

**I INTERNATIONAL EXPERIENCE  
PERUGIA - ITÁLIA**

**INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA  
DIGITAL IV**

**YURI NATHAN DA COSTA LANNES**

**HERON JOSÉ DE SANTANA GORDILHO**

Todos os direitos reservados e protegidos. Nenhuma parte destes anais poderá ser reproduzida ou transmitida sejam quais forem os meios empregados sem prévia autorização dos editores.

**Diretoria - CONPEDI**

**Presidente** - Profa. Dra. Samyra Haydêe Dal Farra Naspolini - FMU - São Paulo

**Diretor Executivo** - Prof. Dr. Orides Mezzaroba - UFSC - Santa Catarina

**Vice-presidente Norte** - Prof. Dr. Jean Carlos Dias - Cesupa - Pará

**Vice-presidente Centro-Oeste** - Prof. Dr. José Querino Tavares Neto - UFG - Goiás

**Vice-presidente Sul** - Prof. Dr. Leonel Severo Rocha - Unisinos - Rio Grande do Sul

**Vice-presidente Sudeste** - Profa. Dra. Rosângela Lunardelli Cavallazzi - UFRJ/PUCRio - Rio de Janeiro

**Vice-presidente Nordeste** - Prof. Dr. Raymundo Juliano Feitosa - UNICAP - Pernambuco

**Representante Discente:** Prof. Dr. Abner da Silva Jaques - UPM/UNIGRAN - Mato Grosso do Sul

**Conselho Fiscal:**

Prof. Dr. José Filomeno de Moraes Filho - UFMA - Maranhão

Prof. Dr. Caio Augusto Souza Lara - SKEMA/ESDHC/UFMG - Minas Gerais

Prof. Dr. Valter Moura do Carmo - UFERSA - Rio Grande do Norte

Prof. Dr. Fernando Passos - UNIARA - São Paulo

Prof. Dr. Edinilson Donisete Machado - UNIVEM/UENP - São Paulo

**Secretarias**

**Relações Institucionais:**

Prof. Dra. Claudia Maria Barbosa - PUCPR - Paraná

Prof. Dr. Heron José de Santana Gordilho - UFBA - Bahia

Profa. Dra. Daniela Marques de Moraes - UNB - Distrito Federal

**Comunicação:**

Prof. Dr. Robison Tramontina - UNOESC - Santa Catarina

Prof. Dr. Liton Lanes Pilau Sobrinho - UPF/Univali - Rio Grande do Sul

Prof. Dr. Lucas Gonçalves da Silva - UFS - Sergipe

**Relações Internacionais para o Continente Americano:**

Prof. Dr. Jerônimo Siqueira Tybusch - UFSM - Rio Grande do Sul

Prof. Dr. Paulo Roberto Barbosa Ramos - UFMA - Maranhão

Prof. Dr. Felipe Chiarello de Souza Pinto - UPM - São Paulo

**Relações Internacionais para os demais Continentes:**

Profa. Dra. Gina Vidal Marcilio Pompeu - UNIFOR - Ceará

Profa. Dra. Sandra Regina Martini - UNIRITTER / UFRGS - Rio Grande do Sul

Profa. Dra. Maria Claudia da Silva Antunes de Souza - UNIVALI - Santa Catarina

**Educação Jurídica**

Profa. Dra. Viviane Coêlho de Séllos Knoerr - Unicuritiba - PR

Prof. Dr. Rubens Beçak - USP - SP

Profa. Dra. Livia Gaigher Bosio Campello - UFMS - MS

**Eventos:**

Prof. Dr. Yuri Nathan da Costa Lannes - FDF - São Paulo

Profa. Dra. Norma Sueli Padilha - UFSC - Santa Catarina

Prof. Dr. Juraci Mourão Lopes Filho - UNICHRISTUS - Ceará

**Comissão Especial**

Prof. Dr. João Marcelo de Lima Assafim - UFRJ - RJ

Profa. Dra. Maria Creusa De Araújo Borges - UFPB - PB

Prof. Dr. Antônio Carlos Diniz Murta - Fumec - MG

Prof. Dr. Rogério Borba - UNIFACVEST - SC

I61

Inteligência Artificial: Desafios da Era Digital IV [Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Heron José de Santana Gordilho, Yuri Nathan da Costa Lannes. – Florianópolis: CONPEDI, 2025.

Inclui bibliografia

ISBN: 978-65-5274-093-9

Modo de acesso: [www.conpedi.org.br](http://www.conpedi.org.br) em publicações

Tema: Inteligência Artificial e Sustentabilidade na Era Transnacional

1. Direito – Estudo e ensino (Pós-graduação) – Encontros Internacionais. 2. Inteligência Artificial. 3. Desafios da Era Digital. I International Experience Perugia – Itália. (1: 2025 : Perugia, Itália).

CDU: 34



# I INTERNATIONAL EXPERIENCE PERUGIA - ITÁLIA

## INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL IV

---

### **Apresentação**

O I International Experience Perugia - Itália, organizado pelo CONPEDI, UNIPG e UNIVALI, e com apoio da FDF, PPGD da UFSC e PPGD da UIVALI, apresentou como temática central “Inteligência Artificial e Sustentabilidade na Era Transnacional”. Esta questão suscitou intensos debates desde o início e, no decorrer do evento, com a apresentação dos trabalhos previamente selecionados, fóruns e painéis que na no ambiente digital ocorreram.

Os trabalhos contidos nesta publicação foram apresentados como artigos no Grupo de Trabalho “Inteligência Artificial: Desafios da Era Digital IV”, realizado nos dias 29 e 30 de maio de 2025, que passaram previamente por dupla avaliação cega por pares. Encontram-se os resultados de pesquisas desenvolvidas em diversos cursos de Direito pelos graduandos e também por Programas de Pós-Graduação em Direito, que retratam parcela relevante dos estudos que têm sido desenvolvidos na temática central do Grupo de Trabalho.

As temáticas abordadas decorrem de intensas e numerosas discussões que acontecem, com temas que reforçam a diversidade cultural e as preocupações que abrangem problemas relevantes e interessantes.

Os trabalhos estão organizados em cinco eixos temáticos:

#### 1. Inteligência Artificial, Ética e Governança

Este eixo reúne trabalhos que exploram as implicações éticas, os desafios de governança e a necessidade de regulação da Inteligência Artificial, abordando questões como vieses algorítmicos, transparência, responsabilidade e o papel do judiciário.

A NEUTRALIDADE ALGORÍTMICA EM XEQUE: VIESES, DISCRIMINAÇÃO E TRANSPARÊNCIA (Jéssica Cindy Kempfer, Mariana Emília Bandeira)

DA INTELIGÊNCIA ARTIFICIAL ÀS METÁFORAS DA HUMANIZAÇÃO: POR UMA ÉTICA PARA O USO DA IA PELO JUDICIÁRIO(Léo Peruzzo Júnior, Gilson Bonato, Gabriela Cristine Buzzi)

COMPLIANCE E DUE DILIGENCE NA GESTÃO ALGORÍTMICA DA MOBILIDADE URBANA: DESAFIOS E IMPACTOS ÉTICOS NAS SMART CITIES (Luiz Dalago Júnior, Cristiani Fontanela, Giovanni Olsson)

A REGULAÇÃO TRANSNACIONAL DAS PLATAFORMAS DIGITAIS E SUAS FERRAMENTAS DE INTELIGÊNCIA ARTIFICIAL: UMA EMERGÊNCIA GLOBAL (Álvaro Luiz Poglia)

A INTELIGÊNCIA ARTIFICIAL E O DISCURSO JURÍDICO PERANTE OS TRIBUNAIS (Andre Lipp Pinto Basto Lupi)

INTELIGÊNCIA ARTIFICIAL, BIG DATA E DEMOCRACIA: DESAFIOS, RISCOS E O FUTURO DA GOVERNANÇA DIGITAL (Jéssica Cindy Kempfer, Mariana Emília Bandeira)

LA CORRELAZIONE FRA SPIEGABILITÀ ED INNOVAZIONE: NUOVE FRONTIERE DELLA RESPONSABILITÀ CIVILE (José Luiz de Moura Faleiros Júnior, Andre Vecchi, Victor Rezende Goulart)

RESPONSABILITÀ CIVILE E SISTEMI DI INTELLIGENZA ARTIFICIALE: PANORAMICA REGOLATORIA E POSSIBILI APPROCCI IN RELAZIONE AL NESSO DI IMPUTAZIONE (José Luiz de Moura Faleiros Júnior, Andre Vecchi, Victor Rezende Goulart)

## 2. Proteção de Dados e Direitos Fundamentais na Era Digital

Este eixo aborda a intersecção entre a proteção de dados, a LGPD, o Big Data e a salvaguarda de direitos fundamentais, especialmente no contexto da exposição de dados e do uso da Inteligência Artificial.

I LIMITI DELLA LEGGE GENERALE SULLA PROTEZIONE DEI DATI (LGPD) DEL BRASILE CONSIDERANDO LE BUONE PRATICHE ESG AI FINI DELLA PROTEZIONE DEI DATI SENSIBILI NEI SISTEMI DI INTELIGÊNCIA ARTIFICIALE (Grace Ladeira Garbaccio, Francisco Leonardo Silva Neto, Consuêla Félix De Vasconcelos Neta)

BIG DATA E DIREITOS FUNDAMENTAIS SOB A PERSPECTIVA DA LEI GERAL DE PROTEÇÃO DE DADOS (LGPD): DIREITO À INTIMIDADE NA ERA DA EXPOSIÇÃO

MUNDIAL DE DADOS E INTELIGÊNCIA ARTIFICIAL (Andréa Arruda Vaz, Angela Rank Linzmeier, Tais Martins)

### 3. Direitos Humanos e Vulnerabilidades no Cenário Digital

Este eixo congrega trabalhos que analisam o impacto das tecnologias digitais, incluindo a IA, na efetivação dos direitos humanos, com foco em grupos vulneráveis, desinformação e novas formas de discriminação.

INTELIGÊNCIA ARTIFICIAL GENERATIVA E COLONIALISMO JURÍDICO: DESAFIOS PARA O DIREITO À AUTOIDENTIFICAÇÃO INDÍGENA NO JUDICIÁRIO BRASILEIRO (Adriana Da Silva Chaves, Ana Beatriz Gonçalves Carvalho, Andre Augusto Salvador Bezerra)

INOVAÇÃO DIGITAL E DIREITOS HUMANOS: INTELIGÊNCIA ARTIFICIAL E BLOCKCHAIN COMO MOTORES DE TRANSFORMAÇÃO CONSTITUCIONAL (Mariela Sanchez Salas)

DESINFORMAÇÃO CIENTÍFICA NA SAÚDE: A DEVIDA DILIGÊNCIA EM DIREITOS HUMANOS NAS PLATAFORMAS DIGITAIS NA ERA DA IA (Anna Luisa Walter de Santana, Cinthia Obladen de Almendra Freitas)

ANÁLISE DA INCLUSÃO PREVIDENCIÁRIA DA POPULAÇÃO EM SITUAÇÃO DE RUA FRENTE À AUTOMAÇÃO DO SERVIÇO PÚBLICO BRASILEIRO (Alice Arlinda Santos Sobral, Nicolle Patrice Pereira Rocha)

A EFETIVAÇÃO DOS DIREITOS FUNDAMENTAIS DAS PESSOAS VULNERÁVEIS NA ERA DA INTELIGÊNCIA ARTIFICIAL (Claudia Isabele Freitas Pereira Damous)

DA UTILIZAÇÃO DE INTELIGÊNCIA ARTIFICIAL PARA FINS DE FACILITAÇÃO DOS PROCESSOS DE ADOÇÃO: A EXPERIÊNCIA DO PROGRAMA “FAMILY MATCH” À LUZ DOS DIREITOS FUNDAMENTAIS E DA PERSONALIDADE (Raissa Arantes Tobbin, Valéria Silva Galdino Cardin, Tereza Rodrigues Vieira)

O IMPACTO DAS DEEPPAKES AOS DIREITOS FUNDAMENTAIS E DA PERSONALIDADE DE CRIANÇAS E ADOLESCENTES (Raissa Arantes Tobbin, Valéria Silva Galdino Cardin, Tereza Rodrigues Vieira)

#### 4. Liberdade de Expressão e Desafios da Sociedade Digital

Este eixo agrupa pesquisas que abordam a liberdade de expressão no ambiente digital, os impactos das plataformas nas mídias sociais e a análise de conceitos jurídicos em um cenário de rápida evolução tecnológica.

**LIBERDADE DE EXPRESSÃO E SUAS IMPLICAÇÕES NAS MÍDIAS SOCIAIS** (Andréa Arruda Vaz, Gleyziele De Oliveira Aragao Mascarenhas, Janaina Leite Polchlopek)

#### 5. Direito Comparado e Análises Sociojurídicas

Este eixo reúne estudos que utilizam a perspectiva comparada ou que realizam análises mais amplas sobre fenômenos sociojurídicos, não diretamente ligados à IA, mas relevantes para o contexto do direito contemporâneo.

**JUSTIÇA E LIBERDADE NAS POLÍTICAS DE COMBATE À POBREZA: ANÁLISE COMPARATIVA DO BRASIL, ÍNDIA E ÁFRICA DO SUL SOB A ÓTICA DE RAWLS E SEM** (Isabela Domingos, Gabriela Trentin Zandoná)

**EXCLUSÃO E DISCRIMINAÇÃO DIGITAL: INSIGHTS COMPARATIVOS DO BRASIL E DA ITÁLIA NO CENÁRIO DIGITAL EM EVOLUÇÃO** (Isabela Domingos, José Sérgio da Silva Cristóvam, Roberto Miccú)

Espera-se, então, que o leitor possa vivenciar parcela destas discussões por meio da leitura dos textos. Agradecemos a todos os pesquisadores, colaboradores e pessoas envolvidas nos debates e organização do evento pela sua inestimável contribuição e desejamos uma proveitosa leitura!

Prof. Dr. Heron José de Santana Gordilho - Universidade Federal da Bahia - UFBA

Prof. Dr. Yuri Nathan da Costa Lannes – Faculdade de Direito de Franca/FDF

# LA CORRELAZIONE FRA SPIEGABILITÀ ED INNOVAZIONE: NUOVE FRONTIERE DELLA RESPONSABILITÀ CIVILE

## THE CONVERGENCE OF EXPLAINABILITY AND INNOVATION: NEW FRONTIERS IN CIVIL LIABILITY

José Luiz de Moura Faleiros Júnior <sup>1</sup>

Andre Vecchi <sup>2</sup>

Victor Rezende Goulart <sup>3</sup>

### Resumo

Il presente lavoro esamina la correlazione tra la spiegabilità algoritmica (explainability) e la responsabilità civile, con particolare riferimento al principio di accountability delineato nell'AI Act europeo. Lo studio argomenta come la trasparenza, seppur necessaria, non sia sufficiente a garantire un'efficace tutela dei diritti e una prevenzione adeguata del danno in un contesto di crescente complessità tecnologica. Integrando prospettive normative e dottrinali, emerge la centralità di un approccio multilayered, che superi la semplice distinzione tra responsabilità soggettiva e oggettiva, per abbracciare un modello in cui i soggetti coinvolti (sviluppatori, fornitori, utilizzatori e autorità) cooperino in base a doveri di rendicontazione e tracciabilità. In tale quadro, la funzione preventiva della responsabilità civile acquisisce ulteriore rilevanza: l'adozione di audit costanti, la predisposizione di schemi assicurativi mirati e la possibilità di effettuare valutazioni d'impatto consentono, infatti, di ridurre i rischi ex ante, valorizzando il ruolo pedagogico della responsabilità. L'analisi evidenzia altresì come gli obblighi di spiegabilità possano favorire una più equa ripartizione delle responsabilità tra i vari attori, mitigando le criticità connesse all'opacità dei processi decisionali automatizzati. In conclusione, l'unione di spiegabilità e accountability costituisce un elemento fondamentale per la costruzione di sistemi di IA affidabili, etici e rispettosi del quadro giuridico europeo.

**Palavras-chave:** Spiegabilità, Accountability, Responsabilità civile, Funzione preventiva, Intelligenza artificiale

### Abstract/Resumen/Résumé

This paper examines the relationship between algorithmic explainability and civil liability, with a focus on the principle of accountability set forth in the European AI Act. The analysis suggests that transparency, although necessary, is insufficient to ensure effective protection of rights and adequate prevention of harm in an increasingly complex technological context.

---

<sup>1</sup> Doutor em Direito pela Universidade de São Paulo. Professor da Faculdade Milton Campos. E-mail: josefaleirosjr@outlook.com

<sup>2</sup> Mestrando em Direito pela Faculdade Milton Campos. E-mail: andre.vecchi.lima@gmail.com

<sup>3</sup> Mestrando em Direito pela Faculdade Milton Campos. E-mail: victorrezende826@gmail.com

By integrating normative and doctrinal perspectives, the paper underscores the centrality of a multilayered approach that transcends the simple dichotomy between subjective and strict liability, promoting a model where stakeholders (developers, providers, users, and authorities) cooperate under specific duties of reporting and traceability. In this framework, the preventive function of civil liability gains renewed importance: constant auditing, specialized insurance schemes, and thorough impact assessments help to reduce risks ex ante, highlighting the pedagogical nature of liability. The study also shows how explainability obligations may foster a fairer allocation of responsibilities among various actors, thus mitigating issues related to the opacity of automated decision-making processes. In conclusion, the synergy of explainability and accountability emerges as a cornerstone for the development of trustworthy, ethical, and legally compliant AI systems in the European regulatory landscape.

**Keywords/Palabras-claves/Mots-clés:** Explainability, Accountability, Civil liability, Preventive function, Artificial intelligence

## 1. INTRODUZIONE

Lo sviluppo esponenziale dell'intelligenza artificiale (IA) ha suscitato dibattiti giuridici e accademici in merito ai parametri di responsabilità, trasparenza e governance dei sistemi algoritmici. In tale prospettiva, emerge il “principio di spiegabilità algoritmica” come meccanismo fondamentale per garantire la comprensione, da parte degli utenti e delle autorità competenti, degli elementi costitutivi e delle conseguenze delle decisioni automatizzate. L'articolazione tra la spiegabilità e altre direttive etiche e legali si rivela imprescindibile per la costruzione di una regolamentazione solida ed efficace, specialmente alla luce delle nuove configurazioni normative in via di definizione in ambito europeo, guidate dal Regolamento n. 1689 del 2024, comunemente noto come Regolamento Europeo sull'Intelligenza Artificiale o *AI Act* dell'Unione Europea.

La presente ricerca propone, come problema di studio, la necessità di definire in modo adeguato il principio di spiegabilità algoritmica, distinguendolo dal principio di trasparenza, solitamente menzionato nelle discussioni affini sull'IA. In particolare, si intende evidenziare come l'approccio europeo, riflesso nel recente *AI Act*, possa ispirare e sorreggere le regolamentazioni in diversi ordinamenti giuridici. Tale questione acquisisce rilevanza nella misura in cui la precisione concettuale si configura come un fattore determinante per la definizione di obblighi chiari e la fissazione di standard minimi di informazione e di verifica, garantendo la dovuta certezza del diritto.

Il suddetto Regolamento n. 1689/2024, il cui scopo concerne la governance dei sistemi intelligenti in molteplici settori, presenta disposizioni ampie in materia di obblighi di trasparenza e spiegabilità, in particolare introducendo richieste tecniche e organizzative rivolte agli sviluppatori e agli operatori di IA. La distinzione tra trasparenza, intesa come messa a disposizione generale di informazioni, e spiegabilità, relativa alla chiarezza del ragionamento algoritmico, rappresenta quindi un punto nodale. La robustezza di questo quadro normativo europeo non solo apporta benefici ai cittadini dell'Unione, ma proietta i suoi effetti a livello globale, influenzando anche le normative emergenti.

In effetti, non si può sottovalutare il potenziale di armonizzazione regolamentare nello scenario internazionale. Diversi Paesi in tutto il mondo stanno elaborando o rivedendo le proprie leggi specifiche sull'IA, attingendo alle buone pratiche internazionali. Tra questi, spicca il Brasile, il cui Progetto di Legge n. 2.338/2023 affronta la regolamentazione

dell'intelligenza artificiale. In tale ottica, l'osservanza e l'eventuale incorporazione dei parametri definiti dall'*AI Act* europeo possono fornire un solido riferimento per la formulazione di regole interne, che tengano conto delle peculiarità economiche, sociali e culturali brasiliane.

Nell'ambito di questo articolo, si adotta una metodologia di tipo qualitativo, basata sul metodo deduttivo e sull'analisi delle norme pertinenti, in particolare il Regolamento Europeo sull'Intelligenza Artificiale (n. 1689/2024) e il Progetto di Legge brasiliano n. 2.338/2023. Lo studio analizza, in modo comparativo e critico, come la disciplina giuridica europea stabilisca parametri per la spiegabilità dei sistemi di IA e in che misura tali direttive possano essere incorporate, in forma adeguata, in altre legislazioni nazionali. Il corpus normativo è esaminato alla luce di riferimenti dottrinali che toccano sia il campo del Diritto Digitale sia le discussioni etiche in ambito IA.

L'adozione del metodo deduttivo si giustifica per la necessità di partire dai principi generali individuati nell'ordinamento europeo e verificare la loro applicabilità negli ordinamenti che a tale fonte regolamentare si ispirano. Pertanto, l'analisi delle disposizioni contenute nell'*AI Act*, in particolare le esigenze di valutazione del rischio, l'istituzione di meccanismi di audit e l'obbligo di fornire spiegazioni comprensibili, verrà messa a confronto con le proposte del progetto legislativo brasiliano, considerando le specificità del contesto nazionale.

Come ipotesi di ricerca, si sostiene che l'influenza europea, fondata sull'*AI Act*, giocherà un ruolo cruciale nella definizione di un concetto giuridicamente coerente e universalizzabile del principio di spiegabilità algoritmica. La premessa è che la precisione concettuale favorisca il superamento delle lacune nella comprensione di come e perché un sistema di IA prenda determinate decisioni, potenziando la tutela dei diritti fondamentali e la mitigazione dei rischi connessi all'uso di tali tecnologie. Inoltre, una volta definiti parametri chiari, risulta più agevole lo sviluppo di tecnologie conformi a standard etici e normativi.

In questa prospettiva, si mira infine a dimostrare che una corretta comprensione e adozione del principio di spiegabilità algoritmica, alla luce dell'esperienza europea, costituisce un requisito imprescindibile per il progresso normativo in materia di intelligenza artificiale. Solo a partire da una nozione solida e articolata, che vada oltre i limiti generici del principio di trasparenza, sarà possibile garantire l'effettività degli obblighi giuridici e tutelare

gli interessi sociali e individuali. Il presente articolo intende pertanto fornire un contributo al dibattito accademico-giuridico, delineando una prospettiva comparativa e approfondita sul tema e indicando percorsi affinché le normative nazionali promuovano simultaneamente l'innovazione e la protezione dei titolari di diritti interessati dalle decisioni automatizzate.

## **2. LA SPIEGABILITÀ ALGORITMICA: CONCETTO E PROSPETTIVE**

Il concetto di spiegabilità algoritmica, previsto nel recente Regolamento Europeo sull'Intelligenza Artificiale (Regolamento n. 1689/2024 o *AI Act*), emerge come risposta alla crescente complessità dei sistemi di IA e alle conseguenze giuridiche e sociali che derivano dalle loro decisioni automatizzate. Stabilendo l'obbligo, per sviluppatori e operatori, di fornire giustificazioni comprensibili sul funzionamento degli algoritmi, l'*AI Act* mira a superare la semplice nozione di trasparenza. Mentre la trasparenza si concentra sulla messa a disposizione di informazioni generali riguardo il sistema, la spiegabilità si focalizza sul come e sul perché vengono ottenuti determinati risultati, rafforzando così la garanzia che le persone interessate comprendano i criteri decisionali.

La ragione critica per l'introduzione di questo principio nell'ordinamento europeo è legata all'imperativo di responsabilità e certezza del diritto. Di fronte alle cosiddette "scatole nere" generate dagli algoritmi, la spiegabilità diventa un parametro essenziale per valutare l'adeguatezza delle decisioni automatizzate e verificarne la compatibilità con i principi basilari del diritto. In tale prospettiva, autori come Coeckelbergh (2020) sottolineano che la possibilità di contestare e pretendere giustificazioni dai sistemi autonomi è strettamente connessa alla legittimità della tecnologia in rapporto ai valori democratici e ai diritti fondamentali.

L'inserimento del principio di spiegabilità nell'*AI Act* si fonda anche sulla necessità di definire più chiaramente la responsabilità per eventuali danni. Questa prospettiva è illustrata da Teubner e Beckers (2023), che analizzano le sfide connesse all'attribuzione di responsabilità rispetto al "comportamento" algoritmico. In contesti complessi, è necessario chiarire se l'errore derivi da un insufficiente addestramento dei dati, da fallimenti nella modellazione matematica o da lacune nella supervisione umana. La spiegabilità consente di circoscrivere le sfere di responsabilità, contribuendo così a ridurre i cosiddetti "vuoti di responsabilità".

In chiave critica, si osserva come l'*AI Act*, distinguendo trasparenza e spiegabilità, evidenzi la necessità di metodologie rigorose per la revisione e la verifica. Autori come Nicholas (2020) sottolineano che l'analisi dettagliata dei criteri decisionali non solo favorisce la comprensione esterna, ma migliora anche il processo di sviluppo stesso, poiché spinge i progettisti a documentare e motivare le scelte di progettazione. Pertanto, la spiegabilità va oltre la mera diffusione di informazioni e si configura quale requisito metodologico.

Alcune prospettive concettuali integrano questo dibattito. Wachter (2022), ad esempio, discute la formazione di “gruppi algoritmici” e indaga in che misura la spiegabilità possa contribuire a contrastare eventuali pratiche discriminatorie risultanti da segmentazioni statistiche non direttamente collegate a categorie tutelate dalla legge. Allo stesso modo, Hildebrandt (2018) richiama l'attenzione sui rischi connessi a un modello di regolamentazione puramente matematico, sostenendo che la spiegabilità debba essere compresa come parte di un ordine giuridico che valorizzi la possibilità di contestazione e tutela dei diritti.

Inoltre, ricerche come quelle di Mökander et al. (2022) e Pasquale (2022) indicano che l'esperienza europea intorno all'*AI Act* possa esercitare un'influenza significativa sulle discussioni regolatorie in altre aree, come nel caso dell'*Algorithmic Accountability Act* statunitense. In tal modo, la definizione solida di spiegabilità nel contesto europeo diventa fonte di ispirazione a livello globale, stabilendo standard di riferimento nella creazione di cornici normative dedicate all'IA.

Per quanto riguarda l'ampliamento della regolamentazione, si può osservare l'importanza di tale concetto per altri Paesi che stanno avviando, autonomamente, la regolamentazione dei sistemi intelligenti. In Brasile, ad esempio, il Projeto de Lei n. 2.338/2023, che disciplina l'intelligenza artificiale, riflette la volontà di allinearsi alle migliori pratiche internazionali. L'influenza europea si rivela determinante per l'adozione di principi volti a garantire i diritti individuali e a prevenire abusi nell'impiego di strumenti algoritmici.

Questa influenza non si limita a meri “trapianti” di norme, ma implica un adeguamento contestualizzato. È dunque fondamentale che l'ordinamento brasiliano, nell'assimilare il principio di spiegabilità, tenga conto delle peculiarità locali, come la diversità socioculturale e le disuguaglianze strutturali. Il concetto deve operare come

salvaguardia, non solo per garantire una trasparenza estesa, ma anche per proteggere i gruppi vulnerabili dalle decisioni discriminatorie basate su profili opachi.

Dal punto di vista metodologico, questa ispirazione europea comporta la necessità di prevedere procedure di valutazione dell'impatto algoritmico, comprese verifiche indipendenti e la pubblicazione di rapporti esplicativi. Tali meccanismi contribuiscono a rafforzare la fiducia del pubblico nell'impiego dei sistemi intelligenti, poiché, come afferma Nicholas (2020), la comprensibilità del processo decisionale costituisce il fondamento della legittimità delle applicazioni di IA agli occhi della società.

Inoltre, come evidenzia Hildebrandt (2018), la spiegabilità svolge il ruolo di “ponte” tra il linguaggio tecnico e i principi giuridici. Oltre a fornire dettagli su codici e modelli, occorre un impegno di traduzione di tali aspetti in termini comprensibili a giudici, avvocati, regolatori e ai singoli interessati. In questo modo, la spiegabilità risulta compatibile con la salvaguardia del due process of law, garantendo un'ampia possibilità di difesa e contraddittorio anche in presenza di decisioni automatizzate.

Per quanto concerne l'avanzamento di iniziative regolatorie a livello globale, si riscontra che la costruzione di un quadro internazionale per la governance dell'IA includa il principio di spiegabilità come elemento trasversale. Tale dinamica, sostenuta dai documenti di organismi multilaterali, favorisce una convergenza intorno a requisiti condivisi per un uso responsabile degli algoritmi. L'influenza dell'*AI Act* europeo si concretizza in numerosi progetti di legge che mirano a bilanciare l'innovazione tecnologica con la tutela dei diritti fondamentali.

Quindi, la concettualizzazione della spiegabilità algoritmica definita nell'*AI Act* europeo riveste un'importanza sia teorica sia pratica, con ricadute rilevanti per le iniziative di regolamentazione in tutto il mondo, in particolare in Brasile. L'ipotesi che l'Europa svolga un ruolo determinante nel dibattito deriva dalla sua esperienza normativa e dal suo approccio pionieristico alla valutazione critica dei sistemi intelligenti. Di conseguenza, la convergenza degli sforzi legislativi e l'adozione di parametri di spiegabilità rafforzano i meccanismi di responsabilizzazione e di protezione dei cittadini, promuovendo nuove fasi di sviluppo etico e giuridico dell'intelligenza artificiale.

Un ulteriore aspetto da considerare riguarda la formazione e la sensibilizzazione degli attori coinvolti nello sviluppo e nell'implementazione di sistemi di IA. Non basta,

infatti, stabilire principi di spiegabilità se sviluppatori, manager, autorità di vigilanza e utenti finali non possiedono la competenza necessaria per interpretare correttamente le spiegazioni fornite. La disponibilità di linee guida chiare e di percorsi formativi specializzati facilita la comprensione delle dinamiche algoritmiche e incrementa la capacità di riconoscere eventuali lacune o disfunzioni nei sistemi adottati.

Parallelamente, assume rilievo la creazione di un ecosistema di ricerca interdisciplinare capace di affrontare sfide sempre più complesse. Giuristi, ingegneri, sociologi, filosofi ed esperti di etica possono collaborare per elaborare metodologie che coniughino rigorosamente aspetti tecnici e principi normativi. Tale sforzo congiunto può generare standard condivisi di valutazione dell'IA e alimentare uno scambio continuo tra il mondo accademico, quello imprenditoriale e le istituzioni, evitando che la regolamentazione diventi obsoleta di fronte a innovazioni rapide.

In definitiva, la spiegabilità algoritmica, come definita nell'*AI Act* e arricchita dai contributi della letteratura internazionale, rappresenta un pilastro centrale per sostenere un processo decisionale automatizzato che sia al contempo efficiente, equo e responsabile. Le prospettive di evoluzione legislativa su scala globale, così come i progetti di legge in Brasile e altrove, sottolineano l'importanza di questo concetto per orientare lo sviluppo di tecnologie che rispettino i diritti fondamentali e le esigenze sociali. Resta ancora molto da fare, ma il dibattito aperto e la convergenza tra diverse discipline indicano un percorso costruttivo verso una governance dell'IA sempre più solida e inclusiva.

### **3. SPIEGABILITÀ ALGORITMICA E RESPONSABILITÀ CIVILE: LA SINERGIA CON IL PRINCIPIO DI *ACCOUNTABILITY***

La crescente rilevanza della spiegabilità algoritmica (*explainability*) nel contesto dell'intelligenza artificiale ha stimolato un ampio dibattito riguardo al suo impatto sul regime di responsabilità civile. In particolare, l'*AI Act* europeo (Regolamento n. 1689/2024) non si limita a chiedere trasparenza nella progettazione e nell'uso dei sistemi di IA, ma introduce esplicitamente l'esigenza di rendere conto (*accountability*) delle scelte algoritmiche in termini di tracciabilità, comprensibilità e giustificabilità (Faleiros Júnior, 2024). Tale esigenza si colloca in un panorama in cui la prevenzione del danno e la corretta allocazione del rischio diventano centrali nell'ottica di una responsabilità che si estende ai diversi soggetti coinvolti.

Il principio di *accountability* impone che gli attori che sviluppano, gestiscono o impiegano sistemi di IA rispondano dei processi decisionali sottostanti, garantendo non solo la disponibilità di spiegazioni tecniche, ma anche di giustificazioni adeguate di fronte alle possibili conseguenze dannose. Come sottolineato da Comandé (2019), il passaggio dalla semplice liability (focalizzata sull'individuazione del responsabile post-danno) all'idea di *accountability* (fondata su misure preventive e audit costanti) risulta cruciale per assicurare che i rischi algoritmici siano monitorati e mitigati ex ante.

Secondo alcuni autori, tra cui Pasquale (2017; 2020), la spiegabilità algoritmica, se adeguatamente implementata, facilita la tracciabilità dell'operato di sistemi di IA, consentendo di attribuire correttamente responsabilità in caso di errore o di danno. Ciò significa che l'*accountability* funge da "ponte" tra il principio di spiegabilità e la responsabilità civile, poiché rende possibili verifiche tecniche in sede giudiziale o amministrativa e permette di individuare se vi sia stata una violazione di obblighi di diligenza nella progettazione, nell'addestramento o nell'applicazione dell'algoritmo.

Nella prospettiva del regolamento europeo, la prevedibilità e la prevenzione del danno diventano elementi centrali: la presenza di una documentazione esplicita sui processi decisionali consente non solo di analizzare gli errori ex post, ma anche di predisporre protocolli di testing e validazione ex ante (Coeckelbergh 2020). L'obbligo di fornire spiegazioni in forma intelligibile, inoltre, contribuisce a ridurre il rischio di "black box", fenomeno che Hildebrandt (2018) definisce come l'oscurità computazionale che ostacola la comprensione del funzionamento dei modelli algoritmici.

D'altro canto, il collegamento tra spiegabilità e responsabilità non può prescindere dalle implicazioni di carattere tecnico-legale. Come evidenziato da Nicholas (2020), la sola disponibilità di informazioni sui parametri di un algoritmo non equivale necessariamente a una spiegabilità effettiva. La responsabilità civile, infatti, si compie pienamente soltanto se le "spiegazioni" facilitano la contestazione di risultati potenzialmente discriminatori o lesivi, soddisfacendo così anche l'esigenza di prevenzione e di risarcibilità del danno.

Nel quadro dell'*AI Act*, la definizione di explicability include una descrizione delle motivazioni e dei criteri di base adottati dal sistema, ma lascia spazio a implementazioni differenziate a seconda del livello di rischio associato all'applicazione. Su questo punto, Wachter (2022) esprime la necessità di ricondurre l'uso di gruppi algoritmici e profilazioni

complesse a una forma di *accountability* più strutturata, tale da prevenire disparità di trattamento. In tal senso, si avverte l'importanza di disposizioni capaci di rafforzare la responsabilità civile in un'ottica non solo riparatoria, ma anche preventiva.

Sul piano scientifico, una prima chiave di lettura è che l'*accountability* diventa uno strumento funzionale a sostenere le molteplici funzioni della responsabilità civile. Al di là della funzione compensativa, la dottrina sottolinea la dimensione preventiva della responsabilità, in cui l'assunzione di obblighi espliciti di rendicontazione e di verifica conduce a una riduzione dei comportamenti dannosi (Teubner; Beckers 2023). In questo quadro, la spiegabilità algoritmica funge da leva di trasparenza, ponendo in essere un deterrente contro configurazioni altrimenti opache dei sistemi di IA.

La seconda funzione rilevante è quella sanzionatoria, ossia la possibilità di comminare conseguenze negative (risarcitorie, ma anche reputazionali) a chi non abbia rispettato i requisiti di spiegabilità e di *accountability*. La prospettiva è che un soggetto che ometta di predisporre adeguati meccanismi di controllo e di spiegazione degli algoritmi risponda civilmente di eventuali danni, potendo essere tenuto a rimborsare le vittime. In tale scenario, l'*AI Act* rafforza la posizione dei danneggiati, in quanto l'inadempimento agli obblighi di chiarimento può costituire un argomento a sfavore del responsabile.

La terza funzione della responsabilità civile, di natura ripristinatoria o compensativa, viene arricchita dall'adozione del principio di *accountability*: l'adozione di procedure di auditing continuo, di risk assessment periodici e la predisposizione di spiegazioni contestualizzate favoriscono un rapido accertamento della causa del danno (Faleiros Júnior, 2024). Tale meccanismo agevola la dimostrazione del nesso di causalità e, di conseguenza, facilita l'ottenimento di un risarcimento equilibrato, adeguando la misura del danno all'effettiva responsabilità del soggetto coinvolto.

Sotto il profilo preventivo, la responsabilità civile esprime la sua efficacia quando l'obbligo di *accountability* si traduce in protocolli di progettazione e sviluppo di algoritmi conformi a standard di sicurezza, affidabilità e correttezza. Come rileva Comandé (2018), l'inserimento di procedure di "*explainability by design*" e "*accountability by design*" nelle fasi iniziali di sviluppo rappresenta un baluardo contro possibili malfunzionamenti o bias che sfocerebbero in responsabilità civile.

La quarta funzione, sovente trascurata, è quella deterrente o pedagogica: attraverso sanzioni esemplari e l'obbligo di fornire spiegazioni pubbliche, la responsabilità civile educa gli operatori a implementare in modo rigoroso il principio di *accountability*. In effetti, il rischio di danni d'immagine derivanti dalla scarsa trasparenza diviene un incentivo potente a investire in metodologie di IA che siano spiegabili e monitorabili (Mökander et al. 2022).

Da un punto di vista metodologico, la correlazione fra spiegabilità e *accountability* implica una rivisitazione dei canoni classici di responsabilità per colpa, che spesso faticano a inquadrare algoritmi complessi. L'adozione di regole presumptive, l'inversione dell'onere della prova o l'introduzione di forme di responsabilità oggettiva possono trovare una giustificazione ulteriore qualora l'operatore non ottemperi all'obbligo di rendere il sistema comprensibile e tracciabile. In ciò, si scorge il potenziale di un approccio multistrato, in cui la responsabilità dell'operatore si accompagna a quella del produttore e del programmatore (Comandé 2019; 2018).

Un aspetto centrale emerso nella letteratura più recente concerne il ruolo dei fondi di garanzia e delle assicurazioni specializzate, strutturate attorno a regole di *accountability*. Secondo Teubner e Beckers (2023), la creazione di schemi assicurativi calibrati sui rischi algoritmici favorisce la ripartizione dei costi di risarcimento e, al contempo, incentiva gli stakeholder ad adottare misure efficaci di risk management. La possibilità di verificare ex post l'effettiva implementazione di procedure di spiegabilità è, in tali contesti, un requisito essenziale per stabilire premi assicurativi più o meno vantaggiosi.

Dal punto di vista funzionale, la responsabilità civile preventiva acquisisce ulteriore rilevanza: la predisposizione di processi di *explainability* e di report periodici previene l'emergere di violazioni gravi, indirizzando i soggetti verso una condotta diligente e conforme agli standard di buona tecnica. Analogamente, l'*AI Act* incoraggia lo sviluppo di codici di condotta che promuovano l'*accountability* come elemento fondante di una IA affidabile, indicando dettagliatamente le procedure di audit e i requisiti di trasparenza (Nicholas 2020; Pasquale 2022).

In conclusione, la sinergia tra spiegabilità e responsabilità civile, mediata dal principio di *accountability* codificato nell'*AI Act*, costituisce un paradigma di riferimento per una regolamentazione equilibrata e orientata alla tutela effettiva degli interessi coinvolti.

L'approccio scientifico e normativo converge nel definire la responsabilità civile come strumento di prevenzione, riparazione e incoraggiamento di comportamenti virtuosi in ambito algoritmico, in un contesto in cui la rapida evoluzione tecnologica richiede soluzioni flessibili ma rigorose. Tale prospettiva rafforza la comprensione collettiva delle sfide poste dai sistemi di IA e conferma il ruolo cruciale della responsabilità civile e dell'*accountability* nell'assicurare un impiego etico e sostenibile dell'intelligenza artificiale.

#### **4. PROSPETTIVE FUTURE TRA SPIEGABILITÀ E INNOVAZIONE**

Nel dibattito contemporaneo sull'intelligenza artificiale (IA), la tensione fra la necessità di garantire la spiegabilità algoritmica e l'esigenza di preservare spazi di innovazione rappresenta uno dei nodi critici più rilevanti (Nicholas 2020). Da un lato, i principi di *accountability* e trasparenza impongono la creazione di sistemi intelligibili e verificabili; dall'altro, la ricerca e lo sviluppo di nuovi modelli di IA si basano spesso su tecnologie di complessità crescente e metodologie di apprendimento profondo che sfuggono a rendicontazioni immediate.

In quest'ottica, la spiegabilità non va intesa soltanto come un mero obbligo di rivelazione del codice sorgente o delle logiche di funzionamento. Piuttosto, essa assume la valenza di uno strumento di equilibrio tra la protezione dei diritti fondamentali e la libertà di sperimentazione tecnologica (Wachter 2022). Così facendo, il discorso sulla tutela del singolo si intreccia con la promozione di un contesto in cui le imprese possano continuare a innovare, senza temere eccessivi vincoli normativi.

Dal punto di vista pratico, il futuro delle interfacce spiegabili si lega strettamente alle esigenze del mercato e delle autorità regolatorie. Le imprese dovranno investire in soluzioni in grado di offrire report sintetici e comprensibili, integrando modelli autoesplicativi e meccanismi di auditing indipendente (Coeckelbergh 2020). Allo stesso tempo, le normative come l'AI Act europeo evidenziano un impegno a definire standard internazionali che uniscano chiarezza e tutela dell'innovazione.

La prospettiva critica, tuttavia, mette in luce il rischio di un approccio "dualista", che veda da un lato la spiegabilità come ostacolo all'evoluzione tecnologica e, dall'altro, l'innovazione come potenziale fonte di opacità (Pasquale 2022). Invece, un'analisi più approfondita rivela che la compatibilità tra i due aspetti è non solo possibile, ma auspicabile:

lo sviluppo di metodologie spiegabili può diventare un fattore competitivo per le imprese, incrementando la fiducia degli utenti e la sicurezza giuridica.

Un ulteriore elemento di riflessione riguarda la cooperazione internazionale e la costruzione di un ecosistema di conoscenza condivisa. La messa a punto di linee guida globali, sostenute da organismi sovranazionali, può infatti contribuire a creare un terreno fertile per lo scambio di best practice (Mökander et al. 2022). In tal modo, spiegabilità e innovazione si rafforzano a vicenda, generando un circolo virtuoso in cui la ricerca avanzata incontra la domanda di responsabilità.

In definitiva, la definizione di “spiegabilità innovativa” implica la capacità di ideare strumenti di interpretazione dei modelli di IA che non blocchino, ma anzi incentivino la creatività tecnica. Tale concetto ridefinisce il paradigma del rapporto tra legalità e crescita economica, offrendo prospettive per un futuro in cui sia possibile coniugare progresso, tutela del consumatore e garanzia di *accountability* (Comandé 2019).

L'importanza di questi due temi si relaziona in modo diretto con la nozione di *accountability* e il suo contributo allo sviluppo della responsabilità civile in tutte le sue funzioni, in particolare quella preventiva. Già diversi autori hanno messo in luce come la condivisione di obblighi di trasparenza e tracciabilità sia un antidoto contro potenziali abusi o danni derivanti dall'utilizzo di algoritmi complessi (Teubner; Beckers 2023).

Con riferimento all'innovazione, la *accountability* offre la possibilità di mantenere elevati standard di sicurezza, pur consentendo l'elaborazione di soluzioni tecnologiche avanzate. In questa logica, la responsabilità civile non opera esclusivamente sul versante ex post – ossia risarcitorio – ma si estende a una sfera ex ante, in cui i soggetti coinvolti sono incentivati a prevenire attivamente i rischi (Hildebrandt 2018). La funzione preventiva diventa così protagonista.

Dal punto di vista metodologico, l'adozione di pratiche di audit e di monitoraggio continuo dei sistemi di IA consente di individuare e correggere errori prima che producano effetti dannosi. Ciò si traduce in una responsabilità civile che funge da stimolo alla qualità dei prodotti, contribuendo a un contesto competitivo sano e orientato alla protezione degli utenti (Pasquale 2017).

In questo scenario, l'innovazione non risulta frenata, bensì qualificata: la complessità degli algoritmi diventa un elemento di valore nel momento in cui è accompagnata

da procedure di valutazione del rischio e da un corpus regolatorio che definisce chiaramente ruoli e obblighi di rendicontazione (Faleiros Jr 2024). In tal senso, *accountability* significa anche predisporre strutture interne all'azienda capaci di gestire eventuali problematiche in modo efficace e tempestivo.

Relativamente alla funzione preventiva della responsabilità civile, occorre porre l'accento sulla possibilità di sviluppare forme di assicurazione o fondi di garanzia specifici, che incentivino gli operatori a dotarsi di sistemi spiegabili e controllabili (Comandé 2018). Tale meccanismo, unito a un costante dialogo con le autorità regolatorie, favorisce l'adozione di buone pratiche e riduce il rischio di controversie.

Il passaggio da una responsabilità puramente compensativa a una "responsabilità strategica" investe anche il profilo pedagogico: le imprese e i progettisti di IA, consapevoli delle possibili conseguenze negative, mirano a conformarsi volontariamente a standard di *accountability* elevati. In tal modo, la prevenzione del danno diviene elemento integrante del processo stesso di innovazione, producendo un effetto di miglioramento costante sulla qualità degli algoritmi (Nicholas 2020).

Ulteriori sviluppi potrebbero riguardare la creazione di consorzi pubblico-privati per la ricerca sull'IA spiegabile, finalizzati a elaborare protocolli condivisi e aggiornabili secondo l'evoluzione delle tecnologie. In quest'ottica, *accountability* e spiegabilità non solo prevengono rischi, ma consentono di accumulare conoscenze utili a progettare sistemi ancora più efficaci e sicuri. La responsabilità civile si conferma così un volano di progresso.

Sotto il profilo giuridico, l'integrazione di *accountability* e spiegabilità potrebbe tradursi in revisioni sostanziali dei codici di condotta e dei regolamenti interni alle organizzazioni, promuovendo una cultura della responsabilità diffusa in tutti i livelli decisionali (Comandé 2019). Tale evoluzione non avverrà senza ostacoli, ma la spinta verso la trasparenza, combinata con incentivi economici e reputazionali, appare un fattore determinante per la resilienza dei mercati digitali.

In conclusione, le prospettive future tra spiegabilità e innovazione delineano un percorso in cui la responsabilità civile diventa un perno regolatorio non soltanto per la riparazione dei danni, ma anche per la promozione di una crescita tecnologica consapevole. Il rafforzamento della funzione preventiva, mediato dall'*accountability*, offre una chiave di lettura nuova, in cui le sfide dell'opacità algoritmica possono essere trasformate in

opportunità di miglioramento condiviso. Il risultato auspicato è un sistema equilibrato, in cui la tutela degli utenti e il progresso tecnologico si sostengono a vicenda, consolidando la fiducia nell'IA e la sua adozione responsabile.

## 5. CONCLUSIONE

L'analisi condotta ha messo in luce la funzione centrale della spiegabilità algoritmica nel fondare un quadro regolatorio che sia al contempo solido e duttile, in cui le istanze di tutela dei diritti fondamentali si sposino con l'esigenza di promuovere l'innovazione. La saldatura tra il principio di *accountability* e l'assetto della responsabilità civile risulta, in questa prospettiva, il fulcro di un paradigma evoluto, in grado di integrare modelli di previsione, prevenzione e riparazione del danno.

Le tensioni presenti tra trasparenza, tutela concorrenziale e sovranità tecnologica implicano un ripensamento delle categorie classiche del diritto della responsabilità. Tale ripensamento deve valorizzare l'idea di responsabilità esplicativa, che includa un dovere di chiarezza e giustificazione, in contrapposizione a un mero obbligo informativo, incapace di aggredire la complessità dei sistemi di IA.

L'evidente ricaduta preventiva di un simile approccio si manifesta nel rafforzamento della fiducia collettiva verso tecnologie sofisticate, riducendo l'opacità dell'"algoritmo oscuro" e incoraggiando l'adozione di protocolli condivisi, in cui la diligenza si estende dalla fase di progettazione fino all'implementazione e al monitoraggio costante.

La parallela necessità di tutela dei consumatori e degli utenti trova compimento nell'applicazione di modelli ibridi, capaci di bilanciare correttamente i principi di responsabilità per colpa, i regimi di responsabilità oggettiva e la supervisione pubblica mediante autorità specializzate. In tale contesto, l'*AI Act* europeo funge da riferimento normativo di primaria importanza.

La letteratura di settore, attraverso i contributi di diversi autori, evidenzia come l'inclusione di dispositivi di *accountability* e spiegabilità sin dalla fase di design incrementi la robustezza legale dei sistemi IA, prefigurando un panorama in cui la responsabilità sia condivisa fra soggetti eterogenei, ma sempre ancorata a standard di condotta chiaramente definiti.

Le dinamiche di risk management trovano concreta attuazione in questo modello, poiché l'adozione di valutazioni d'impatto e la conduzione di audit periodici consentono di individuare anomalie e lacune, anticipando l'insorgere di potenziali danni e orientando le politiche assicurative. Ciò favorisce un'ampia sinergia fra imprese, cittadini e pubbliche autorità.

Sul piano giuridico, l'approccio multilayered, sostenuto da svariati studiosi, completa la cornice della responsabilità civile attraverso un intelligente intreccio di principi generali e regole presuntive, favorendo una distribuzione del rischio e dei relativi costi in modo proporzionale alla partecipazione di ciascun operatore.

L'assetto interdisciplinare che emerge da questa analisi, incardinato sui concetti di trasparenza, spiegabilità e *accountability*, conferma la necessità di un dialogo costante fra ingegneri, giuristi e policy-maker, affinché si consolidi un ecosistema di IA eticamente e giuridicamente sostenibile, in cui il danno venga prevenuto prima ancora che compensato.

La formulazione di buone pratiche e codici di condotta, ispirati al principio di spiegabilità, contribuisce non solo a regolare la complessità, ma anche a trasformarla in opportunità di innovazione responsabile e di crescita economica. In questo modo, la responsabilità civile si arricchisce di una missione di promozione del benessere collettivo.

Si può concludere, dunque, che l'integrazione fra spiegabilità e *accountability* costituisce uno snodo ineludibile per il futuro della responsabilità civile in materia di intelligenza artificiale. Solo un sistema che sappia implementare efficacemente tali istanze potrà garantire il necessario equilibrio fra protezione dei diritti, slancio tecnologico e giustizia sostanziale.

## **BIBLIOGRAFIA**

COECKELBERGH, M. Artificial Intelligence, responsibility attribution, and a relational justification of explainability. **Science and Engineering Ethics**, v. 26, n. 4, p. 2051–2068, 2020.

COMANDÉ, G. Intelligenza artificiale e responsabilità tra liability e accountability: Il carattere trasformativo dell'IA e il problema della responsabilità. In: NUZZO, A.; OLIVIERI, G. (a cura di). **Analisi Giuridica dell'Economia – Studi e Discussioni sul Diritto dell'Impresa**. Bologna: Il Mulino, 2019.

COMANDÉ, G. Multilayered (Accountable) Liability for Artificial Intelligence. In: LOHSSE, S.; SCHULZE, R.; STAUDENMAYER, D. (eds.). **Liability for Robotics and in**

**the Internet of Things.** Münster Colloquia on EU Law and the Digital Economy. Baden-Baden: Nomos, 2018.

FALEIROS JR, J. L. M. Algorithmic Torts: The Role of Explainability. **Brazilian Journal of Law, Technology and Innovation**, v. 2, n. 1, p. 210–230, 2024.

HILDEBRANDT, M. Algorithmic regulation and the rule of law. **Philosophical Transactions of the Royal Society**, v. 376, n. 2128, 2018.

MÖKANDER, J.; JUNEJA, P.; WATSON, D. S.; FLORIDI, L. The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: what can they learn from each other? **Minds and Machines**, v. 32, p. 751–758, 2022.

NICHOLAS, G. Explaining Algorithmic Decisions. **Georgetown Law Technology Review**, v. 4, n. 2, p. 711–730, 2020.

PASQUALE, F. The US Algorithmic Accountability Act of 2022. **Minds and Machines**, 2022.

TEUBNER, G.; BECKERS, A. Responsibility for Algorithmic Misconduct: Unity or Fragmentation of Liability Regimes? **Yale Journal of Law & Technology**, Special Issue, p. 76, 2023.

WACHTER, S. The Theory of Artificial Immutability: Protecting Algorithmic Groups under Anti-Discrimination Law. **Tulane Law Review**, 2022–2023 (in corso di pubblicazione).