparenza e fiducia.

La governance del progetto si articola in dieci spoke tematici, coordinati da otto università statali e due istituzioni di ricerca, ai quali partecipano ulteriori istituzioni affiliate, tra cui università, enti vigilati dal Ministero dell'Università e della Ricerca e soggetti privati. Attraverso questa struttura, la Fondazione FAIR si configura come un modello innovativo di collaborazione tra ricerca pubblica e industria, volto a consolidare il ruolo dell'Italia nel panorama scientifico internazionale e a promuovere uno sviluppo tecnologico etico, sostenibile e orientato al benessere collettivo.

Le aree tematiche affrontate dagli spoke di FAIR hanno implicato diversi livelli di possibile interazione, in quanto il progetto aveva come obbiettivo quello di risolvere problemi complessi che attraversano ambiti differenti. Per conseguire questo obiettivo, è stato necessario condividere approcci, materiali e informazioni tra i vari team di ricerca. La cooperazione tra le aree tematiche all'interno di FAIR è stata pertanto garantita dai Progetti Trasversali (TP), gruppi di lavoro dinamici che hanno riunito ricercatori impegnati su sfide comuni e che hanno operato trasversalmente ai singoli *spoke*, o gruppi di ricerca. I TP hanno favorito sinergie tra i diversi ambiti di ricerca e gettato le basi per ulteriori collaborazioni, coordinando attività e conoscenze condivise.

2. *IL TP1*

In questo contesto, il TP1 – coordinato a livello nazionale da Carlo Casonato, Adriano Fabris e Giovanni Sartor – si era proposto di sviluppare metodologie di progettazione e contenuti volti alla realizzazione di sistemi di intelligenza artificiale (IA) affidabili, lungo le diverse dimensioni della «trustworthiness» così come delineate dai documenti europei sull'intelligenza artificiale (IA) e dall'*AI Act*. L'obiettivo principale è stato quello di verificare la conformità dei sistemi AI ai valori etici e giuridici fondamentali – quali privacy, equità, spiegabilità e controllo umano – al fine di accrescerne l'affidabilità e promuovere una maggiore fiducia nell'adozione di tali tecnologie.

Le entità computazionali pervasive, autonome e intelligenti costituiscono ormai parte integrante dell'ecologia sociale contemporanea. Attualmente, la maggior parte dei dispositivi connessi è abilitato all'IA, fenomeno che sottolinea l'urgenza di sviluppare sistemi in grado di rispettare principi etici e giuridici coerenti con una visione *human-centered*. La conformità a tali principi, lungi dall'essere un vincolo accessorio, diviene quindi un elemento fondativo dell'IA stessa, contribuendo al rafforzamento dei valori etici e giuridici nella società.

Da questa prospettiva, il rapporto tra IA e principi etico-legali rappresenta una questione trasversale che coinvolge tutti gli *spoke* del progetto. Ciascun *work package*, seppur in misura differente, è stato chiamato a confrontarsi con problematiche giuridiche ed etiche, rendendo necessaria un'attenta coordinazione tra i diversi *work packages*. A tal fine, sono stati previsti negli anni di svolgimento del progetto numerosi workshop tematici di verifica, oltre a quello conclusivo che ora presentiamo. Essi sono stati finalizzati ad approfondire tematiche quali: l'IA affidabile e *human-centered* nei diversi contesti progettuali (by-design, in-design, per i progettisti); lo sviluppo di modelli etici per IA affidabile in scenari incerti e multi-agente; l'analisi delle implicazioni etiche degli algoritmi in settori e scenari paradigmatici, tra cui IA resiliente, adattiva, simbiotica ed *edge-exascale*; i fondamenti etici per l'autonomia regolabile; i metodi per diritto ed etica computabili, mediante tecnologie innovative per la governance giuridica ed etica dell'IA; e infine la realizzabilità dell'*AI Act*, con la verifica delle attività progettuali rispetto ai principi della proposta di regolamento europeo. Tale approccio ha voluto garantire che la progettazione e l'implementazione dei sistemi di IA fossero coerenti con un quadro etico e normativo solido, promuovendo al contempo un ecosistema tecnologico affidabile, responsabile e

centrato sull'essere umano, valorizzando la cooperazione trasversale tra gli *spoke* e l'efficacia del Progetto Trasversale.

Più precisamente, i diversi *spoke* che hanno cooperato nel TP1 hanno indagato le dimensioni giuridiche, etiche e tecniche dell'intelligenza artificiale (IA) in una vasta gamma di ambiti applicativi. Pur mantenendo ciascuno una prospettiva disciplinare e settoriale ben determinata – dalla sanità alla giustizia, dal mercato del lavoro al monitoraggio ambientale – sono emerse alcune linee di ricerca comuni, segno di una crescente convergenza attorno alla necessità di sviluppare sistemi di IA conformi alla normativa vigente, eticamente solidi e tecnicamente affidabili. In particolare, i partecipanti al TP1, nel periodo della ricerca, si sono concentrati su quattro principali direttrici condivise:

1. *L'AI Act: interpretazione, applicazione e criticità*

L'analisi e l'applicazione del *Regolamento Europeo sull'Intelligenza Artificiale* (*AI Act*), all'interno del complesso ecosistema normativo dedicato alla IA, ha rappresentato uno dei temi centrali delle attività di ricerca del TP1. I ricercatori hanno approfondito il testo da prospettive giuridiche, etiche e tecniche, esaminandone vantaggi, limiti e implicazioni pratiche. Lo Spoke 1 si è occupato di analizzare il *risk-based approach* adottato dal regolamento e, connesso a ciò, il ruolo della nozione di «vulnerabilità» in correlazione con esso. Particolare attenzione è stata dedicata a questioni cruciali quali i *bias* e la trasparenza (Spoke 2), le proibizioni previste dall'articolo 5 (Spoke 3) e i meccanismi di governance, come la *coregolamentazione*, gli standard armonizzati e i *sandbox* per l'IA (Spoke 8).

La ricerca ha inoltre considerato diversi contesti di applicazione, come la Pubblica Amministrazione (Spoke 3), il settore sanitario (Spoke 2), i sistemi giudiziari e di *law enforcement*, nonché l'interazione dell'AI Act con altri strumenti regolatori europei, in particolare il GDPR e il DMA (Spoke 8). L'obiettivo comune è stato quello di valutare come l'*AI Act* possa essere correttamente interpretato e efficacemente implementato, garantendo un equilibrio tra innovazione e tutela dei diritti fondamentali.

2. *Sistemi di IA etici e valutazione dei rischi*

Un altro filone trasversale della ricerca ha riguardato la costruzione di sistemi di IA eticamente accettabili e giuridicamente responsabili. Diversi *spoke* hanno affrontato il tema con approcci complementari: lo Spoke

6 ha introdotto modelli di *Assumption-Based Argumentation* e di rischio etico, finalizzati a formalizzare un ragionamento «value-sensitive» nei sistemi di IA. Lo Spoke 3 ha invece analizzato i rischi occupazionali legati all'uso dell'IA nella gestione del lavoro e nella valutazione delle prestazioni, mentre lo Spoke 1 ha proposto un quadro teorico ispirato al «pensiero lento» e all'umanesimo digitale, volto a riflettere sulle trasformazioni culturali e morali indotte dall'automazione e dall'uso crescente di algoritmi decisionali.

Queste ricerche convergono nella definizione di un approccio integrato che consideri congiuntamente il rischio legale e l'accettabilità etica delle soluzioni di IA.

3. *Trasparenza, spiegabilità e* accountability

La necessità di garantire trasparenza, spiegabilità e responsabilità nei sistemi di IA rappresenta un punto di convergenza fondamentale tra gli *spokes*. Sul piano tecnico, lo Spoke 7 ha sviluppato controllori basati su *Deep Reinforcement Learning* interpretabili, capaci di trasformare la logica «black box» in sistemi decisionali fondati su regole esplicite. Sul versante giuridico e sociale, lo Spoke 8 ha invece elaborato modelli di ragionamento giuridico basati su regole e schemi di argomentazione causale per affrontare situazioni di incertezza e conflitto normativo. Parallelamente, lo Spoke 2 ha evidenziato come la trasparenza costituisca un requisito non solo tecnico, ma anche etico e legale, in linea con i principi dell'*AI Act*, mentre lo Spoke 1 ha approfondito gli effetti dell'IA generativa sulla percezione e sulla responsabilità, segnalando i rischi legati alla manipolazione delle immagini e alla distorsione della realtà.

4. *L'impiego dell'IA in domini sensibili*

Un ulteriore tema comune ha riguardato l'uso dell'intelligenza artificiale in ambiti ad alta sensibilità giuridica ed etica. Lo Spoke 6 ha analizzato le applicazioni giudiziarie e di *law enforcement*, concentrandosi in particolare sui rischi derivanti dalle classificazioni preventive e dall'uso di logiche predittive nei settori fiscale e penale. Lo Spoke 3 ha studiato la gestione algoritmica del lavoro, i diritti dei lavoratori e i rischi professionali, nonché l'uso dell'IA in agricoltura, nelle campagne elettorali e nella governance regionale. Gli Spoke 1 e 2 si sono focalizzati sui sistemi di IA in ambito sanitario, valutandone gli aspetti regolatori e di policy, mentre gli Spoke

7 e 8 hanno indagato le applicazioni legate alla gestione ambientale, alla conformità legale e al settore assicurativo.

Nel complesso, queste linee di ricerca hanno delineato una visione comune: quella volta a promuovere un ecosistema europeo dell'IA che sia non solo innovativo, ma anche giusto, trasparente e rispettoso dei valori fondamentali dell'Unione.

3. *Le ricerche dei vari* spoke

Andando ancora più nello specifico, e considerando le attività di ogni singolo *spoke*, i gruppi di ricerca del TP1 hanno approfondito i seguenti temi:

Spoke 1

Il gruppo di ricerca dell'Università di Pisa, coordinato per quanto riguarda il profilo etico da Adriano Fabris, ha concentrato i propri studi sull'etica dell'intelligenza artificiale lungo tre linee tra loro interconnesse: i fondamenti metaetici dell'IA, la trasformazione della comunicazione nell'era digitale e le implicazioni etiche delle immagini artificiali. L'obiettivo complessivo è stato quello di promuovere una *human-centered AI*, capace di potenziare la capacità umana di decidere e agire responsabilmente, valorizzando sia l'individuo, sia la collettività. Ciò implica la necessità di ripensare le basi stesse dell'IA per orientarla verso una progettazione affidabile ed etica *by design*, fondata su co-progettazione, validazione e uso consapevole dei programmi e dei dispositivi.

Sul piano metaetico, la ricerca ha approfondito i concetti di «ambiente digitale», «rischio» e «vulnerabilità». L'ambiente digitale è stato interpretato come un sistema di relazioni tra esseri umani e tecnologie, che può assumere forme strumentali, mediali o immersive. Da questa prospettiva nasce il concetto di *ecologia dell'IA*, che mira a gestire le vulnerabilità generate dall'interazione con sistemi intelligenti, applicando il principio di vulnerabilità come fondamento normativo per costruire relazioni basate su cura, responsabilità e consapevolezza. In tale contesto, il gruppo ha proposto una filosofia del digitale che promuove un «pensiero lento» e riflessivo, capace di bilanciare la velocità e l'astrazione dei processi algoritmici, in vista di un nuovo umanesimo che unisca immaginazione tecnica, etica e responsabilità.

Un'attenzione particolare è stata poi rivolta alla comunicazione nell'era *post-verità*, segnata dall'erosione della verità condivisa e dalla mediazione

algoritmica dell'informazione. Il gruppo ha analizzato modelli etici inclusivi e umanocentrici per la comunicazione mediata dall'IA, al fine di preservare il dialogo democratico e la qualità del dibattito pubblico.

Infine, la ricerca ha affrontato l'etica delle immagini artificiali, analizzando come l'IA generativa trasformi immaginazione e percezione visiva, specialmente in ambiti come pubblicità, giornalismo e medicina. Sono stati esaminati i rischi legati a *bias*, *deepfake* e manipolazioni digitali, evidenziando la necessità di integrare principi etici e regolatori europei per prevenire la «defattualizzazione» della realtà e favorire un uso consapevole delle tecnologie visive emergenti.

Spoke 2

Il gruppo di ricerca dell'Università di Trento, coordinato per quanto riguarda il profilo giuridico da Carlo Casonato, ha approfondito in modo sistematico l'analisi del quadro normativo dell'AI Act, concentrandosi sia sulle sue potenzialità sia sulle criticità legate all'attuazione. Particolare attenzione è stata dedicata alle vulnerabilità, alla trasparenza e al principio del controllo umano, oltre che alla regolamentazione dei modelli generali di IA, con un focus sulle implicazioni etiche e giuridiche del loro impiego in ambito medico. In tale contesto, sono state esaminate anche linee guida e codici deontologici.

L'attività si è caratterizzata per un forte approccio interdisciplinare, grazie alla collaborazione tra giuristi, eticisti, medici e informatici, che ha consentito di affrontare in modo concreto temi quali i *bias* algoritmici, la trasparenza e l'evoluzione del quadro regolatorio europeo. Oltre alla partecipazione a diversi convegni nazionali, alcuni membri del gruppo hanno da ultimo partecipato a un seminario scientifico a Bilbao, dedicato proprio a questi aspetti, contribuendo al dibattito internazionale sull'uso responsabile dell'IA in ambito sanitario.

Parallelamente, l'Università di Trento ha svolto un ruolo di riferimento nell'organizzazione della serie biennale di webinar *AI Act 1.01* (marzo-aprile 2024 e giugno 2025), che ha favorito la diffusione e la discussione dei risultati della ricerca: infatti, le registrazioni dei webinar, rese disponibili nella piattaforma YouTube, hanno complessivamente raccolto circa 3000 visualizzazioni. Gli incontri hanno inoltre goduto di un'ampia partecipazione e del contributo di relatori provenienti da discipline e *spoke* diversi. Inoltre, la Fondazione Bruno Kessler (FBK) ha sviluppato una piattaforma di *AI-enactment* basata su infrastruttura MLOps, concepita per supportare

le Pubbliche Amministrazioni nello sviluppo sostenibile di applicazioni IA, integrando strumenti di *accountability*, metriche di qualità e algoritmi per la valutazione etica dei dati e dei modelli.

Spoke 3

Lo Spoke 3, dell'Università di Napoli «Federico II», coordinato da Giovanna De Minico, ha condotto un organico programma di ricerca sulla regolazione europea e nazionale, con particolare attenzione al diritto costituzionale e a quello del lavoro. Esso si è articolato in due aree principali distinte, ma strettamente coordinate tra loro.

A) Nel campo del diritto costituzionale, la De Minico ha approfondito l'evoluzione delle fonti giuridiche nell'*e-society* dall'eteronomia classica, passando per la *co-regulation*, quindi, gli standard armonizzati fino alla tecnica *as source of law*: su questo tema la coordinatrice ha già pubblicato alcuni suoi saggi in riviste italiane e internazionali. Inoltre, il gruppo ha analizzato l'applicazione iniziale del Regolamento (UE) 2024/1689 con particolare riferimento alle proibizioni dell'art. 5 e alla *Comunicazione della Commissione sulle pratiche di IA vietate* (C(2025) 884 final), tema questo al quale si è dedicata particolarmente la Tuozzo. Infine, è stato esaminato il ruolo della governance dell'IA all'interno delle istituzioni italiane ed europee, anche attraverso il contributo al Board sull'Intelligenza Artificiale, istituito dall'AGCOM, di cui la De Minico è componente.

A completamento dei temi generali prima indicati, si menzionano gli studi intorno alla regolamentazione della comunicazione politica e delle campagne elettorali, con attenzione alle decisioni dei tribunali europei e rumeni (Tuozzo); la sostenibilità dell'IA in agricoltura e nella governance dei dati, collegando i diritti individuali e collettivi con quelli sociali e con il mercato unico (De Tullio); l'impatto dell'IA sulla governance regionale e sui rapporti Stato-Regioni (Grossi); l'uso pubblico dell'IA e la necessità di modelli di *Responsible AI* (Palomba); la libertà di espressione e le sfide etiche poste dall'IA generativa nella comunicazione e nella creatività artistica (Parente).

Quanto agli studi in proiezione futura si indica quello sui *neuroright*, che sta sviluppando la De Minico, avvalendosi peraltro di un periodo di studio, svolto all'*University of Pennsylvania*, dove ha contribuito con *speech* e Relazioni ai lavori del *Center of Neuroscience and Society*, in https://neuroethics.upenn.edu/about-us/people/visiting-scholars/. Il secondo studio in atto riguarda invece la recente legge italiana sull'IA n.132/2025, condotto principalmente da Palomba e Parente.

B) Il secondo fronte è quello del diritto del lavoro, la cui ricerca, coordinata da Laura Tebano, ha analizzato l'impatto dei sistemi di gestione algoritmica sul lavoro e sui diritti collettivi, nel contesto dell'*AI Act* e della normativa europea. Sono stati esaminati i sistemi di IA ad alto rischio impiegati nei luoghi di lavoro, inclusi quelli di riconoscimento delle emozioni, con particolare attenzione alla trasparenza, alla tutela della privacy e alla sicurezza dei lavoratori. Ulteriori approfondimenti hanno riguardato le politiche attive per l'occupazione, la prevenzione dei rischi professionali nel lavoro da remoto e i diritti fondamentali degli atleti di *e-sports* in relazione alle normative antidoping.

Spoke 6

Il gruppo di ricerca dell'Università di Bari «Aldo Moro» coordinato da Francesca Alessandra Lisi ha sviluppato un articolato programma di ricerca sul tema dell'accettabilità della Symbiotic Artificial Intelligence (SAI), contribuendo in modo significativo agli obiettivi di TP1. Le attività si sono concentrate su tre dimensioni principali: legale, etico-filosofica e applicativa.

Sul piano giuridico, a partire dall'*AI Act*, i ricercatori del gruppo hanno analizzato i sistemi di IA ad alto rischio destinati all'uso giudiziario e di *law enforcement*, approfondendo in particolare i rischi connessi all'impiego di logiche algoritmiche e di decisioni predittive nei settori della fiscalità e del diritto penale. La riflessione ha messo in luce la debolezza concettuale di un approccio basato su una valutazione *ex ante* del rischio in un sistema, come quello giudiziario, che per sua natura opera *ex post*. Inoltre, sono stati esplorati scenari sensibili, come l'uso dell'IA nei casi di violenza di genere, con l'obiettivo di garantire maggiore tutela e trasparenza.

Dal punto di vista etico-filosofico, la ricerca ha affrontato questioni centrali quali vulnerabilità, inclusione, coscienza artificiale, libero arbitrio e *IA sociale*, proponendo un approccio multilivello alla comprensione del rapporto umano-macchina e alla decostruzione della centralità antropocentrica. Sono stati inoltre condotti studi sui *bias* algoritmici e sulla *fairness*, con particolare attenzione alla prospettiva di genere e ai processi di esclusione o discriminazione insiti nei sistemi automatizzati.

Sul piano applicativo, in collaborazione con l'Università del Salento (progetto eDef-AI), è stato sviluppato un modello di *Assumption-Based Argumentation* arricchito da pesi e calcolato tramite *Answer Set Programming*, utile per il ragionamento etico in ambito sanitario. È stato inoltre proposto un metodo formale per la specifica e la validazione di modelli decisionali etici basati

su *fuzzy Petri nets*, illustrato attraverso un caso di studio in ambito medico (care robot per il monitoraggio del paziente). Parallelamente, lo Spoke ha contribuito alla diffusione dei risultati in ambiti interdisciplinari e nel dibattito pubblico, promuovendo una cultura tecnologica più riflessiva e inclusiva.

Spoke 7

Lo Spoke 7, attivo al Politecnico di Torino, è stato coordinato da Barbara Caputo. Al suo interno, il WP7.6, (Ethical, Legal, Economical and Societal issues in Edge and Exascale AI), coordinato da Andrea Maria Lingua ha sviluppato un articolato programma di ricerca sull'applicazione dell'intelligenza artificiale a dati geospaziali, ambientali e 3D, con particolare attenzione a privacy, equità e interpretabilità. Le attività si sono articolate in più linee di ricerca interconnesse.

Un primo gruppo di ricercatori ha condotto numerosi audit di sistemi di IA (in campo assicurativo, in sistemi di classificazione volti, e in LLM) e studi empirici sulla documentazione dei dataset come mezzo per favorire trasparenza e accountability. Inoltre, sono state condotte tredici interviste con stakeholder per individuare criticità e indicazioni utili alla definizione di un quadro metodologico e di un protocollo condiviso per la generazione di dati sintetici, basato su modelli di ultima generazione adattabili a diversi contesti. È stato inoltre elaborato un sistema di codifica e avviata l'analisi delle trascrizioni delle interviste.

Un altro gruppo ha integrato modelli AI per la segmentazione semantica e la classificazione automatica di immagini a supporto del processo peritale assicurativo. Tali modelli si basano su ontologie di dominio finalizzate a guidare l'etichettatura automatica delle immagini, consentendo una rappresentazione coerente e interpretabile delle variabili decisionali chiave visibili nelle immagini. I risultati sperimentali sono stati discussi con esperti del settore, al fine di validare le raccomandazioni del sistema e di verificarne l'efficacia nel migliorare la trasparenza e l'uniformità del processo decisionale. Parallelamente, sono stati sviluppati modelli generativi e sistemi basati su *PointLLM* per la segmentazione semantica di ambienti urbani e *indoor*, il riconoscimento di specie marine (dataset SardinIA) e la classificazione automatica di immagini satellitari e iperspettrali per l'individuazione del degrado nei beni storici. Inoltre, sono stati sperimentati approcci *NeRF* e *Gaussian Splatting* per la ricostruzione 3D del patrimonio culturale subacqueo e approcci neuro-simbolici per migliorare la segmentazione semantica delle nuvole di punti architettoniche.

Infine, una terza linea di ricerca ha riguardato lo sviluppo di un sistema di controllo avanzato per impianti HVAC, basato su *Deep Reinforcement Learning*. I modelli addestrati sono stati trasformati in controllori interpretabili tramite alberi decisionali, capaci di garantire prestazioni elevate, efficienza energetica e trasparenza operativa. Le attività sono state accompagnate da workshop e corsi universitari sull'uso responsabile dell'IA nei dati geospaziali e ambientali.

Spoke 8

Lo Spoke 8 dell'Università di Bologna coordinato da Giovanni Sartor ha condotto infine un'ampia attività di ricerca dedicata agli aspetti regolatori, giuridici e applicativi dell'intelligenza artificiale, con un approccio interdisciplinare che unisce diritto, etica e tecnologia.

Sul piano regolatorio, il gruppo ha analizzato in profondità il ruolo dell'Unione Europea nel panorama normativo internazionale dell'IA, esaminando gli sviluppi multilaterali e bilaterali e il loro impatto sull'agenda digitale europea. Particolare attenzione è stata dedicata all'interazione tra *AI Act*, *Digital Markets Act* (DMA) e GDPR, con riferimento al consenso degli interessati e alle sue implicazioni giuridiche. È stato inoltre elaborato un framework concettuale per la valutazione qualitativa dei rischi legali dei sistemi di IA, pensato per garantire conformità normativa e tutela dei diritti fondamentali, soprattutto in relazione ai sistemi ad alto rischio e ai modelli di IA generali. La ricerca ha affrontato anche la «*twin transition*» – digitale ed ecologica – studiando l'impatto dell'IA sui modelli occupazionali e sulle trasformazioni del lavoro, tema che sarà approfondito in vista del World Congress 2027.

Nell'ambito della governance dell'IA, sono state esaminate strategie per l'implementazione di *AI regulatory sandboxes* a livello europeo, anche attraverso il progetto EUSAiR, coordinato da A. Rotolo, al fine di promuovere la cooperazione tra Stati membri e standardizzare i processi di sperimentazione regolatoria.

Sul versante più tecnico-giuridico, lo Spoke ha sviluppato modelli di ragionamento giuridico basati su regole, trattando il *legal reasoning* come un compito di classificazione binaria. Sono stati integrati la logica deontica e l'*argumentation theory* per gestire obblighi conflittuali e causalità legale, e proposti strumenti di *causal model checking* per analizzare la responsabilità giuridica. Infine, la ricerca ha esplorato l'uso dei Large Language Models (LLM) nel diritto, valutandone il potenziale nel rilevamento

di clausole vessatorie e nell'allineamento tra logica legale, linguaggio naturale e standard normativi, come il Codice della Strada.

English title: What ethical and legal criteria should apply to Artificial Intelligence? The work of WP 1 of the FAIR project

Abstract

The contribution outlines the work carried out and the results achieved within the Transversal Project (TP) of the PNRR FAIR Project. It begins by providing an overall presentation of the project, followed by an in-depth analysis of the specific features of the TP in all its components, namely the research groups from various Italian universities that belong to some of the Spokes. The TP's ethical and legal focus is addressed by these groups through different approaches, all aimed at creating a digital ecosystem that is ethical, legally compliant, and human-centered.

Keywords: AI; FAIR prject; trasversal project; ethical and legal values.

Adriano Fabris
Università di Pisa
adriano.fabris@unipi.it

Carlo Casonato
Università di Trento
carlo.casonato@unitn.it