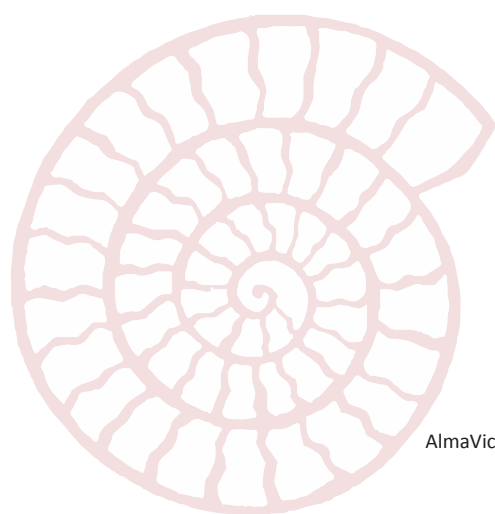




Cooperation and Artificial Intelligence: A Critical Review across Governance, Ethics, and Democracy

Cooperazione e Intelligenza Artificiale: Una revisione critica tra Governance, Etica e Democrazia

Elisabetta Miraglio

AlmaVico; Alma Mater Università di Bologna (Italy); e.miraglio@coode.it – <https://orcid.org/0009-0008-1490-8454>

Luca Ferrari

Department of Education Studies “G. M. Bertin”; Alma Mater Università di Bologna (Italy); l.ferrari@unibo.it

<https://orcid.org/0000-0001-7853-449X>

Claudia Beatrice Elena Iormetti

Head of Social Innovation; Open Group Coop. Soc. (Italy); claudia.iormetti@opengroup.eu

<https://orcid.org/0009-0005-0060-5119>



DOUBLE BLIND PEER REVIEW

ABSTRACT

This paper offers a scoping review of how cooperation and artificial intelligence (AI) intersect, examining the ways cooperative principles can inform the design, governance, and ethical development of intelligent technologies. Drawing on a systematic analysis of 70 academic contributions published between 2016 and 2025, the study identifies four main research areas: Cooperative AI and social dilemmas, cooperative governance of AI, Artificial Intelligence for Ethical Cooperation (AI4EC), and Democratic AI. The findings highlight that cooperation provides a transversal framework – technical, political, and educational – capable of promoting more inclusive, participatory, and ethically grounded approaches to digital transformation. In this perspective, education is not limited to formal instructional settings but extends to processes of collective learning, capacity-building, and socio-technical awareness involving citizens, educators, and organizations. By mapping research gaps and emerging trends, the article argues that placing cooperation at the heart of AI development means reaffirming the collective role of these actors in shaping technologies that serve democratic and social well-being.

Il contributo propone una scoping review delle relazioni tra cooperazione e intelligenza artificiale (IA), analizzando come i principi cooperativi possano orientare la progettazione, la governance e lo sviluppo etico delle tecnologie intelligenti. A partire da un'analisi sistematica di 70 contributi scientifici pubblicati tra il 2016 e il 2025, il lavoro individua quattro principali aree di ricerca: Cooperative AI e dilemmi sociali, governance cooperativa dell'IA, Artificial Intelligence for Ethical Cooperation (AI4EC) e Democratic AI. I risultati mostrano come la cooperazione offra un quadro trasversale – tecnico, politico ed educativo – in grado di promuovere approcci inclusivi, partecipativi ed eticamente fondati alla trasformazione digitale. In questa prospettiva, la dimensione educativa non si limita ai contesti formali di istruzione, ma include processi diffusi di apprendimento collettivo, costruzione di competenze e consapevolezza socio-tecnica che coinvolgono cittadini, educatori e organizzazioni. Mettendo in luce tendenze e vuoti di ricerca, l'articolo sostiene che riportare la cooperazione al centro del dibattito sull'IA significa riaffermare il ruolo collettivo di questi attori nella costruzione di tecnologie orientate al benessere democratico e sociale.

KEYWORDS

Cooperative artificial intelligence, AI governance and democracy, Ethical cooperation, Data cooperatives and digital commons, Collective intelligence in education

Intelligenza artificiale cooperativa, Governance e democrazia dell'IA, Cooperazione etica, Cooperative di dati e beni comuni digitali, Intelligenza collettiva nell'educazione

Citation: Miraglio, E., Ferrari, L., & Iormetti, C.B.E. (2026). Cooperation and Artificial Intelligence: A Critical Review across Governance, Ethics, and Democracy. *Formazione & insegnamento*, 24(2), 182-198. https://doi.org/10.7346/-fei-XXIV-02-26_19

Copyright: © 2026 Author(s).

License: Attribution 4.0 International (CC BY 4.0).

Conflicts of interest: The Authors declare no conflicts of interest.

Acknowledgments: Questo articolo nasce dal lavoro congiunto tra il Dipartimento di Scienze dell'Educazione “G.M. Bertin” dell'Alma Mater Università di Bologna e Legacoop Bologna, con il supporto di AlmaVico e della Fondazione PICO. Gli autori ringraziano i colleghi e le colleghe che hanno partecipato ai percorsi formativi e di riflessione collettiva nell'ambito del progetto “Cooperative Digital Education - COO.DE”.

Editorial Note:

Authorship: All authors contributed equally to the conceptualisation and writing of the manuscript.

DOI: https://doi.org/10.7346/-fei-XXIV-02-26_19

Submitted: October 7, 2025 • **Accepted:** May 22, 2026 • **Published on-line:** June 3, 2026

Pensa MultiMedia: ISSN 2279-7505 (online)

1. Introduzione

1.1. Contesto

Negli ultimi decenni, le innovazioni digitali più rilevanti si sono sviluppate all'interno di contesti caratterizzati da forme di collaborazione e interazione tra soggetti, più che come esito del lavoro isolato di singoli individui (Ferrari & Ingrosso, 2024). In questa prospettiva, il concetto di intelligenza collettiva (Lévy, 1997) offre una chiave interpretativa particolarmente utile per leggere i processi contemporanei di produzione e condivisione della conoscenza: una forma di intelligenza distribuita che trova nella rete digitale il proprio ambiente di sviluppo e che oggi assume un ruolo sempre più centrale nell'affrontare le sfide poste dall'intelligenza artificiale (IA). Come sottolineano Malone e Bernstein (2015), tali sfide richiedono non solo avanzamenti tecnologici, ma anche la capacità di sviluppare modelli efficaci di coordinamento e valorizzazione dell'intelligenza umana collettiva.

In questo scenario, i recenti sviluppi dei Large Language Models (LLM) hanno contribuito ad ampliare il dibattito sul ruolo dell'IA in diversi ambiti dell'esperienza umana, tra cui quello educativo (Kasneci et al., 2023; Panciroli & Rivoltella, 2023; Shi et al., 2026). In particolare, nel campo dell'educazione, l'IA non si configura più soltanto come strumento a supporto della didattica o come insieme di competenze tecniche da acquisire, ma come un ambito che sollecita una riflessione più ampia, di natura pedagogica, culturale ed etica. In questa direzione, si evidenzia la necessità di formare professionisti dell'educazione capaci di comprendere e governare i processi decisionali legati allo sviluppo e all'uso delle tecnologie digitali. La mancata assunzione di questa responsabilità rischia di produrre nuove forme di esclusione e marginalità, che riguardano tanto gli studenti quanto gli stessi educatori (Ranieri, 2020).

Negli ultimi anni, a partire da queste trasformazioni, sono stati sviluppati diversi framework internazionali per sostenere lo sviluppo delle competenze digitali degli educatori. Tra i principali, il DigCompEdu (Redecker & Punie, 2017) e le più recenti proposte dell'UNESCO (Fengchun & Mutlu, 2024) contribuiscono a delineare un quadro di riferimento che pone al centro la persona e la comunità professionale. Questo approccio è guidato da una solida dimensione etica e orienta la scelta e la progettazione di proposte pedagogico-didattiche mediate e potenziate dall'uso dei sistemi di intelligenza artificiale (Ellerani & Ferrari, 2024).

In questa direzione diversi contributi nel settore pedagogico sottolineano come la costruzione di connessioni significative tra IA ed educazione richieda un processo di contaminazione reciproca tra ricerca e pratica, fondato su una negoziazione condivisa di modelli, valori e pratiche (Panciroli & Macaudo, 2021; Luckin & Cukurova, 2019). L'educazione all'IA si colloca, infatti, all'intersezione tra dimensioni tecniche e questioni più ampie legate alla democrazia, alla giustizia sociale e all'inclusione.

Parallelamente, nel campo degli studi sull'IA si sta affermando un filone di ricerca interdisciplinare, con una crescente attenzione alle ricadute educative, noto come Cooperative AI. Tale prospettiva è orientata allo sviluppo di sistemi artificiali capaci di collaborare efficacemente tra loro e con gli esseri umani, promuovendo dinamiche di fiducia, equità e coordinamento (Dafoe et al., 2021). Se da un lato questo ambito si sviluppa prevalentemente all'interno delle scienze computazionali, dall'altro presenta interessanti punti di contatto con il patrimonio teorico e pratico del movimento cooperativo, che da tempo sperimenta forme di governance condivisa e partecipazione. Evidentemente, l'incontro tra

educazione all'IA e cultura cooperativa consente di aprire nuove piste di lavoro per la progettazione di percorsi formativi che non si limitino alla comprensione tecnica degli algoritmi, ma che favoriscano la capacità di interpretare e governare le tecnologie in chiave sociale e partecipativa (Dafoe et al., 2021).

Alla luce della necessità di ripensare i percorsi formativi rivolti alle figure educative, sia in ambito scolastico sia extra-scolastico, è stato sviluppato il corso di alta formazione "Cooperative Digital Education - Coo.de". A partire dal 2020, una collaborazione tra università e impresa – sostenuta da sostenuta da Legacoop Bologna, con il supporto di AlmaVico e della Fondazione PICO – ha dato origine a questo percorso, configurandosi come uno spazio di ricerca-formazione in cui progettazione didattica, riflessione critica e produzione di conoscenza si intrecciano in modo sistematico (Ferrari & Ingrosso, 2024). Tale esperienza, riconducibile alla tradizione della ricerca-azione partecipativa (Reason & Bradbury, 2008), ha consentito di sviluppare percorsi di formazione continua e moduli formativi dedicati al digitale e all'IA, organizzati attorno ad alcune aree tematiche – consapevolezza digitale, cooperazione e inclusione, comunicazione e documentazione, salute e benessere – con particolare attenzione alle competenze mediaeducative (Buckingham, 2019; Rivoltella, 2017, Pasta & Rondonotti, 2020).

È proprio all'interno di questo contesto che è emersa l'esigenza di sistematizzare la letteratura relativa al rapporto tra IA e cooperazione in ambito educativo, mantenendo una postura di ricerca interdisciplinare. A fronte di una crescente attenzione verso questi temi, appare infatti ancora limitata una mappatura sistematica dei contributi che mettono in relazione IA, cooperazione e processi educativi. Da questa esigenza prende forma la presente *scoping review*, che si propone di analizzare la letteratura prodotta negli ultimi dieci anni, con l'obiettivo di individuare le principali traiettorie di ricerca e le aree ancora poco esplorate.

Il lavoro si pone un duplice obiettivo: da un lato, raccogliere e sistematizzare i contributi esistenti; dall'altro, contribuire alla definizione di possibili agende di ricerca orientate allo sviluppo di un'IA intesa come fenomeno sociale, oltre che tecnologico, e coerente con principi di partecipazione, equità e cooperazione. I risultati della *review* si intrecciano con il ripensamento di alcuni moduli del corso "Cooperative Digital Education - Coo.de", con l'intento di costruire percorsi di *media education* che pongano al centro la persona e la sua capacità agentiva (Sen, 1999; Biesta & Tedder, 2007; Margiotta, 2018; Fabbri, 2025), all'interno di una visione della tecnologia orientata in senso democratico e cooperativo.

1.1. Obiettivi

Le domande di ricerca della presente *scoping review* sono state strutturate utilizzando il framework PCC (Population, Concept and Context - Popolazione, Concetto e Contesto) (Peters et al., 2020; Pollock et al., 2023), al fine di garantire un approccio sistematico e completo. In linea con le linee guida del Joanna Briggs Institute (JBI), questo framework ha consentito di affinare gli obiettivi dello studio e di stabilire criteri di eleggibilità chiari, offrendo una base strutturata per la definizione del focus della revisione, l'elaborazione della strategia di ricerca e la determinazione dei criteri di inclusione ed esclusione, in un'ottica di trasparenza e replicabilità metodologica.

Il framework PCC si rivela particolarmente adatto alle revisioni di *scoping*, poiché consente una formulazione delle domande di ricerca al tempo stesso strutturata e flessibile,

caratteristica essenziale per mappare ambiti ampi, emergenti ed eterogenei come quello dell'etica e della governance cooperativa dell'intelligenza artificiale (Arksey & O'Malley, 2005). A differenza delle revisioni sistematiche, che richiedono quesiti circoscritti e criteri di inclusione stringenti, le revisioni di *scoping* mirano a restituire l'ampiezza della letteratura esistente, a individuare eventuali lacune e a orientare sviluppi futuri della ricerca (Westphal et al., 2021).

In questo lavoro, tale approccio assume anche una valenza educativa: la mappatura del campo non è finalizzata unicamente alla sistematizzazione teorica, ma contribuisce a costruire un quadro di riferimento utile per informare e arricchire percorsi educativi sull'intelligenza artificiale, intesi non solo in chiave didattica, ma come spazi di formazione critica, partecipazione e consapevolezza socio-tecnica.

In questa prospettiva, il framework PCC supporta l'analisi delineando in modo esplicito le dimensioni chiave dello studio, sintetizzate nella Tabella 1. La categorizzazione della do-

manda di ricerca in termini di popolazione, concetto e contesto ha garantito coerenza metodologica e chiarezza nei criteri di eleggibilità, favorendo al contempo l'inclusione di una pluralità di attori, pratiche e contesti. Questi, pur non collocandosi sempre esplicitamente nel campo educativo, contribuiscono a definire le condizioni culturali, organizzative e politiche entro cui si sviluppano i processi di apprendimento e la costruzione di competenze legate all'intelligenza artificiale.

La presente revisione di *scoping* ha quindi mappato il panorama della ricerca, analizzato il ruolo delle dinamiche cooperative nella progettazione e nell'uso dell'intelligenza artificiale e identificato le principali lacune conoscitive, con l'obiettivo di offrire una base interpretativa utile allo sviluppo di pratiche educative più consapevoli e situate.

Le seguenti domande di ricerca sono state formulate sulla base del framework PCC:

Componente PCC	Descrizione	Operazionalizzazione nello studio
Population	Attori e sistemi coinvolti	Ricercatori, sviluppatori di sistemi di IA, organizzazioni cooperative, responsabili politici ed educatori coinvolti nella progettazione, nella governance e nell'uso dell'intelligenza artificiale.
Concept	Concetto	Dinamiche cooperative e strategie di governance cooperativa applicate alla progettazione e all'uso dell'IA, con riferimento a Cooperative IA, governance democratica, etica della cooperazione e Democratic IA.
Context	Contesto di analisi	Contesti organizzativi, educativi, istituzionali e socio-tecnici in cui l'IA viene sviluppata e adottata; diversi quadri normativi, culturali e disciplinari a livello nazionale e internazionale.

Tabella 1. Applicazione del framework PCC

La domanda di ricerca principale che orienta la presente revisione è la seguente: in che modo le dinamiche cooperative e le strategie di governance cooperativa possono orientare la progettazione e l'uso dell'intelligenza artificiale? A partire da essa, sono state formulate quattro domande specifiche:

- **RQ1:** In che modo i sistemi di intelligenza artificiale vengono progettati per affrontare dilemmi sociali e promuovere comportamenti cooperativi tra agenti artificiali e tra agenti artificiali ed esseri umani? Questa domanda esplora i modelli teorici, le architetture algoritmiche e i contesti applicativi in cui la Cooperative AI viene mobilitata per gestire situazioni di interdipendenza, conflitto di interessi e coordinamento collettivo.
- **RQ2:** Quali modelli di governance cooperativa vengono proposti o sperimentati per orientare lo sviluppo, la regolazione e l'uso dei sistemi di intelligenza artificiale? Questa domanda analizza le strutture di partecipazione, le forme di controllo distribuito e i meccanismi decisionali condivisi che si ispirano ai principi del movimento cooperativo per ripensare la gestione delle tecnologie intelligenti.
- **RQ3:** In che modo il rapporto tra intelligenza artificiale e cooperazione si configura come bidirezionale sul piano etico? Questa domanda esamina, da un lato, le applicazioni di IA progettate per sostenere pratiche collaborative, rafforzare la fiducia tra attori umani e contribuire al raggiungimento di obiettivi collettivi nel rispetto dei principi di equità e responsabilità; dall'altro, indaga in che misura i valori e le pratiche proprie della tradizione

cooperativa, quali la reciprocità e la gestione condivisa, possano contribuire a orientare lo sviluppo di un'IA più etica, trasparente e socialmente responsabile.

- **RQ4:** Quali approcci teorici e pratici vengono sviluppati per rendere l'intelligenza artificiale più democratica, partecipativa e allineata ai valori della cittadinanza attiva? Questa domanda esplora le proposte volte a garantire che i sistemi di IA siano progettati e governati in modo inclusivo, trasparente e responsabile, con particolare attenzione al coinvolgimento dei cittadini, alla riduzione delle asimmetrie di potere e alla promozione della giustizia sociale.

2. Metodologia

2.1. Protocollo e criteri di eleggibilità

Il protocollo della presente *scoping review* è stato sviluppato preliminarmente in conformità con il framework PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) (Tricco et al., 2018), al fine di garantire trasparenza metodologica e coerenza nei criteri di eleggibilità, nella costruzione della strategia di ricerca, nei processi di estrazione dei dati e nella sintesi dei risultati.

La *scoping review* non è stata pre-registrata. Questa scelta è motivata dalla natura stessa dell'oggetto d'indagine, che si colloca all'intersezione tra etica dell'intelligenza artificiale, cooperazione, educazione e governance socio-tecnica, un ambito ancora emergente e non pienamente coperto da

piattaforme di registrazione come PROSPERO, tradizionalmente orientate verso studi in ambito clinico-sanitario. Inoltre, il carattere esplorativo dello studio, strutturato attraverso il framework PCC, ha richiesto un approccio flessibile e ciclico, in cui le categorie analitiche sono state progressivamente affinate durante il processo di analisi. In questo senso, una preregistrazione rigida avrebbe potuto limitare la capacità di adattare il disegno di ricerca alla complessità del fenomeno indagato.

Al fine di garantire comunque pertinenza e rigore metodologico, i criteri di inclusione ed esclusione sono stati definiti in modo esplicito e coerente con il framework PCC e con gli obiettivi della ricerca.

Sono stati inclusi articoli peer-reviewed e *conference proceedings* pubblicati negli ultimi dieci anni (2016–2025), periodo che coincide con l'emersione e il progressivo consolidamento del campo di ricerca. La delimitazione temporale consente di cogliere le traiettorie più recenti del dibattito scientifico, evitando al contempo di estendere l'analisi a contributi antecedenti all'emergere di tali filoni d'indagine.

Per quanto riguarda la lingua, sono stati inclusi esclusivamente contributi redatti in inglese. Tale scelta è motivata da ragioni sia empiriche sia metodologiche. Sul piano empirico, l'analisi preliminare delle banche dati selezionate ha mostrato che la quasi totalità della produzione scientifica pertinente al tema è pubblicata in inglese, che rappresenta la lingua franca della comunicazione accademica internazionale nei settori dell'informatica, dell'etica dell'IA e degli studi sulla governance tecnologica. Sul piano metodologico, la limitazione a una sola lingua riduce il rischio di errori interpretativi derivanti da traduzioni non standardizzate di terminologie tecniche e concettuali particolarmente sensibili al contesto, come quelle proprie del dibattito sulla Cooperative AI o sulla Democratic AI. Si riconosce tuttavia che questa scelta introduce un potenziale *bias* linguistico, che potrebbe comportare la sottorappresentazione di contributi significativi prodotti in altre lingue, in particolare in contesti nazionali con una forte tradizione cooperativa, come quelli dell'Europa continentale o dell'America Latina. Tale limitazione viene esplicitamente dichiarata come uno dei principali vincoli della presente revisione e costituisce al contempo un'indicazione per sviluppi futuri: un'estensione della ricerca a contributi in altre lingue rappresenta, infatti, una delle direzioni prioritarie per il proseguimento di questa linea di indagine.

Sono stati inoltre inclusi contributi su ambiti affini e trasversali al tema centrale, quali processi decisionali collettivi, modelli organizzativi cooperativi, *data commons* e democrazia tecnologica, a condizione che adottassero approcci qualitativi, teorici, mixed-methods o di review e che presentassero un'esplicita connessione con le dinamiche di governance o cooperazione nell'ambito dell'IA.

Sono stati invece esclusi capitoli di libri e contributi non sottoposti a peer-review, in quanto non garantiscono gli stessi standard di validazione scientifica. Sono stati altresì esclusi articoli esclusivamente tecnici su sistemi di IA privi di qualsiasi connessione con il tema della governance o della cooperazione, analisi di piattaforme digitali strutturate secondo logiche non cooperative, come i modelli tipici della *gig economy* tradizionale, riflessioni filosofiche generali sull'IA prive di implicazioni pratiche per la governance cooperativa, nonché studi duplicati o caratterizzati da una metodologia non trasparente o non ricostruibile.

2.2. Fonti informative e strategia di ricerca

La ricerca della letteratura è stata condotta su cinque database accademici selezionati per garantire una copertura interdisciplinare. ERIC (Education Resources Information Center) è stato utilizzato per reperire articoli sull'impatto sociale e formativo dell'IA in contesti cooperativi o educativi. Springer ha offerto accesso a numerosi contributi sul rapporto tra IA, cooperazione e scienze sociali, mentre SSRN (Social Science Research Network) ha permesso di includere contributi emergenti con attenzione alla dimensione sociopolitica ed etica dell'IA. Per coprire l'ambito tecnico e ingegneristico sono stati consultati IEEE Xplore e la ACM Digital Library.

L'inclusione di SSRN è stata motivata dalla natura ancora in consolidamento del campo di ricerca: trattandosi di un ambito emergente, tale database consente di intercettare *working papers* e studi in fase di sviluppo, spesso rilevanti per cogliere traiettorie teoriche innovative.

L'interrogazione dei database è stata effettuata nel periodo compreso tra febbraio 2025 e aprile 2025, al fine di includere i contributi più aggiornati disponibili al momento della raccolta dei dati.

È stata implementata una strategia di ricerca strutturata e replicabile, costruita attraverso la combinazione di parole chiave relative a intelligenza artificiale, cooperazione, governance ed educazione.

Le interrogazioni sono state effettuate nei diversi database adottando criteri coerenti ma adattati alle specificità tecniche di ciascuna piattaforma. In particolare:

- nei database Springer ed ERIC, le ricerche sono state condotte in modalità *"all fields"*, al fine di ampliare la copertura e intercettare contributi anche in contesti interdisciplinari;
- nei database SSRN, IEEE Xplore e ACM Digital Library, le stringhe sono state inserite utilizzando i filtri *title*, *abstract* e *keywords*, così da mantenere una maggiore precisione nel recupero delle fonti.

Le stringhe di ricerca sono state definite attraverso successivi affinamenti. In una fase iniziale, sono state testate query più articolate, ma queste hanno prodotto un numero molto limitato di risultati, escludendo contributi rilevanti per gli obiettivi della ricerca. Per questo motivo, si è scelto di adottare stringhe intenzionalmente più semplici e inclusive, limitando l'uso restrittivo degli operatori booleani (in particolare AND) al fine di privilegiare la sensibilità della revisione rispetto alla sua specificità (Tabella 2), coerentemente con l'obiettivo di mappatura della letteratura.

Stringhe utilizzate	Springer	ERIC	SSRN	IEEE Xplore	ACM Digital Library
"Cooperative AI"	39 risultati	0 risultati	4 risultati	5 risultati	7 risultati
"Artificial Intelligence" AND "Cooperative governance"	27 risultati	0 risultati	5 risultati	0 risultati	0 risultati
"Platform Cooperativism" AND "Artificial intelligence"	9 risultati	0 risultati	0 risultati	0 risultati	0 risultati
"Artificial Intelligence" AND "Democracy" AND "Cooperativism"	5 risultati	0 risultati	0 risultati	0 risultati	0 risultati
"Democratic AI"	18 risultati	0 risultati	0 risultati	0 risultati	1 risultato
"Data cooperatives" AND "Artificial intelligence"	12 risultati	0 risultati	3 risultati	1 risultato	0 risultati
"Data commons" AND "Data cooperatives"	7 risultati	0 risultati	0 risultati	0 risultati	0 risultati
"AI commons"	2 risultati	8 risultati	2 risultati	0 risultati	0 risultati
"AI curriculum" AND "Governance"	19 risultati	0 risultati	0 risultati	0 risultati	0 risultati
"AI Ethicist" AND "Skills"	11 risultati	0 risultati	0 risultati	0 risultati	0 risultati

Tabella 2. Stringhe di ricerca e database

Questa strategia ha permesso di bilanciare ampiezza e pertinenza, garantendo da un lato una copertura sufficientemente estesa della letteratura e dall'altro la possibilità di intercettare contributi rilevanti in un campo ancora in fase di definizione.

2.3. Selezione e analisi dei dati

La ricerca ha identificato complessivamente 185 record provenienti dai cinque database selezionati. Dopo la rimozione di 5 duplicati, sono stati considerati 180 studi unici per lo screening iniziale. In questa fase, basata sull'analisi di titolo, abstract e parole chiave, sono stati esclusi 71 articoli in quanto non pertinenti rispetto agli obiettivi della ricerca o non sottoposti a peer review.

I 109 studi rimanenti sono stati successivamente sottoposti a valutazione full-text, che ha portato all'esclusione di ulteriori 39 contributi per scarsa pertinenza rispetto al focus della *scoping review* o per un livello di tecnicità tale da non consentire un'analisi significativa in relazione alle dimensioni di governance e cooperazione. La revisione include pertanto un totale di 70 articoli.

L'intero processo di selezione è stato condotto seguendo una procedura in due fasi, screening preliminare e valutazione approfondita, ed è stato realizzato da più ricercatori in modo indipendente. Nei casi di disaccordo, le decisioni sono state discusse collettivamente fino al raggiungimento di un consenso. Il workflow completo del processo di selezione è rappresentato nella Figura 1 (diagramma PRISMA).

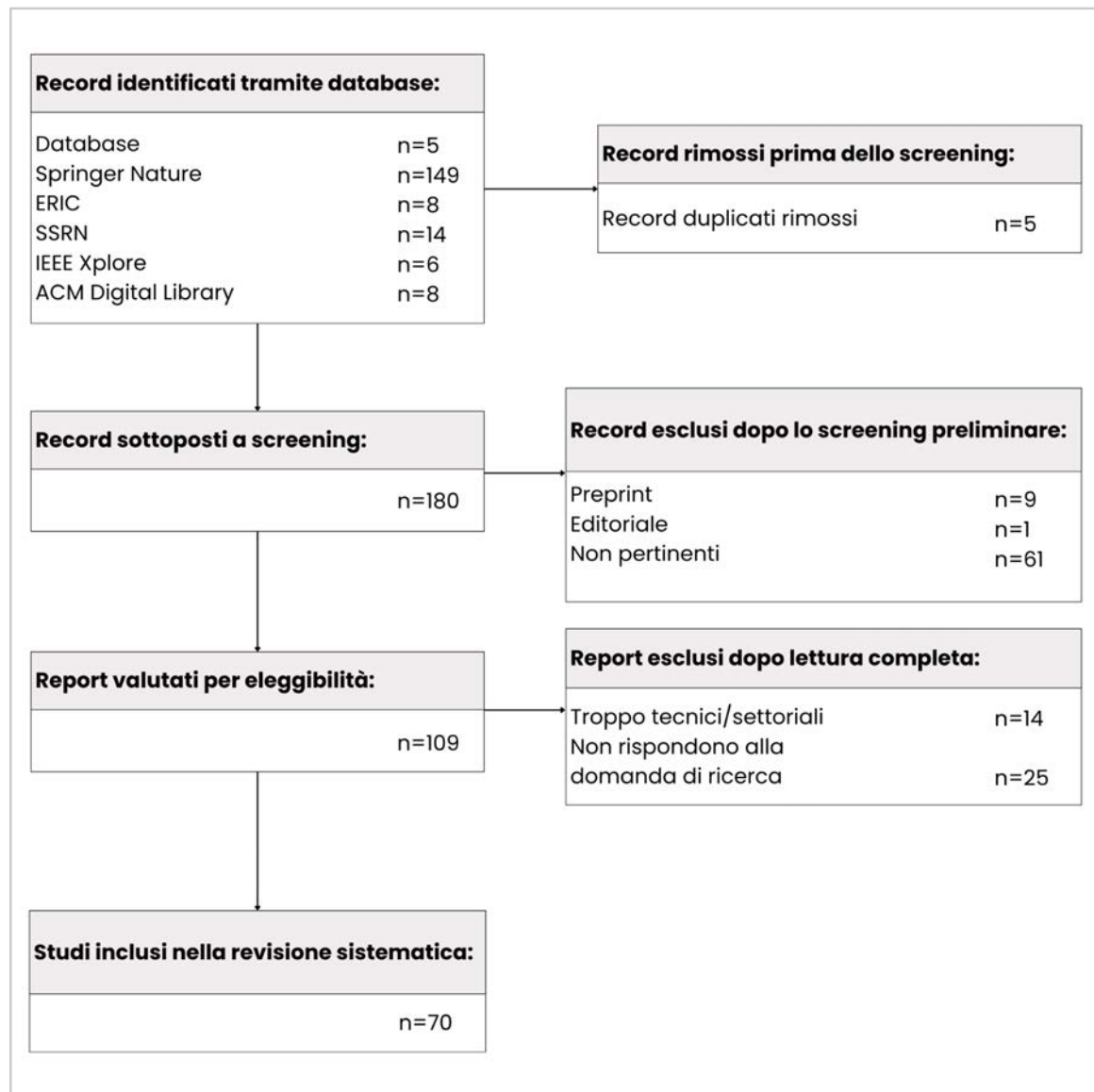


Figura 1. Diagramma di flusso PRISMA che illustra il processo di selezione, comprese le ragioni delle esclusioni e le fonti finali incluse

Per l'analisi dei dati è stato impiegato un modulo di raccolta strutturato, volto a garantire un'estrazione sistematica delle informazioni rilevanti. In questa prospettiva, gli studi esaminati non sono stati considerati soltanto per il loro contributo tecnico o teorico, ma anche per la loro più ampia valenza educativa: pur non appartenendo necessariamente alla didattica in senso stretto, offrono chiavi interpretative, modelli e pratiche capaci di arricchire i percorsi educativi sull'intelligenza artificiale in ambito cooperativo.

Le principali caratteristiche degli studi raccolte hanno incluso i dati bibliografici, il contesto e le specificità di ciascun contributo, l'approccio metodologico adottato e il tipo di apporto rispetto alle *research questions*. La struttura del modulo ha così consentito non solo di descrivere i lavori inclusi, ma anche di metterli a confronto in modo approfondito.

Per mantenere un equilibrio tra leggibilità del testo e trasparenza del processo analitico, i dati completi relativi agli studi inclusi e la relativa valutazione qualitativa sono stati resi disponibili in forma aperta in un'unica tabella online.¹

Il processo di analisi si è sviluppato in modo iterativo e riflessivo: le categorie interpretative non sono state rigidamente predefinite, ma progressivamente affinate attraverso la lettura approfondita degli studi e il confronto tra i ricercatori. Questo approccio ha permesso di adattare il framework analitico alla complessità e all'eterogeneità del materiale esaminato, favorendo l'emersione di connessioni, convergenze e tensioni tra i diversi contributi.

2.4. Valutazione della qualità degli studi inclusi

Non è stata condotta una valutazione formale della qualità degli studi inclusi, in linea con le indicazioni del framework PRISMA-ScR (Tricco et al., 2018). La presente *scoping review*, infatti, si propone di mappare il panorama della letteratura esistente piuttosto che di valutarne sistematicamente la validità o il rischio di *bias*. Considerata la natura interdiscipli-

1 Disponibile in forma cristallizzata come *Supplementary materials* al presente contributo (<https://ojs.pensamultimedia.it/index.php/siref/article->

[/view/8398/7707](view/8398/7707)). Disponibile anche in forma dinamica su piattaforma esterna (<https://docs.google.com/spreadsheets/d/1jzGFF7xyTgmFsGLg-pfiDmLlloBCj5Z9aoNFaL4-Jde4/edit?usp=sharing>).

nare ed emergente dell'oggetto di studio, collocato all'intersezione tra intelligenza artificiale, cooperazione, educazione e governance socio-tecnica, la scelta di privilegiare l'ampiezza ha consentito di includere contributi eterogenei per approcci, tradizioni disciplinari e livelli di maturità empirica.

L'adozione di criteri valutativi rigidi avrebbe rischiato di escludere studi preliminari, teorici o interdisciplinari particolarmente rilevanti in un campo ancora in fase di consolidamento, limitando la capacità della review di restituire una visione articolata e rappresentativa dello stato dell'arte.

A supporto dell'interpretazione dei risultati, è stata comunque condotta una valutazione qualitativa semplificata delle caratteristiche metodologiche degli studi inclusi, disponibile nella tabella online sopra richiamata insieme ai dati completi del corpus. Tale valutazione non ha avuto funzione escludente, ma interpretativa, ed è stata utilizzata per supportare una lettura critica e comparativa dei contributi in coerenza con la natura esplorativa della *scoping review*. In particolare, sono state considerate due dimensioni principali: (1) la trasparenza metodologica, intesa come grado di esplicitazione di protocolli, procedure, dimensioni campionarie e criteri analitici; e (2) i limiti degli studi, identificati sia sulla base di quanto esplicitamente dichiarato dagli autori, sia attraverso un processo di analisi interpretativa condotto dal gruppo di ricerca.

Nel complesso, gli studi analizzati presentano una marcata eterogeneità metodologica, coerente con la pluralità di approcci che ha caratterizzato il campo. I contributi di natura computazionale ed empirica rappresentano circa la metà del corpus e mostrano livelli differenziati di rigore. Tra i lavori più robusti si collocano esperimenti comportamentali su larga scala con oltre 7.000 partecipanti (von Schenk et al., 2023), simulazioni multi-agente con iterazioni multiple finalizzate alla stabilità statistica (Rachum et al., 2024; Dasgupta & Musolesi, 2025; Orzan et al., 2024), studi qualitativi longitudinali che coprono un arco temporale di vent'anni (Trinidad, 2025) e survey cross-nazionali con strumenti psicometrici validati (Okada et al., 2025). Accanto a questi, emergono contributi dal carattere più ricognitivo, tra cui studi su campioni ridotti (Veitch et al., 2024), prototipi ingegneristici non ancora validati in contesti urbani reali (Zhang et al., 2019) e simulazioni computazionali fortemente dipendenti dalle assunzioni modellistiche (Woods, 2025).

La componente teorico-normativa risulta altrettanto rilevante e include sia analisi filosofiche radicate in tradizioni consolidate (Benton, 2024; Cecchini et al., 2025; Steinhoff, 2024; Verdegem, 2022; Gandorfer, 2022), sia contributi di natura programmatica, come manifesti disciplinari e vision paper (Dafoe et al., 2021; Shi et al., 2025).

Questa varietà riflette una caratteristica strutturale del campo: la governance dell'IA e la Cooperative AI si configurano come ambiti in cui coesistono legittimamente ricerca empirica, modellizzazione formale e riflessione critica, ciascuna con un proprio valore epistemico.

Per quanto riguarda la trasparenza metodologica, i contributi computazionali e misti presentano generalmente una documentazione più esplicita in termini di dataset, protocolli di codifica, algoritmi di addestramento e metriche di valuta-

zione (Orzan et al., 2024; Orzan et al., 2025; von Schenk et al., 2023). Nei lavori teorici, invece, la solidità è stata valutata sulla base della coerenza argomentativa, dell'ancoraggio ai riferimenti teorici e della capacità di offrire un contributo significativo alle domande di ricerca. Tutti i paper inclusi sono stati considerati rilevanti, indipendentemente dal loro grado di rigore formale, in linea con l'approccio della *scoping review* che privilegia l'ampiezza e la rappresentatività del panorama disciplinare rispetto alla selezione per qualità metodologica (Tricco et al., 2018).

3. Risultati

3.1. Organizzazione tematica dei risultati

I risultati della revisione sono stati organizzati in una matrice tematica che ha permesso di mettere in evidenza le principali aree di indagine e le connessioni emergenti tra di esse. Questa articolazione non è stata definita a priori, ma è emersa progressivamente dall'analisi dei contributi in relazione alle quattro domande di ricerca formulate a partire dal framework PCC. Pertanto, gli assi individuati rappresentano una traduzione analitica delle traiettorie di risposta alla domanda principale – in che modo le dinamiche cooperative e le strategie di governance cooperativa possono orientare la progettazione e l'uso dell'intelligenza artificiale – e alle sue articolazioni specifiche.

A partire da questo processo interpretativo, la letteratura selezionata è stata ricondotta a quattro assi di ricerca:

- Cooperative AI e dilemmi sociali;
- Governance cooperativa dell'IA;
- Artificial Intelligence for Ethical Cooperation (AI4EC);
- Democratic AI.

I 70 contributi inclusi nella revisione sistematica sono stati classificati lungo questi assi. Tuttavia, la natura fortemente interdisciplinare del campo e le frequenti sovrapposizioni tematiche hanno reso evidente come molti studi non siano riconducibili a una singola categoria, ma attraversino più dimensioni contemporaneamente.

Per rendere esplicita e trasparente questa complessità, è stata costruita una matrice di classificazione degli studi per assi tematici (Tabella 3), che consente di visualizzare in modo sistematico la collocazione di ciascun contributo rispetto ai quattro ambiti di analisi. Questa matrice rappresenta uno strumento complementare alla sintesi narrativa: da un lato evita un'eccessiva semplificazione dei risultati, dall'altro permette di cogliere con maggiore precisione le sovrapposizioni, le ricorrenze e le specificità dei diversi studi, evidenziando il grado di concentrazione o dispersione delle evidenze all'interno dei vari assi.

La classificazione è stata condotta in modo indipendente da più ricercatori e, in caso di disaccordo, le assegnazioni sono state discusse fino al raggiungimento di un consenso.

<i>Studio (Autore, anno)</i>	<i>Cooperative AI</i>	<i>Governance cooperativa</i>	<i>AI4EC</i>	<i>Democratic AI</i>
Asa et al., 2023		X		
Benton, 2024				X
Blasimme et al., 2018		X		
Bruder et al., 2021	X			
Calzada, 2023		X		
Calzada, 2024a		X		
Calzada, 2024b		X		
Calzati & van Loenen, 2023		X		
Carballa Smichowski, 2019		X		
Cecchini et al., 2025	X		X	
Chen et al., 2022	X			
Cheng et al., 2024	X			
Cherner et al., 2025			X	X
Cocchiaro et al., 2025			X	
Cross, 2024		X		X
Dafoe et al., 2021	X			
Dasgupta & Musolesi, 2025	X			
de Souza & Bhardwaj, 2024		X		
Decker et al., 2025		X		X
Dua et al., 2025			X	X
Eke & Stahl, 2024		X	X	
Farisco et al., 2024			X	
Gambelin, 2020			X	
Gandorfer, 2022		X	X	X
Garcia, 2024	X			
Gerstgrasser et al., 2023	X			
Gil et al., 2025		X		
Guan et al., 2025	X			
Hagendorff, 2023			X	
Huang et al., 2025			X	X
Keane & Allen, 2022	X			
Koster et al., 2022				X
LaCroix & Mohseni, 2022	X	X		
Makandar, 2024			X	X
Marvin et al., 2023			X	X
McKee, 2024	X			
McKee et al., 2024	X			
Mikhaylovskaya, 2024				X
Montes, 2022	X		X	
Mun et al., 2025				X
Noorman & Swierstra, 2023				X
Okada et al., 2025		X	X	
Orzan et al., 2024	X			
Orzan et al., 2025	X			
Perera et al., 2025	X			
Rachum et al., 2024	X	X		
Ribeiro-Navarrete et al., 2024				X
Saheb, 2024			X	X
Sato & Osawa, 2020	X			
Sætra et al., 2022				X

Schrepel & Potts, 2024		X		
Shi et al., 2025	X			
Simões et al., 2020	X			
Stathis & van den Herik, 2025		X	X	
Steinhoff, 2024			X	X
Su et al., 2024	X			
Tartaro et al., 2024			X	
Taylor et al., 2025				X
Tomašev et al., 2020		X	X	X
Trinidad, 2025		X		
Tsamados et al., 2024	X			
Veitch et al., 2024	X			
Verdegem, 2022		X	X	
von Schenk et al., 2023	X			
Wang et al., 2024	X			
Woods, 2025		X	X	
Xu, 2022		X		
Zhang et al., 2019	X			
Zheng et al., 2023		X		
Zygmuntowski, 2023		X	X	
Totale studi	27	25	22	19

Tabella 3. Distribuzione dei 70 studi selezionati nella *scoping review* secondo quattro assi tematici: Cooperative AI, Governance cooperativa, AI4EC e Democratic AI

Questa complessità è ulteriormente rappresentata nella Figura 2, che mostra un diagramma di Venn della distribuzione dei contributi. In particolare, l'area azzurra raccoglie i lavori sulla Cooperative AI e dilemmi sociali (27 paper), l'area arancione quelli sulla Governance cooperativa dell'IA (25 paper), l'area viola quelli relativi all'AI4EC (22 paper), mentre l'area verde rappresenta gli studi sulla Democratic AI (19 paper). Le intersezioni tra le aree rendono visibili le connessioni tra i diversi ambiti: 23 contributi sono dedicati esclusivamente alla Cooperative AI, 13 alla Governance cooperativa, 8 alla Democratic AI e 5 all'AI4EC. Allo stesso tempo, emer-

gono zone di sovrapposizione significative: 2 paper si collocano tra Cooperative AI e AI4EC, 2 tra Cooperative AI e Governance cooperativa, 6 tra AI4EC e Governance cooperativa, 7 tra AI4EC e Democratic AI, 2 tra Governance cooperativa e Democratic AI, mentre 2 contributi integrano prospettive provenienti da tre assi contemporaneamente. Questa visualizzazione restituisce con chiarezza la natura intrinsecamente interconnessa del dibattito scientifico contemporaneo su IA, cooperazione, governance e democrazia.

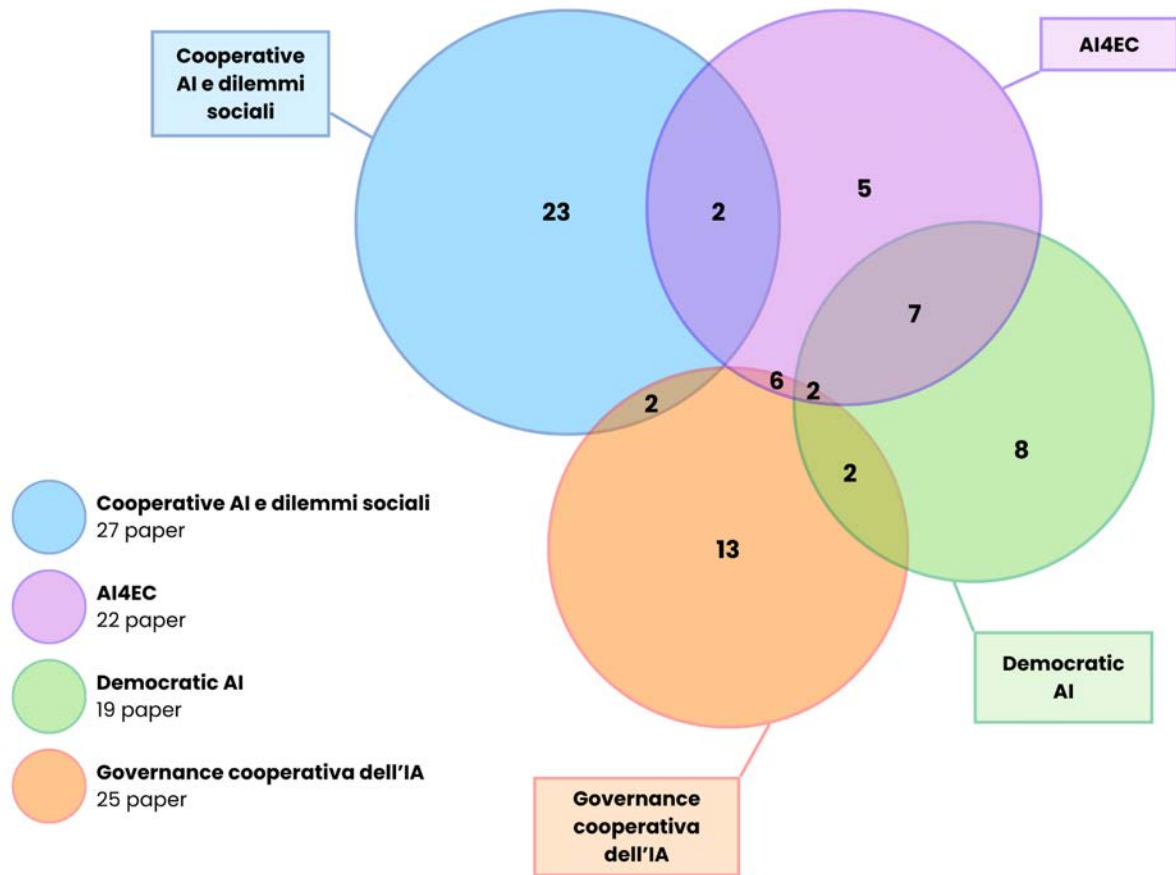


Figura 2. Diagramma di Venn della distribuzione dei contributi per asse d'indagine

Ciascun asse si caratterizza inoltre per specifiche traiettorie disciplinari. L'asse Cooperative AI e dilemmi sociali raccoglie contributi prevalentemente radicati nelle scienze computazionali, nella teoria dei giochi e nell'apprendimento multi-agente, con un focus sui meccanismi di cooperazione emergente e sul coordinamento tra agenti artificiali e umani. L'asse Governance cooperativa dell'IA si colloca invece tra diritto, scienze politiche ed economia, esplorando modelli alternativi di gestione dei dati, come *data cooperatives* e *data commons*, e forme di governance partecipativa e decentralizzata. L'asse AI4EC si sviluppa all'intersezione tra etica applicata, filosofia della tecnologia e scienze sociali, concentrandosi su temi come l'allineamento valoriale, la responsabilità degli attori e l'impatto sociale dell'IA. Infine, l'asse Democratic AI integra prospettive di teoria politica e studi sulla democrazia digitale, interrogandosi su inclusione, trasparenza algoritmica e legittimazione democratica dei sistemi intelligenti.

Nel loro insieme, questi assi non rappresentano compartimenti rigidi, ma spazi interpretativi dinamici, che consen-

tono di leggere il campo nella sua complessità e di individuare sia tendenze consolidate sia vuoti di ricerca ancora aperti.

Accanto alla distribuzione tematica, è stata analizzata anche la dimensione temporale della produzione scientifica. Come illustrato nella Figura 3, che riporta il trend delle pubblicazioni per anno, emerge una crescita inizialmente contenuta tra il 2018 e il 2021 (con rispettivamente 1 contributo nel 2018, 2 nel 2019, 4 nel 2020 e 2 nel 2021), seguita da un'accelerazione significativa a partire dal 2022 (9 pubblicazioni). Questa crescita si intensifica ulteriormente nel biennio successivo, con 11 contributi nel 2023 e un picco di 23 nel 2024. Il 2025 registra già 18 pubblicazioni, nonostante la revisione includa articoli solo fino ad aprile, suggerendo un'ulteriore espansione del campo. Nel complesso, il trend evidenzia una forte accelerazione negli ultimi tre anni, segnalando l'urgenza crescente di affrontare le questioni legate a governance, etica e cooperazione nell'ecosistema dell'intelligenza artificiale.

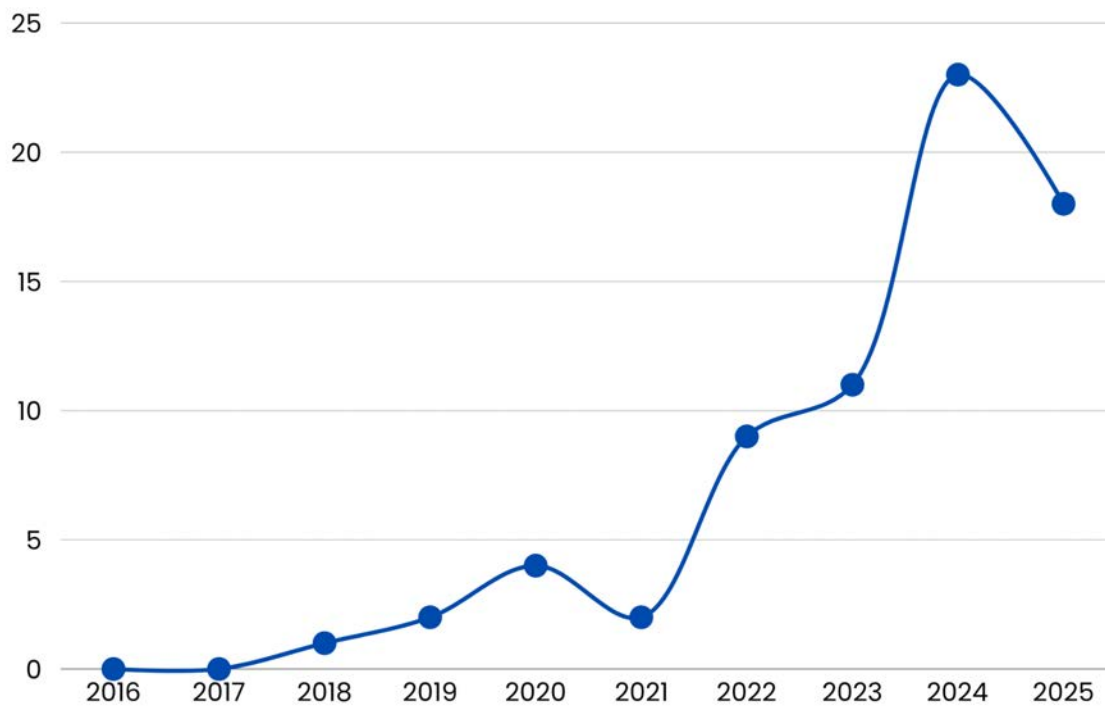


Figura 3. Analisi temporale: trend lineare delle pubblicazioni per anno (2016–2025)

3.2. RQ1: Cooperative AI e dilemmi sociali

Il termine *Cooperative AI* compare per la prima volta nell'articolo di Dafoe et al. (2020) *Open problems in cooperative AI* e successivamente in Dafoe et al. (2021), *Cooperative AI: Machines must learn to find common ground*. Questi contributi pongono le basi teoriche di un nuovo ambito di ricerca dedicato a rafforzare la capacità cooperativa dell'intelligenza artificiale avanzata, con l'obiettivo di produrre benefici condivisi. Il concetto, tuttavia, affonda le sue radici in una fase precedente: riflessioni e sperimentazioni erano già in corso nei campi della cooperazione multi-agente e dell'etica dell'IA. La formulazione del termine ha dunque svolto la funzione di consolidare e rendere riconoscibile un insieme di idee già emergenti all'interno di un quadro teorico comune.

Gli autori sottolineano l'urgenza di avviare un nuovo percorso di ricerca, volto alla costruzione di una vera e propria "scienza dell'intelligenza artificiale cooperativa" (Dafoe et al., 2021, p. 34):

"We need to build a science of cooperative AI. As researchers in the field and its governance, we argue that it is time to prioritize the development of cooperative intelligence that has the ability to promote mutually beneficial joint action, even when incentives are not fully aligned. Just as psychologists studying humans have found that the infant brain does not develop fully without social interaction, progress towards socially valuable AI will be stunted unless we put the problem of cooperation at the centre of our research" (Dafoe et al., 2021, p. 34).

Secondo Dafoe et al. (2021), la ricerca sull'IA cooperativa deve articolarsi attorno a tre ambiti distinti:

- Cooperazione IA-IA;
- Cooperazione uomo-IA;
- Cooperazione uomo-uomo mediata dall'IA.

In ciascuno di questi contesti, i ricercatori sono chiamati a riprodurre scenari realistici, che non si riducano né a situazioni di conflitto puro, prive di margine di contrattazione e negoziazione, né a contesti di perfetta armonia di interessi, che rappresentano il caso più semplice. Le relazioni del mondo reale, infatti, combinano elementi di interesse comune e di conflitto. Riprodurre e analizzare scenari misti rappresenta, per Dafoe et al., il modo più promettente per consentire all'IA di produrre valore sociale.

Si tratta dei cosiddetti dilemmi sociali, situazioni in cui nessun attore (umano o artificiale) ha un incentivo individuale a cooperare, pur essendo la cooperazione reciproca la condizione che massimizza il beneficio collettivo. Come mostrano Orzan et al. (2024, 2025), uno dei principali ostacoli all'emergere della cooperazione è l'incertezza, che assume due forme:

- Incertezza *ambientale*, legata alla variabilità delle risorse disponibili;
- Incertezza *sociale*, legata all'imprevedibilità del comportamento altrui.

Da queste tensioni derivano i nodi centrali della cooperazione umana – contrattazione, fiducia e sfiducia, inganno e comunicazione credibile, impegno e garanzie, politica e coalizioni, norme e istituzioni – che anche gli agenti di intelligenza artificiale dovranno imparare a gestire.

Per affrontare l'incertezza nei dilemmi sociali e sviluppare agenti autonomi orientati a valori prosociali, la ricerca sulla *cooperative AI* si divide in due principali approcci (Montes, 2022):

- Approccio *collettivo e sistemico*, che analizza la teoria dei giochi e i meccanismi sociali capaci di incentivare comportamenti a beneficio dell'intera comunità di agenti (Dasgupta & Musolesi, 2025; LaCroix & Mohseni, 2022);
- Approccio *individuale*, che mira a potenziare le capacità interne dei singoli agenti, ad esempio attraverso il coor-

dinamento implicito (Keane & Allen, 2022), l'empatia artificiale (Chen et al., 2022) e la condivisione selettiva delle esperienze rilevanti (Gerstgrasser et al., 2023).

Entrambi gli approcci si concentrano su elementi che Dafoe et al. (2021) individuano come cruciali per ottenere esiti reciprocamente vantaggiosi: comprensione, comunicazione, impegno, norme e istituzioni.

Questo filone di indagine ridefinisce le finalità di ricerca tradizionali dell'IA: non più solo competere o ottimizzare, ma cooperare per servire il bene sociale e affrontare le grandi sfide globali. Ne emerge una prospettiva di un futuro in cui gli agenti artificiali diventino parte integrante di strutture sociali umane, capaci di interagire eticamente, negoziare obiettivi comuni e contribuire alla costruzione di società più giuste e inclusive.

3.3. RQ2: Governance cooperativa dell'IA

Da un lato, la ricerca sulla *cooperative AI* utilizza la teoria dei giochi per individuare le variabili sociali e organizzative che possono incentivare comportamenti cooperativi. Dall'altro, un filone critico sottolinea come l'attuale modello di sviluppo dell'intelligenza artificiale rimanga ancorato a logiche estrattive, segnato da opacità e concentrazione monopolistica.

In questo contesto, parlare di governance dell'IA significa innanzitutto affrontare la questione della governance dei dati, indispensabili per addestrare i sistemi, applicare l'addestramento e migliorarne progressivamente le prestazioni. Allo stesso tempo, significa anche confrontarsi con i cosiddetti *data-opolies*: termine coniato da Stucke (2022), dall'unione di "*data*" e "*monopolies*", che descrive l'accentramento del potere in poche piattaforme capaci di controllare accesso, gestione e valorizzazione dei dati.

Zygmuntowski (2023) mostra come la governance dei dati, inizialmente concepita come funzione tecnica interna alle organizzazioni per ottimizzare la gestione degli asset informativi (archiviazione, qualità, sicurezza), si sia progressivamente trasformata in un dispositivo di controllo e influenza. Oggi essa è una governance sistemica, che coinvolge società civile, mercato e istituzioni, e riguarda direttamente questioni di potere, diritti e distribuzione del valore. Per rappresentare visivamente queste tensioni, l'autore propone il "Data Governance Trilemma" (DGT), che mette in evidenza la difficile conciliazione tra tre obiettivi: tutela dei diritti fondamentali, produzione di valore economico e perseguimento dell'interesse pubblico.

Insieme a Verdegem, Zygmuntowski (2023; Verdegem, 2022) individua nei "*digital commons*" una possibile risposta al trilemma, fondata su partecipazione, trasparenza, redistribuzione e controllo democratico delle risorse. Tale prospettiva si collega alle riflessioni di LaCroix e Mohseni (2022) e Xu (2022), contribuendo a delineare una visione politica ed economica radicale che apre la strada a una governance cooperativa dell'AI.

Secondo Zygmuntowski (2023), i "*data commons*" non sono un modello unico, ma un insieme flessibile di esperimenti accomunati dall'idea di governance collettiva. Tra questi rientrano le cooperative di dati, costituite da individui o organizzazioni che mettono in comune i propri dati per creare risorse condivise.

In questa direzione, Calzada (2024), in "Democratic Erosion of Data-Opolies: Decentralized Web3 Technological Paradigm Shift Amidst AI Disruption", identifica le cooperative di dati come una delle tre strategie per ridisegnare il pano-

rama dei dati e restituire a individui e comunità proprietà, autonomia e controllo. L'autore sottolinea tre aspetti centrali:

- L'architettura decentralizzata e i protocolli blockchain come garanzia di gestione affidabile;
- Le DAO come strumenti in grado di rivoluzionare il processo decisionale collettivo, superando l'autorità centrale;
- Le cooperative di dati come ponte tra proprietà collettiva ed equo compenso.

La convergenza tra blockchain, DAO e cooperative di dati, secondo Calzada (2024), amplifica il loro impatto e apre la strada al superamento dell'estrattivismo dei dati.

Ne consegue che, per andare oltre le logiche estrattive e monopolistiche, una buona governance dei dati, e quindi una governance cooperativa dell'IA, non può separarsi dalla società che quei dati rappresentano. Deve invece trattarli come mezzi democratici, integrando esplicitamente i diritti, i valori e gli obiettivi che essi incorporano.

3.4. RQ3: Artificial Intelligence for Ethical Cooperation (AI4EC)

Se è necessario definire e applicare politiche, procedure e standard per lo sviluppo, l'uso e la gestione corretta dell'IA, è altrettanto importante studiare e valutare i problemi morali legati a dati, algoritmi, pratiche e infrastrutture, così da orientare valori e condotte. Tutto ciò riguarda l'etica dell'IA, che incide sulla governance e sul sistema legislativo attraverso una valutazione morale di ciò che è socialmente accettabile o preferibile.

Come scrive Floridi in "Ethics of artificial intelligence: Principles, challenges, and opportunities" (2023), l'etica digitale può essere intesa in due modi: (i) etica *hard*, che contribuisce a plasmare o modificare il diritto e (ii) etica *soft*, che indica ciò che dovrebbe o non dovrebbe essere fatto al di là della normativa. L'etica, dunque, precede e segue la legge: la rende possibile e, successivamente, la integra, talvolta costringendola a cambiare (Floridi & Taddeo, 2016; Floridi, 2023).

Con l'ascesa dell'IA, sono emerse nuove figure professionali, come l'AI Ethicist, molte organizzazioni hanno avviato iniziative per stabilire principi etici a sostegno di un'IA socialmente vantaggiosa e i governi hanno cominciato a produrre strategie nazionali. Tuttavia, come osservano Dua et al. (2025), permangono vuoti significativi, ad esempio l'assenza di contributi dall'Africa Subsahariana, e le stesse definizioni di "etica" e "IA responsabile" variano da Paese a Paese, spesso plasmate dalle specifiche strutture politiche e dal ruolo della società civile. Nel loro studio "The ethics of national artificial intelligence plans: an empirical lens" (Dua et al., 2025), gli autori analizzano 54 documenti nazionali, mappando la governance etica dell'IA a livello globale. Dai risultati emergono tre principali formazioni geopolitiche:

- Formazione iberoamericana, che privilegia temi come la diversità di genere e la forza lavoro;
- Formazione dei "*science and tech first movers*" (Dua et al., p. 3805), ovvero Stati Uniti e Cina, orientata soprattutto allo sviluppo infrastrutturale per la ricerca, con cenni superficiali a equità, pregiudizi e discriminazione;
- Formazione europea, che adotta un approccio normativo e si concentra su mobilità sociale, questioni di genere, aspetti culturali, sostenibilità, standardizzazione e governance democratica dell'IA.

Indipendentemente dal contesto geografico, il passaggio dall'elaborazione dei principi alla loro applicazione pratica comporta spesso il rischio di comportamenti contrari agli stessi valori dichiarati. Si tratta del fenomeno noto come "ethics washing": la proliferazione di principi e valori etici diventa un mercato in cui attori pubblici e privati scelgono quelli più funzionali a giustificare pratiche già in atto, senza realmente modificarle. In questo modo, gli attori pubblicizzano il più possibile il loro impegno etico senza produrre alcun miglioramento reale.

Steinhoff (2024) interpreta il fenomeno in chiave radicale: non si tratterebbe di deviazioni isolate, ma dell'intera etica dell'IA, intesa come operazione di facciata. L'etica, in questa prospettiva, non renderebbe l'IA più giusta, ma funzionerebbe come una rete di innovazione subordinata, su cui le Big Tech esercitano un controllo indiretto per orientare la ricerca verso finalità produttive e di profitto. Quando gli ideali etici entrano in conflitto con gli interessi economici, prevale sempre il capitale, generando una contraddizione strutturale tra etica e accumulazione, tra benessere umano e profitto (Steinhoff, 2024).

Gandorfer (2022) invita invece a una postura di "complicità critica". Non si tratta di rigettare a priori nuove esperienze etiche e politiche, ma di indagarne le contraddizioni strutturali, ideologiche e materiali. Per l'autrice, non esiste una posizione esterna o neutrale rispetto alle logiche neoliberali: la complicità è una condizione ontologica dell'epoca attuale. Non ci si può tirare fuori dal gioco, occorre piuttosto "sporcarsi le mani" nella sperimentazione politica, accettando l'indeterminatezza e cercando di modulare la propria complicità in modo creativo, invece di inseguire una purezza impossibile (Gandorfer, 2022).

In questo scenario globale e tra queste tensioni emergenti, sorge la domanda su come il movimento cooperativo possa posizionarsi. Se saprà muoversi in questo terreno complesso, costruendo gerarchie e modelli convincenti per le comunità, potrà non limitarsi a denunciare complicità o strutture oppressive, ma agire al loro interno in modo strategico. Non con una critica "pura" o esterna, ma con una critica immanente, gettando le basi per quella che in questo scritto chiamiamo "Artificial Intelligence for Ethical Cooperation" (AI4EC): un'IA che persegue ciò che è socialmente preferibile.

3.5. RQ4: Democratic AI

Un'ulteriore tendenza osservabile tra attori pubblici e privati è quella che potremmo definire "*democratic washing*", ossia la riduzione della democratizzazione dell'IA a una mera questione di accessibilità tecnica. In questa prospettiva, l'IA viene presentata come un servizio da semplificare e diffondere capillarmente con la promessa di risolvere problemi complessi come le disuguaglianze sociali o la corruzione politica.

Sætra et al. (2022) denunciano l'uso sempre più frequente ma fuorviante del termine "democrazia" nella ricerca sull'IA. Secondo gli autori, il concetto viene spesso associato a meccanismi come l'aggregazione di preferenze, le votazioni maggioritarie o l'apertura del codice, ma senza un'adeguata base teorica né un reale ancoraggio alla tradizione della filosofia politica. Tra gli esempi più controversi viene citato il modello di "*Democratic AI*" elaborato da Koster et al. (2022), che riduce la democrazia a un mero dispositivo di ottimizzazione delle preferenze della maggioranza.

A questa critica si ricollegano Noorman e Swierstra (2023) che sottolineano la necessità di un approccio alla democra-

tizzazione dell'IA più sostanziale e radicato nelle dinamiche socio-tecniche. Le decisioni prese dai sistemi IA, infatti, sono collettivamente vincolanti: riguardano tutti i cittadini, anche contro la loro volontà. In questo senso, l'IA non è solo uno strumento tecnico, ma agisce come un vero e proprio agente di allocazione autoritativa di valori, assumendo una funzione politica. Non basta rendere disponibili strumenti tecnici: è necessario interrogarsi su come e da chi tali strumenti vengano progettati, controllati e utilizzati, adottando una prospettiva in cui sociale e tecnico si influenzano reciprocamente e in cui le decisioni sull'IA siano riconosciute come scelte collettive e politiche.

Ispirandosi alla prospettiva socio-tecnica di Warren, Noorman e Swierstra osservano che la democrazia non coincide con un unico insieme di pratiche, ma è un sistema chiamato a svolgere tre funzioni fondamentali:

- "Empowered inclusion", ossia un'inclusione proattiva che consenta a tutti i soggetti coinvolti di incidere sulle decisioni;
- "Collective will formation", la formazione di una volontà collettiva, che richiede sistemi capaci di anticipare e facilitare il dissenso e la trasformazione dei valori;
- "Collective decision capacity", la capacità di prendere decisioni efficaci, applicabili e trasparenti.

Le pratiche che consentono al sistema democratico di svolgere queste funzioni comprendono il riconoscimento reciproco, la resistenza alle ingiustizie, la deliberazione tra eguali, il voto, la rappresentanza, la partecipazione a collettivi attivi e i segnali correttivi inviati alle istituzioni.

In questa cornice si inserisce anche il contributo di Decker et al. (2025), i quali spostano l'attenzione dagli esiti ("*outcomes*") delle decisioni algoritmiche ai processi decisionali stessi ("*procedural fairness*"), ponendo particolare enfasi sulla partecipazione pubblica come elemento normativo e trasformativo. Non è sufficiente correggere i risultati: occorre trasformare le dinamiche relazionali e ridurre le asimmetrie di potere nei processi decisionali.

Ogni fase dello sviluppo algoritmico implica scelte valoriali (dalla raccolta dei dati al "pre-processing", dalla progettazione dell'algoritmo alla definizione delle metriche di equità) e chi prende queste decisioni esercita un potere normativo, stabilendo cosa sia "equo" o "valido". Poiché sviluppatori e progettisti non sono attori neutrali, i loro valori, ipotesi e background influenzano inevitabilmente gli esiti. Per ciascuna di queste fasi, Decker et al. immaginano forme di coinvolgimento pubblico, pur riconoscendone limiti e ostacoli, e ricordano che la democrazia è per sua natura dispendiosa in termini di risorse: rinunciare al coinvolgimento per motivi di complessità non è giustificabile, soprattutto quando le tecnologie incidono profondamente sulla società.

La democrazia è anche uno dei principi fondanti della cooperazione: le cooperative hanno storicamente sviluppato pratiche e regole (come il principio "una testa, un voto") per bilanciare interessi individuali e collettivi, dando forma a processi decisionali inclusivi. In questa prospettiva, la cooperazione può offrire un ancoraggio socio-politico più solido e una tradizione consolidata di governance democratica, contribuendo a ripensare la democratic AI come democrazia sostanziale, e non soltanto procedurale.

5. Limiti della ricerca

La presente *scoping review* presenta alcuni limiti che è necessario esplicitare per una corretta interpretazione dei ri-

sultati. In primo luogo, la varietà terminologica che caratterizza il campo dell'intelligenza artificiale, della cooperazione e della governance socio-tecnica può aver comportato l'esclusione di contributi potenzialmente rilevanti, nonostante l'adozione di una strategia di ricerca ampia e ciclica. A questo si aggiunge un *bias* linguistico, derivante dalla scelta di includere esclusivamente studi pubblicati in lingua inglese. Tale decisione, se da un lato ha garantito coerenza e accesso a una base scientifica ampia e consolidata, dall'altro può aver limitato la rappresentazione di contributi sviluppati in contesti locali o in lingue diverse, in particolare in ambiti, come quello cooperativo, fortemente radicati in specifiche tradizioni culturali e territoriali.

La natura ancora emergente e in continua evoluzione di questo ambito implica inoltre che i risultati qui presentati rappresentino una fotografia situata; una revisione condotta nei prossimi anni potrebbe restituire configurazioni teoriche e traiettorie di ricerca significativamente diverse.

Un ulteriore limite riguarda la possibile sottorappresentazione di pratiche e applicazioni sviluppate in contesti non accademici. In particolare, molte sperimentazioni legate all'IA cooperativa, alla governance dei dati e alle forme di intelligenza artificiale democratica si sviluppano in ambiti progettuali, organizzativi o industriali che non sempre trovano immediata formalizzazione nella letteratura scientifica, generando così uno scarto tra pratiche emergenti e loro documentazione.

Dal punto di vista metodologico, la revisione include studi caratterizzati da elevata eterogeneità per approcci, discipline, contesti e livelli di maturità empirica. Questa varietà, se da un lato rappresenta un valore in termini di ampiezza esplorativa, dall'altro limita la comparabilità sistematica dei risultati e la possibilità di trarre generalizzazioni robuste.

Ulteriori criticità riguardano la qualità e la trasparenza dei dati riportati negli studi, nonché la limitata attenzione, in alcuni casi, agli aspetti metodologici. In diversi contributi, infatti, le modalità di raccolta e analisi dei dati non sono esplicitate in modo dettagliato, rendendo complessa una valutazione approfondita della solidità delle evidenze. A ciò si aggiungono le ben note sfide etiche legate all'intelligenza artificiale, tra cui *bias* algoritmici, opacità dei sistemi decisionali e questioni di accountability, che, pur frequentemente richiamate, non sempre vengono affrontate con adeguato livello di approfondimento.

Nonostante questi limiti, la review offre una mappatura strutturata e aggiornata di un campo di ricerca ancora in fase di consolidamento, contribuendo a evidenziare connessioni emergenti, nodi critici e direzioni di sviluppo future. In questa prospettiva, il lavoro assume anche una valenza orientativa ed educativa in senso ampio, fornendo strumenti interpretativi utili per comprendere e accompagnare l'evoluzione dell'intelligenza artificiale in chiave cooperativa, etica e democratica. Le analisi future potranno rafforzare questi risultati ampliando le strategie di ricerca, includendo letteratura grigia e studi in altre lingue, sviluppando disegni empirici più robusti e approfondendo il dialogo interdisciplinare tra tecnologia, scienze sociali ed educazione.

6. Conclusioni e prospettive

La *scoping review* ha evidenziato come la cooperazione rappresenti una cornice interpretativa trasversale sempre più rilevante nel dibattito contemporaneo sull'IA, articolandosi lungo quattro principali traiettorie di ricerca – Cooperative AI e dilemmi sociali, Governance cooperativa dell'IA, Artificial Intelligence for Ethical Cooperation (AI4EC) e Democratic AI.

Sebbene differenti per impostazione disciplinare, tali aree convergono nel delineare la cooperazione come principio tecnico, politico ed educativo capace di offrire prospettive significative per ripensare criticamente lo sviluppo e l'uso delle tecnologie intelligenti.

Sul piano tecnico, la letteratura suggerisce che la cooperazione costituisca una dimensione centrale nella progettazione di sistemi di IA orientati alla gestione di dilemmi sociali, al coordinamento tra agenti artificiali e umani e alla costruzione di meccanismi di fiducia che eccedano la sola ottimizzazione algoritmica. Sul piano politico-istituzionale, emergono modelli di governance cooperativa – tra cui *data cooperatives*, *platform cooperativism* e *commons* digitali – che propongono alternative rilevanti rispetto ai paradigmi estrattivi dominanti, aprendo nuove possibilità per forme più partecipative e democratiche di gestione tecnologica. Sul piano educativo, i risultati mostrano come l'IA richieda percorsi formativi intesi in senso ampio, capaci di sostenere processi di cittadinanza critica, alfabetizzazione socio-tecnica e partecipazione consapevole, andando oltre una concezione puramente strumentale delle competenze digitali.

Un elemento particolarmente significativo emerso dalla review riguarda il fatto che cooperazione e democrazia, nel contesto dell'IA, non sembrano riducibili a meri dispositivi procedurali o dichiarazioni valoriali, ma richiedono pratiche sostanziali di inclusione, deliberazione e redistribuzione del potere. In questo quadro, i fenomeni di *ethics washing* e *democratic washing* segnalati dalla letteratura indicano come approcci esclusivamente dichiarativi risultino spesso insufficienti, sottolineando invece l'importanza di sviluppare capacità critiche diffuse nei soggetti, nelle organizzazioni e nei contesti educativi.

La revisione ha inoltre permesso di individuare diversi ambiti che richiedono ulteriori approfondimenti: la traduzione dei principi cooperativi in architetture algoritmiche e istituzionali concretamente applicabili; la sostenibilità e la scalabilità di modelli come cooperative di dati, *commons* digitali e piattaforme cooperative; la valutazione empirica dell'efficacia di approcci educativi partecipativi; nonché il superamento delle persistenti asimmetrie geopolitiche nella definizione globale di etica, governance e democrazia dell'IA.

Nel complesso, la *scoping review* non intende offrire soluzioni definitive, ma propone una mappatura critica di un campo di ricerca ancora in consolidamento, mettendo in luce traiettorie emergenti, convergenze interdisciplinari e principali lacune conoscitive. I risultati suggeriscono che collocare la cooperazione al centro della riflessione sull'intelligenza artificiale possa costituire una prospettiva promettente per orientare future ricerche teoriche, empiriche ed educative. In conclusione, le grandi sfide poste dall'intelligenza artificiale offrono un orizzonte condiviso, un ombrello sotto il quale realtà con mission e specificità differenti possano ritrovarsi, favorendo collaborazioni e convergenze generative e inedite: tra ricerca e pratica, tra educazione formale e apprendimento organizzativo, tra movimento cooperativo e governance tecnologica. È in questo intreccio che si gioca, oggi, la possibilità di un'intelligenza artificiale autenticamente cooperativa, democratica ed eticamente fondata.

Riferimenti bibliografici

- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19–32. <https://doi.org/10.1080/1364557032000119616>
- Asa, Y., Kato, T., & Mine, R. (2023). Composite consensus-building

- process: Permissible meeting analysis and compromise choice exploration. In Y. Maemura, M. Horita, L. Fang, & P. Zaraté (Eds.), *Group decision and negotiation in the era of multimodal interactions: GDN 2023* (Lecture Notes in Business Information Processing, Vol. 478, pp. 97–112). Springer. https://doi.org/10.1007/978-3-031-33780-2_7
- Benton, P. (2024). Democratic AI: Justification for a broad view of public reason. In A. Gerber, J. Maritz, & A. W. Pillay (Eds.), *Artificial intelligence research: SACAIR 2024* (Communications in Computer and Information Science, Vol. 2326, pp. 440–455). Springer. https://doi.org/10.1007/978-3-031-78255-8_26
- Biesta, G., & Tedder, M. (2007). Agency and learning in the lifecourse: Towards an ecological perspective. *Studies in the Education of Adults*, 39(2), 132–149. <https://doi.org/10.1080/02660830.2007.11661545>
- Blasimme, A., Vayena, E., & Hafen, E. (2018). Democratizing health research through data cooperatives. *Philosophy & Technology*, 31(3), 473–479. <https://doi.org/10.1007/s13347-018-0320-8>
- Bruder, R., Stockinger, C., Petrat, D., & Subtil, I. (2021). Development of cooperative artificial intelligence (AI) applications to support human work in manufacturing. In N. L. Black, W. P. Neumann, & I. Noy (Eds.), *Proceedings of the 21st Congress of the International Ergonomics Association: IEA 2021* (Lecture Notes in Networks and Systems, Vol. 221, pp. 391–397). Springer. https://doi.org/10.1007/978-3-030-74608-7_49
- Buckingham, D. (2019). *The media education manifesto*. Polity Press.
- Calzati, S., & van Loenen, B. (2023). Towards a citizen- and citizenry-centric digitalization of the urban environment: Urban digital twinning as commoning. *Digital Society*, 2(3), Article 38. <https://doi.org/10.1007/s44206-023-00064-0>
- Calzada, I. (2023). Disruptive technologies for e-diasporas: Blockchain, DAOs, data cooperatives, metaverse, and ChatGPT. *Futures*, 154, Article 103258. <https://doi.org/10.1016/j.futures.2023.103258>
- Calzada, I. (2024a). Artificial intelligence for social innovation: Beyond the noise of algorithms and datafication. *Sustainability*, 16(19), Article 8638. <https://doi.org/10.3390/su16198638>
- Calzada, I. (2024b). Democratic erosion of data-opolies: Decentralized Web3 technological paradigm shift amidst AI disruption. *Big Data and Cognitive Computing*, 8(3), Article 26. <https://doi.org/10.3390/bdcc8030026>
- Carballa Smichowski, B. (2019). Alternative data governance models: Moving beyond one-size-fits-all solutions. *Intereconomics*, 54(4), 222–227. <https://doi.org/10.1007/s10272-019-0828-x>
- Cecchini, D., Pflanzner, M., & Dubljević, V. (2025). Aligning artificial intelligence with moral intuitions: An intuitionist approach to the alignment problem. *AI Ethics*, 5(2), 1523–1533. <https://doi.org/10.1007/s43681-024-00496-5>
- Chen, J., Zhang, D., Qu, Z., & Wang, C. (2022). Modeling adaptive empathy based on neutral assessment: A way to enhance the prosocial behaviors of socialized agents under the premise of self-security. *Applied Intelligence*, 52, 6692–6722. <https://doi.org/10.1007/s10489-021-02712-9>
- Cheng, Y., Zhang, W., Zhang, Z., Zhang, C., Wang, S., & Mao, S. (2024). Towards federated large language models: Motivations, methods, and future directions. *IEEE Communications Surveys & Tutorials*, 27(4), 2733–2764. <https://doi.org/10.1109/COMST.2024.3503680>
- Cherner, T., Foulger, T. S., & Donnelly, M. (2025). Introducing a generative AI decision tree for higher education: A synthesis of ethical considerations from published frameworks & guidelines. *TechTrends*, 69(1), 84–99. <https://doi.org/10.1007/s11528-024-01023-3>
- Cocchiaro, M. Z., Morley, J., Novelli, C., Panai, E., Tartaro, A., & Floridi, L. (2025). Who is an AI ethicist? An empirical study of expertise, skills, and profiles to build a competency framework. *AI Ethics*, 5(4), 3713–3725. <https://doi.org/10.1007/s43681-024-00643-y>
- Cross, M. (2024). Limitations of transparency in democratising and regulating algorithmic management. In M. Ziosi, G. Sartor, J. M. Cunha, A. Trotta, & P. Wicke (Eds.), *AI for people, democratizing AI: CAIP 2023* (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Vol. 591, pp. 86–94). Springer. https://doi.org/10.1007/978-3-031-71304-0_7
- Dafoe, A., Bachrach, Y., Hadfield, G. K., Horvitz, E., Larson, K., & Graepel, T. (2021). Cooperative AI: Machines must learn to find common ground. *Nature*, 593(7857), 33–36. <https://doi.org/10.1038/d41586-021-01170-0>
- Dafoe, A., Hughes, E., Bachrach, Y., Collins, T., McKee, K. R., Leibo, J. Z., Larson, K., & Graepel, T. (2020). Open problems in cooperative AI. *arXiv*. <https://doi.org/10.48550/arXiv.2012.08630>
- Dasgupta, N., & Musolesi, M. (2025). Investigating the impact of direct punishment on the emergence of cooperation in multi-agent reinforcement learning systems. *Autonomous Agents and Multi-Agent Systems*, 39(1), Article 19. <https://doi.org/10.1007/s10458-025-09698-5>
- de Souza, S., & Bhardwaj, K. (2024). India's conception of community data and addressing concerns for access to justice. *Digital Society*, 3(1), Article 16. <https://doi.org/10.1007/s44206-024-00102-5>
- Decker, M., Wegner, L., & Leicht-Scholten, C. (2025). Procedural fairness in algorithmic decision-making: The role of public engagement. *Ethics and Information Technology*, 27(1), Article 1. <https://doi.org/10.1007/s10676-024-09811-4>
- Dua, M., Singh, J. P., & Shehu, A. (2025). The ethics of national artificial intelligence plans: An empirical lens. *AI Ethics*, 5(4), 3803–3831. <https://doi.org/10.1007/s43681-025-00663-2>
- Eke, D., & Stahl, B. (2024). Ethics in the governance of data and digital technology: An analysis of European data regulations and policies. *Digital Society*, 3(1), Article 11. <https://doi.org/10.1007/s44206-024-00101-6>
- Ellerani, P., & Ferrari, L. (2024). The contribution of generative AI ecosystems in micro-instructional design: Opportunities and limitations. *Formazione & insegnamento*, 22(1), 117–124. https://doi.org/10.7346/fe-XXII-01-24_13
- Fabbri, M. (2025). *Dall'empowerment all'agency pedagogico-trasformativa: Prospettive di ricerca e linee per l'azione didattica*. FrancoAngeli.
- Farisco, M., Baldassarre, G., Cartoni, E., Leach, A., Petrovici, M. A., Rosemann, A., Salles, A., Stahl, B., & van Albada, S. J. (2024). A method for the ethical analysis of brain-inspired AI. *Artificial Intelligence Review*, 57(6), Article 133. <https://doi.org/10.1007/s10462-024-10769-4>
- Fengchun, M., & Mutlu, C. (2024). *AI competency framework for teachers*. UNESCO.
- Ferrari, L., & Ingrassio, P. (2024). *COO.DE - Cooperative digital education: Percorsi emancipativi in prospettiva inclusiva*. Erickson.
- Floridi, L. (2023). *Ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.
- Floridi, L., & Taddeo, M. (2016). What is data ethics? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2083), Article 20160360. <https://doi.org/10.1098/rsta.2016.0360>
- Gambelin, O. (2020). Brave: What it means to be an AI ethicist. *AI Ethics*, 1(1), 87–91. <https://doi.org/10.1007/s43681-020-00020-5>
- Gandorfer, D. (2022). Down and dirty in the field of play: Startup societies, cryptostatecraft, and critical complicity. *Law and Critique*, 33, 355–377. <https://doi.org/10.1007/s10978-022-09327-0>
- Garcia, E. V. (2024). The Other in the machine: Diplomacy and the AI conundrum. *Place Branding and Public Diplomacy*, 21, 339–344. <https://doi.org/10.1057/s41254-024-00329-6>
- Gerstgrasser, M., Danino, T., & Keren, S. (2023). Selectively sharing experiences improves multi-agent reinforcement learning. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems (AAMAS '23)* (pp. 2433–2435). International Foundation for Autonomous Agents and Multiagent Systems. <https://dl.acm.org/doi/10.5555/3545946.3598958>
- Gil, R. M., Ryan, M., & García, R. (2025). Sovereignty by design and human values in agriculture data spaces. *Agriculture and Human Values*, 42(3), 1413–1438. <https://doi.org/10.1007/s10460-024-10674-0>
- Guan, C., Xue, K., Fan, C., Chen, F., Zhang, L., Yuan, L., Qian, C., & Yu, Y. (2024). Open and real-world human-AI coordination by heterogeneous training with communication. *Frontiers of Computer Science*, 19, Article 194314. <https://doi.org/10.1007/s11704-024-3797-6>
- Hagendorff, T. (2023). AI ethics and its pitfalls: Not living up to its own standards? *AI Ethics*, 3(1), 329–336. <https://doi.org/10.1007/s43681-022-00173-5>
- Huang, L. T. L., Papyshv, G., & Wong, J. K. (2025). Democratizing

- value alignment: From authoritarian to democratic AI ethics. *AI Ethics*, 5, 11–18. <https://doi.org/10.1007/s43681-024-00624-1>
- Kasneji, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneji, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, Article 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Keane, L., & Allen, R. E. (2022). Any-play: An intrinsic augmentation for zero-shot coordination. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS '22)* (pp. 853–861). International Foundation for Autonomous Agents and Multiagent Systems. <https://dl.acm.org/doi/10.5555/3535850.3535946>
- Koster, R., Balaguer, J., Tacchetti, A., Weinstein, A., Zhu, T., Hauser, O., Williams, D., Campbell-Gillingham, L., Thacker, P., Botvinick, M., & Summerfield, C. (2022). Human-centred mechanism design with Democratic AI. *Nature Human Behaviour*, 6(10), 1398–1407. <https://doi.org/10.1038/s41562-022-01383-x>
- LaCroix, T., & Mohseni, A. (2022). The tragedy of the AI commons. *Synthese*, 200, Article 289. <https://doi.org/10.1007/s11229-022-03763-2>
- Lévy, P. (1997). *Collective intelligence: Mankind's emerging world in cyberspace* (R. Bononno, Trans.). Perseus Books.
- Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology*, 50, 2824–2838. <https://doi.org/10.1111/bjet.12861>
- Makandar, A. (2024). Digital surveillance capitalism and cities: Data, democracy and activism. *Humanities and Social Sciences Communications*, 11(1), Article 1533. <https://doi.org/10.1057/s41599-024-03941-2>
- Malone, T. W., & Bernstein, M. S. (Eds.). (2015). *Handbook of collective intelligence*. MIT Press.
- Margiotta, U. (Ed.). (2018). *Teacher education agenda: Linee guida per la formazione iniziale dei docenti della scuola secondaria*. Erickson.
- Marvin, G., Tamale, M., Kanagwa, B., & Jjingo, D. (2023). Philosophical review of artificial intelligence for Society 5.0. In A. Yadav, S. J. Nanda, & M. H. Lim (Eds.), *Proceedings of International Conference on Paradigms of Communication, Computing and Data Analytics: PCCDA 2023* (Algorithms for Intelligent Systems, pp. 1–15). Springer. https://doi.org/10.1007/978-981-99-4626-6_1
- McKee, K. R. (2024). Human participants in AI research: Ethics and transparency in practice. *IEEE Transactions on Technology and Society*, 5(3), 279–288. <https://doi.org/10.1109/TTS.2024.3446183>
- McKee, K. R., Bai, X., & Fiske, S. T. (2024). Warmth and competence in human-agent cooperation. *Autonomous Agents and Multi-Agent Systems*, 38(1), Article 23. <https://doi.org/10.1007/s10458-024-09649-6>
- Mikhaylovskaya, A. (2024). Enhancing deliberation with digital democratic innovations. *Philosophy & Technology*, 37(1), Article 3. <https://doi.org/10.1007/s13347-023-00692-x>
- Montes, N. (2022). Engineering pro-social values in autonomous agents: Collective and individual perspectives. In D. Baumeister & J. Rothe (Eds.), *Multi-agent systems: EUMAS 2022* (Lecture Notes in Computer Science, Vol. 13442, pp. 431–434). Springer. https://doi.org/10.1007/978-3-031-20614-6_26
- Mun, J., Jiang, L., Liang, J., Chung, I., DeCario, N., Choi, Y., Kohno, T., & Sap, M. (2025). Particip-AI: A democratic surveying framework for anticipating future AI use cases, harms and benefits. In *Proceedings of the 2024 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 997–1010). AAAI Press. <https://dl.acm.org/doi/10.5555/3716662.3716749>
- Noorman, M., & Swierstra, T. (2023). Democratizing AI from a sociotechnical perspective. *Minds & Machines*, 33(4), 563–586. <https://doi.org/10.1007/s11023-023-09651-z>
- Okada, A., Sherborne, T., Panselinas, G., & Kolonis, G. (2025). Fostering transversal skills through open schooling supported by the CARE-KNOW-DO pedagogical model and the UNESCO AI competencies framework. *International Journal of Artificial Intelligence in Education*, 35(4), 1953–1998. <https://doi.org/10.1007/s40593-025-00458-w>
- Orzan, N., Acar, E., Grossi, D., & Rădulescu, R. (2024). Emergent cooperation under uncertain incentive alignment. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS '24)* (pp. 1521–1530). International Foundation for Autonomous Agents and Multiagent Systems. <https://dl.acm.org/doi/10.5555/3635637.3663012>
- Orzan, N., Acar, E., Grossi, D., & Rădulescu, R. (2025). Learning in public goods games: The effects of uncertainty and communication on cooperation. *Neural Computing & Applications*, 37, 18899–18932. <https://doi.org/10.1007/s00521-024-10530-6>
- Panciroli, C., & Macaudo, A. (2021). *Intelligenza artificiale in una prospettiva educativo-didattica*. Bononia University Press.
- Panciroli, C., & Rivoltella, P. C. (2023). *Pedagogia algoritmica: Per una riflessione educativa sull'intelligenza artificiale*. Scholè.
- Pasta, S., & Rondonotti, M. (2020). Competenze mediaeducative e tecnologie di comunità: Nuove direzioni per il lavoro socioeducativo territoriale. *Rivista Italiana di Educazione Familiare*, 17(2), 377–398. <https://doi.org/10.13128/rief-9557>
- Perera, I., de Nijs, F., & García, J. (2025). Learning to cooperate against ensembles of diverse opponents. *Neural Computing & Applications*, 37, 18835–18849. <https://doi.org/10.1007/s00521-024-10511-9>
- Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. *JBIM Evidence Synthesis*, 18(10), 2119–2126. <https://doi.org/10.11124/JBIES-20-00167>
- Pollock, D., Peters, M. D. J., Khalil, H., McInerney, P., Alexander, L., Tricco, A. C., Evans, C., Brandão de Moraes, É., Godfrey, C. M., Pieper, D., Saran, A., Stern, C., & Munn, Z. (2023). Recommendations for the extraction, analysis, and presentation of results in scoping reviews. *JBIM Evidence Synthesis*, 21(3), 520–532. <https://doi.org/10.11124/JBIES-22-00123>
- Rachum, R., Nakar, Y., Tomlinson, B., Alon, N., & Mirsky, R. (2024). Emergent dominance hierarchies in reinforcement learning agents. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS '24)* (pp. 2426–2428). International Foundation for Autonomous Agents and Multiagent Systems. <https://dl.acm.org/doi/10.5555/3635637.3663182>
- Ranieri, M. (2020). *Tecnologie per educatori socio-pedagogici: Metodi e strumenti*. Carocci Faber.
- Reason, P., & Bradbury, H. (2008). *The SAGE handbook of action research: Participative inquiry and practice* (2nd ed.). SAGE.
- Redecker, C., & Punie, Y. (2017). *Digital competence of educators*. Joint Research Centre. <https://dx.doi.org/10.2760/159770>
- Ribeiro-Navarrete, B., Saura, J. R., & Simón-Moya, V. (2024). Setting the development of digitalization: State-of-the-art and potential for future research in cooperatives. *Review of Managerial Science*, 18(5), 1459–1488. <https://doi.org/10.1007/s11846-023-00663-8>
- Rivoltella, P. C. (2017). *Media education: Idea, metodo, ricerca*. ELS La Scuola.
- Sætra, H. S., Borgebund, H., & Coeckelbergh, M. (2022). Avoid diluting democracy by algorithms. *Nature Machine Intelligence*, 4(10), 804–806. <https://doi.org/10.1038/s42256-022-00537-w>
- Saheb, T., & Saheb, T. (2024). Mapping ethical artificial intelligence policy landscape: A mixed method analysis. *Science and Engineering Ethics*, 30(2), Article 9. <https://doi.org/10.1007/s11948-024-00472-6>
- Sato, E., & Osawa, H. (2020). Reducing partner's cognitive load by estimating the level of understanding in the cooperative game Hanabi. In T. Cazenave, J. van den Herik, A. Saffidine, & I.-C. Wu (Eds.), *Advances in computer games: ACG 2019* (Lecture Notes in Computer Science, Vol. 12516, pp. 11–23). Springer. https://doi.org/10.1007/978-3-030-65883-0_2
- Schrepel, T., & Potts, J. (2024). Measuring the openness of AI foundation models: Competition and policy implications. *Information & Communications Technology Law*, 34(3), 279–304. <https://doi.org/10.1080/13600834.2025.2461953>
- Sen, A. (1999). *Development as freedom*. Oxford University Press.
- Shi, M., Wang, G., Zhang, S., Jiang, J., & Xiang, Y. (2025). AISCW: Artificial intelligence supported cooperative work: The new frontier of CSCW. In H. Sun, H. Fan, Y. Gao, X. Wang, D. Liu, B. Du, & T. Lu (Eds.), *Computer supported cooperative work and social computing: ChineseCSCW 2024* (Communications in Computer and Information Science, Vol. 2343, pp. 3–17). Springer.

- https://doi.org/10.1007/978-981-96-2373-0_1
- Shi, Y., Yu, K., Dong, Y., & Chen, F. (2026). Large language models in education: A systematic review of empirical applications, benefits, and challenges. *Computers and Education: Artificial Intelligence*, 10, Article 100529. <https://doi.org/10.1016/j.caeai.2025.100529>
- Simões, M. A. C., da Silva, R. M., & Nogueira, T. A. (2020). Dataset schema for cooperative learning from demonstration in multi-robot systems. *Journal of Intelligent & Robotic Systems*, 99(3), 589–608. <https://doi.org/10.1007/s10846-019-01123-w>
- Stathis, G., & van den Herik, J. (2025). Ethical and preventive legal technology. *AI Ethics*, 5(2), 1069–1086. <https://doi.org/10.1007/s43681-023-00413-2>
- Steinbock, J. (2024). AI ethics as subordinated innovation network. *AI & Society*, 39(4), 1995–2007. <https://doi.org/10.1007/s00146-023-01658-5>
- Stucke, M. E. (2022). *Breaking away: How to regain control over our data, privacy, and autonomy*. Oxford University Press. <https://doi.org/10.1093/oso/9780197617601.003.0001>
- Su, E., Raffaele, W., Mathieson, L., & Wang, Y. (2024). Better understanding of humans for cooperative AI through clustering. In *2024 IEEE Conference on Games (CoG)* (pp. 1–8). IEEE. <https://doi.org/10.1109/CoG60054.2024.10645647>
- Tartaro, A., Panai, E., & Cocchiato, M. Z. (2024). AI risk assessment using ethical dimensions. *AI Ethics*, 4(1), 105–112. <https://doi.org/10.1007/s43681-023-00401-6>
- Taylor, R. R., Murphy, J. W., Hoston, W. T., & Senkaiahliyan, S. (2025). Democratizing AI in public administration: Improving equity through maximum feasible participation. *AI & Society*, 40(5), 3653–3662. <https://doi.org/10.1007/s00146-024-02120-w>
- Tomašev, N., Cornebise, J., Hutter, F., Mohamed, S., Picciariello, A., Connelly, B., Belgrave, D. C. M., Ezer, D., van der Haert, F. C., Mugisha, F., Abila, G., Arai, H., Almiraat, H., Proskurnia, J., Snyder, K., Otake-Matsuura, M., Othman, M., Glasmachers, T., de Wever, W., Teh, Y. W., Khan, M. E., De Winne, R., Schaul, T., & Clopath, C. (2020). AI for social good: Unlocking the opportunity for positive impact. *Nature Communications*, 11(1), Article 2468. <https://doi.org/10.1038/s41467-020-15871-z>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., ... Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M18-0850>
- Trinidad, J. E. (2025). From accountability to algorithms: Interorganizational learning and the transformation of quantification in education. *Qualitative Sociology*, 48(2), 243–277. <https://doi.org/10.1007/s11133-024-09590-w>
- Tsamados, A., Floridi, L., & Taddeo, M. (2024). Human control of AI systems: From supervision to teaming. *AI and Ethics*, 5, 1535–1548. <https://doi.org/10.1007/s43681-024-00489-4>
- Veitch, E., Dybvik, H., Steinert, M., & Alsos, O. A. (2024). Collaborative work with highly automated marine navigation systems. *Computer Supported Cooperative Work*, 33(1), 7–38. <https://doi.org/10.1007/s10606-022-09450-7>
- Verdegem, P. (2022). Dismantling AI capitalism: The commons as an alternative to the power concentration of Big Tech. *AI & Society*, 39, 727–737. <https://doi.org/10.1007/s00146-022-01437-8>
- von Schenk, A., Klockmann, V., & Köbis, N. (2023). Social preferences towards humans and machines: A systematic experiment on the role of machine payoffs. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.4145868>
- Wang, L., Fu, F., & Chen, X. (2024). Mathematics of multi-agent learning systems at the interface of game theory and artificial intelligence. *Science China Information Sciences*, 67(6), Article 166201. <https://doi.org/10.1007/s11432-024-3997-0>
- Westphal, K. K., Regoeczi, W., Masotya, M., Vazquez-Westphal, B., Lounsbury, K., McDavid, L., Lee, H., Johnson, J., & Ronis, S. D. (2021). From Arksey and O'Malley and beyond: Customizations to enhance a team-based, mixed approach to scoping review methodology. *MethodsX*, 8, Article 101375. <https://doi.org/10.1016/j.mex.2021.101375>
- Woods, D. (2025). Stag hunt in the digital wilds: Legitimizing global AI governance amidst diverse terrains. *Fudan Journal of the Humanities and Social Sciences*, 19, 65–96. <https://doi.org/10.1007/s40647-025-00438-3>
- Xu, Z. (2022). From algorithmic governance to govern algorithm. *AI & Society*, 39(3), 1141–1150. <https://doi.org/10.1007/s00146-022-01554-4>
- Zhang, L., Wu, J., Mumtaz, S., Li, J., Gacanin, H., & Rodrigues, J. J. P. C. (2019). Edge-to-edge cooperative artificial intelligence in smart cities with on-demand learning offloading. In *2019 IEEE Global Communications Conference (GLOBECOM)* (pp. 1–6). IEEE. <https://doi.org/10.1109/GLOBECOM38437.2019.9013878>
- Zheng, X., Zhou, G., & Zeng, D. D. (2023). Platform governance in the era of AI and the digital economy. *Frontiers of Engineering Management*, 10(1), 177–182. <https://doi.org/10.1007/s42524-022-0241-1>
- Zygmuntowski, J. O. (2023). Data governance in a trilemma: A qualitative analysis of rights, values, and goals in building data commons. *Digital Society*, 2(2), Article 30. <https://doi.org/10.1007/s44206-023-00058-y>