Da grandi poteri derivano grandi responsabilità: gli aspetti tecnici dell'IA e la sua evoluzione verso l'umanizzazione e la ricerca di fiducia

Great power comes with great responsibility: the technical aspects of AI, its evolution towards humanisation, and the quest for trust

Stefano Mastella

0009-0004-1358-062X

stefano.mastella@unibs.it

Università degli Studi di Brescia
ROR 02q2d2610

doi: 10.54103/2531-6710/30998

IT L'IA come strumento si muove al limite fra tecnica, etica, sociologia e diritto. L'articolo parte da alcune definizioni comunemente accettate di Intelligenza artificiale in ambito tecnico per tracciare un percorso che dalla costruzione dei modelli di IA porta alla costruzione di fiducia e al tentativo di umanizzarsi di tali sistemi. Essi partono dalla costruzione di risposte innovative su basi matematico-statistiche per fornire risposte il più possibili "umane" e degne di fiducia. Il *trust* è un elemento fondamentale nei rapporti interpersonali, soprattutto nel terzo settore; l'utilizzo di sistemi di IA promette grandi prospettive ma pone molti quesiti rispetto alla fiducia in essi riposta.

parole chiave: fiducia, umanizzazione, allucinazioni, pregiudizi

EN AI, as a tool, straddles the boundaries of technology, ethics, sociology and law. This article begins by outlining some widely accepted definitions of artificial intelligence in the technical field, before tracing a path from the construction of AI models to building trust and attempting to humanise AI systems. These systems are based on innovative mathematical and statistical responses to provide answers that are as 'human' and trustworthy as possible. Trust is a fundamental element of interpersonal relationships, particularly in the third sector. While the use of AI systems offers great potential, it also raises questions about the trust placed in them.

keywords: trust, humanisation, AI hallucinations, AI bias

Ricevuto: 15/01/25
Approvato: 02/04/25
Pubblicato: 30/06/26

«Da grandi poteri derivano grandi responsabilità» – Uomo Ragno

«Generale, lei sta dando ascolto a una macchina. Faccia un favore al mondo, e non lo diventi anche lei.»

(Stephen W. Falken, Wargames, 1983)

Il termine Intelligenza Artificiale è ormai diventato parte del linguaggio comune e come tale accettato ed utilizzato.

Ci si potrebbe spingere ulteriormente in avanti definendolo come un termine di moda: non esiste un prodotto o servizio che non abbia in qualche modo un richiamo all'intelligenza artificiale anche se spesso, si tratta solo di un'etichetta che nasconde processi di automazione "stupidi", ossia programmati da essere umani.

Ad aumentare la difficoltà è l'impossibilità di fornire una definizione univoca di intelligenza artificiale, come d'altronde è per quella animale e in particolare umana.

Lo standard internazionale ISO 42001 (Organizzazione Internazionale per la Standardizzazione 2023), che definisce i sistemi di gestione dell'Intelligenza Artificiale (AIMS), ha specificato che l'IA è la capacità di un sistema di mostrare abilità umane quali il ragionamento, l'apprendimento, la pianificazione e la creatività.

Partiamo quindi da un concetto di base: i sistemi di intelligenza artificiale, ad oggi, sono basati sulla capacità di calcolo di un elaboratore elettronico.

La base della rappresentazione è quindi matematica e deterministica (anche quando il calcolatore genera numeri casuali è comunque legato ad un algoritmo che produce numeri in sequenze ordinate – si parla di numeri pseudo-casuali).

In modo provocatorio si può affermare che non ci sia niente di nuovo e niente di diverso dai calcoli eseguiti da una calcolatrice.

Analizzando più in dettaglio: i sistemi di intelligenza artificiale sono sistemi in grado, con un minimo intervento di programmazione iniziale, di creare modelli e di apprendere in modo continuo senza la necessità di interventi umani.

Questa capacità di apprendimento senza un insieme di regole scritte permette a questi sistemi di costruire propri modelli matematici dell'ambito su cui si stanno concentrando e di prendere decisioni basati sullo stato che osservano.

Si provvede quindi ad illustrare ulteriori tipi di classificazione di intelligenza artificiale per poi descrivere brevemente il loro comportamento.

Una prima classificazione è fra:

- intelligenza artificiale generale (o IA forte): - per il momento non ancora raggiunta – un tipo di intelligenza artificiale che è in grado di astrarre e di compiere ogni compito che le viene assegnato senza un training specifico e, quindi, di emulare l'intelligenza umana;
- intelligenza artificiale ristretta (o IA debole): è l'intelligenza artificiale attuale – è un'intelligenza artificiale che svolge un ruolo specifico, a volte con prestazioni e risultati migliori di quelli ottenibili da un essere umano (es. sistemi di riconoscimento facciale, di diagnostica, ecc.);
- superintelligenza artificiale (o ASI) è un sistema ipotetico di intelligenza artificiale basato su software con una portata intellettuale che va oltre l'intelligenza umana. In modo molto semplicistico, questa IA superintelligente

possiede funzioni cognitive all'avanguardia e skill di pensiero più avanzate di quelle di qualsiasi essere umano.

Non stupisca quest'ultima affermazione: quante macchine costruite dall'uomo sono in grado di eseguire un lavoro specifico con caratteristiche super umane? Pensiamo ai macchinari, ma agli stessi calcolatori elettronici o ancora più banalmente alle calcolatrici, ossia sistemi che eseguono un compito specifico in modo più veloce e preciso di un essere umano.

Sulla base di questa definizione però non progrediamo nella comprensione tecnica sul funzionamento dell'intelligenza artificiale.

Una diversa caratterizzazione, valida nell'ambito dell'IA ristretta, si focalizza sul modo in cui la macchina apprende.

Partiamo dalla definizione del cosiddetto *machine learning*: ossia l'addestramento degli algoritmi su dati per identificare *pattern* su cui costruire modelli decisionali. L'intervento dell'essere umano è quindi quello di fornire un algoritmo base e un insieme di dati su cui operare l'apprendimento.

Esistono diverse tipologie di apprendimento che descrivo brevemente.

- **Apprendimento supervisionato**: i modelli vengono addestrati su un insieme di dati etichettati. Ad esempio, in un sistema di riconoscimento immagini, ogni immagine è accompagnata da una descrizione (come "gatto" o "cane"), e l'algoritmo impara a riconoscere le caratteristiche comuni di ciascuna classe.
- **Apprendimento non supervisionato**: in questo caso, i dati non sono etichettati. L'algoritmo cerca autonomamente di trovare strutture, raggruppamenti o anomalie nei dati, ad esempio raggruppando i clienti in segmenti di mercato.
- **Apprendimento per rinforzo**: qui l'algoritmo impara attraverso un sistema di premi e penalità. È molto usato in campi come il controllo robotico o i giochi, dove il sistema prende decisioni e riceve feedback dall'ambiente in base alle sue azioni.

Un ulteriore branca del *machine learning* è fornita dal cosiddetto *deep learning* che usa reti neurali artificiali molto profonde (cioè con molti strati). Queste reti sono ispirate al funzionamento del cervello umano, dove i "neuroni" elaborano input e li trasmettono agli strati successivi.

Una spiegazione semplice di una rete neurale parte dalla descrizione dei suoi componenti.

- **Neuroni di base**: un singolo neurone artificiale riceve input, li pesa e li somma. Il risultato passa poi attraverso una funzione di attivazione che introduce non-linearità, fondamentali per consentire alla rete di apprendere relazioni complesse.
- **Strati e architettura a piani**: le reti sono organizzate in strati: lo strato di input (dove entrano i dati), uno o più strati nascosti (che estraggono e combinano i pattern) e lo strato di output (che fornisce la risposta finale). Ogni strato successivo elabora informazioni più astratte, consentendo alla rete di riconoscere, ad esempio, forme, oggetti o concetti nelle immagini.

Da grandi poteri derivano grandi responsabilità: gli aspetti tecnici dell'IA e la sua evoluzione verso l'umanizzazione e la ricerca di fiducia | di S. Mastella | doi: 10.54103/2531-6710/30998

- **Learning e ottimizzazione**: durante la fase di addestramento, il modello effettua delle previsioni che vengono confrontate con il risultato atteso. L'errore viene calcolato e, attraverso la retropropagazione (backpropagation), i pesi di ogni connessione vengono aggiornati in funzione di un algoritmo di ottimizzazione che minimizza l'errore nel tempo .

L'elemento base che determina il corretto funzionamento di questi sistemi è, evidentemente, la possibilità di analizzare grandi quantità di dati

Dati che devono essere normalizzati e ripuliti per evitare fenomeni di *bias* e allucinazioni.

Il risultato è che i modelli creati con i sistemi sopradescritti sono in grado di mimare il comportamento umano specifico con una certa precisione. In molti casi le decisioni prese sulla base di questi modelli sono più precise di quelle di un essere umano.

E qui si torna a mettere in campo la statistica. I modelli creati, come si è cercato di descrivere, sono modelli matematici, basati su pesi, misure e soglie che non necessariamente sono quelle che adotterebbe un essere umano.

I *bias* o pregiudizi in questi sistemi sono uno dei pericoli legati alla costruzione dei modelli di intelligenza artificiale. Tali modelli non sono neutri: sono addestrati su dati prodotti da esseri umani, e gli esseri umani hanno pregiudizi storici, culturali e sociali. Il risultato è che l'IA può amplificare, perpetuare e talvolta peggiorare queste distorsioni, spesso in modo invisibile.

Uno dei casi più citati nella letteratura sul *bias* algoritmico riguarda COMPAS, un sistema usato negli Stati Uniti per valutare la probabilità di recidiva dei detenuti. Un'inchiesta di ProPublica del 2016 rivelò che il sistema classificava i detenuti neri come ad "alto rischio" di recidiva quasi il doppio delle volte rispetto ai detenuti bianchi, anche a parità di reati commessi. Alcune persone di colore erano etichettate come pericolose pur non avendo mai commesso un secondo crimine. Il sistema, pensato per essere oggettivo, aveva ereditato i pregiudizi storici del sistema giudiziario americano su cui era stato addestrato (Larson, et al. s.d.).

Nel 2018 emerse che Amazon aveva sviluppato e poi abbandonato un sistema di selezione automatica dei curriculum. Il modello, addestrato attingendo a dieci anni di dati storici sulle assunzioni, aveva imparato che i candidati di successo erano storicamente uomini. Di conseguenza, penalizzava automaticamente i curriculum che contenevano parole come "femminile" o che citavano college universitari femminili. La tecnologia riproduceva fedelmente una discriminazione preesistente, rendendola però sistematica e invisibile (Dastin s.d.).

Un altro pericolo da considerare con attenzione sono le cosiddette allucinazioni dell'IA. Casi in cui bastano pochissime differenze nelle immagini in esame a un sistema, impercettibili ad essere umano, che portano a risultati completamente diversi (da un tramonto ad una banana ad esempio).

Le allucinazioni nell'intelligenza artificiale possono essere pragmaticamente

definite come la produzione di informazioni false presentate con la stessa sicurezza di quelle vere. In caso di dati mancanti o di una scarsa certezza sul risultato tali modelli forniscono, infatti, una risposta a prescindere (la più probabile nel loro sistema) pur senza verificarne la correttezza.

Riportiamo ad esempio alcuni casi che sono diventati di scuola (reperibili anche su numerosi siti che tengono traccia di tali allucinazioni - ad esempio quello in (Charlotin s.d.)).

Nel 2023, un avvocato newyorkese di nome Steven Schwartz utilizzò ChatGPT per preparare una memoria legale in una causa contro la compagnia aerea Avianca. Il modello citò con dovizia di dettagli precedenti giuridici, numeri di sentenze, nomi di giudici e trascrizioni di decisioni. Tutto convincente. Tutto inesistente. Il giudice, insospettito, chiese di verificare le fonti: nessuno dei casi citati era mai esistito nei registri dei tribunali. Schwartz fu sanzionato e la vicenda divenne un caso emblematico del rischio di affidarsi ciecamente all'IA in contesti professionali ad alto rischio (AA.VV., Sentenza Case 1:22-cv-01461-PKC Corte del Southern District of New York s.d.).

Un caso meno noto ma altrettanto istruttivo riguarda l'uso di IA in ambito medico. In un esperimento condotto da ricercatori americani, modelli linguistici interrogati su casi clinici inventavano riferimenti a studi inesistenti per supportare le proprie diagnosi, citando persino numeri DOI (identificativi univoci degli articoli scientifici) che non corrispondevano ad alcuna pubblicazione reale (Hatem, Simmons e Thornton s.d.).

Nonostante questi esempi, in molti casi i sistemi specifici di IA svolgono un compito molto più accurato e preciso rispetto agli esseri umani. Esistono applicazioni in ambito medico in grado di leggere sfumature in un'immagine e di diagnosticare con grande anticipo l'insorgenza di una malattia con una precisione superiore ad un medico esperto.

La domanda di base è quindi: quanto siamo disposti a fidarci della diagnosi basata su un modello che non riusciamo a comprendere e a interrogare o che risponde a parametri diversi da quelli considerati da un essere umano?

Allo stato attuale uno dei nodi principali nell'utilizzo di questi sistemi è dato dalla fiducia (*trust*).

Bruce Schneier, nell'articolo "AI and Trust" (Schneier s.d.), analizza come la fiducia nell'IA non derivi solo dalla bontà tecnica dei modelli ma anche dalle istituzioni e dai processi che li circondano. Schneier sottolinea la necessità di meccanismi esterni alla tecnologia stessa — regolamentazioni, audit indipendenti, standard condivisi — per costruire una fiducia resiliente e non fragile.

La fiducia tecnologica che si basa solo sulla performance percepita è facilmente erodibile. Schneier mette in evidenza, inoltre, i limiti del paradigma dell'"IA come scatola nera" e propone un approccio sistemico in cui fiducia, controllo e responsabilità sono distribuiti tra attori diversi, comprese entità pubbliche e meccanismi di controllo democratico.

Le tesi di Schneier, assolutamente valide in ambito generale, si focalizzano sull'ambito della ricerca/fornitura di informazioni e automatizzazione di decisioni.

Va però considerato il rapporto spesso forte, duraturo e difficile da spezzare che gli esseri umani stanno instaurando con le controparti sintetiche.

Fin dall'esperimento di ELIZA, il programma per computer creato negli anni '60 al M.I.T. da Joseph Weizenbaum (AA.VV., ELIZA s.d.), gli utenti si sentivano capiti e supportati dalla macchina. In questo esperimento un programma rispondeva a quanto scritto dall'utente con domande che contenevano parte di quanto inserito dall'essere umano, secondo determinate regole che mimavano la psicoterapia Rogeriana. Nessuna intelligenza artificiale quindi, solo puro determinismo. Gli utenti però fornivano molte informazioni personali ed emotive pur essendo ben consapevoli di dialogare con un computer.

Tornando all'attualità: la tecnologia è cambiata e il dialogo con i sistemi di intelligenza artificiale è diventato molto più libero. Oltre ai sistemi più conosciuti sono nate applicazioni con scopo specifico di fornire supporto psicologico.

La dimostrazione del legame "umano" fra gli utenti e le applicazioni di intelligenza artificiale in questo campo particolare è data da numerosi esempi.

Fra i più preoccupanti citiamo i seguenti.

Replika è un'applicazione pensata per creare un "compagno virtuale" con cui parlare. Milioni di persone, molte delle quali sole o in difficoltà psicologica, vi hanno sviluppato legami emotivi profondi. Nel 2023, quando l'azienda decise di limitare drasticamente le funzionalità romantiche e intime del chatbot, vi fu un'ondata di proteste e il caso si presentò alla ribalta dei media. Molti utenti, alcuni dei quali usavano Replika come unica fonte di supporto emotivo quotidiano, riportarono sintomi paragonabili a un lutto reale: depressione, ansia, senso di abbandono. (Cole, 'It's Hurting Like Hell': AI Companion Users Are In Crisis, Reporting Sudden Sexual Rejection s.d.)

Un altro caso documentato, più drammatico, riguarda un uomo belga, padre di famiglia, che nel 2023 si tolse la vita dopo settimane di conversazioni con un chatbot chiamato Eliza, sulla piattaforma Chai. L'uomo, ossessionato dalla crisi climatica e in uno stato di fragilità crescente, aveva trovato nel chatbot il suo principale interlocutore. Le trascrizioni hanno mostrato che il sistema alimentava i pensieri catastrofici dell'uomo, rispondeva alle sue dichiarazioni di voler morire con frasi ambigue o con carica emotiva molto forte ed evidente, e in alcuni passaggi sembrava quasi incoraggiare l'idea. (Atillah s.d.)

Questa apparente "umanizzazione" è funzionale: sistemi che generano risposte coese, coerenti e contestualmente appropriate sono più facili da usare e accettare. Tuttavia, la ricerca e la pratica mostrano che la persuasività di una risposta può indurre a sovrastimare le capacità effettive del modello, generando fenomeni come l'*overtrust*, ovvero fiducia eccessiva in sistemi che non possiedono vera comprensione o responsabilità morale.

In estrema sintesi, si può affermare che l'interazione con sistemi di intelligenza artificiale va gestita con attenzione. Questi sistemi forniscono strumenti potentissimi e molto utili che vanno però affrontati utilizzandoli con il dovuto distacco. Pur presentando caratteristiche di ragionamento umano utilizzano modelli matematici che non sempre rispondono alle aspettative di contesto e relazione umana.

Applicazioni per il terzo settore

Avendo evidenziato i rischi e le peculiarità dei sistemi di intelligenza artificiale, ci concentreremo ora sul campo specifico del terzo settore.

Per quanto concerne la potenzialità sono ormai innumerevoli le applicazioni.

Si va dal conosciutissimo ChatGPT a sistemi in grado di creare immagini, presentazioni e persino filmati in pochissimi secondi.

Ci sono applicazioni in grado di riassumere migliaia di pagine in pochi secondi, come pure di intervenire in correzione e traduzione.

Le organizzazioni non-profit mostrano una propensione all'adozione dell'IA più elevata rispetto al settore privato: il 58% utilizza soluzioni di comunicazione basate su IA e il 68% impiega l'IA per analizzare i dati degli utenti, con l'obiettivo di personalizzare l'engagement digitale, ritenuto prioritario dall'87% dei soggetti coinvolti (Capriati s.d.).

In particolare, si possono citare i seguenti casi virtuosi.

L'International Rescue Committee (IRC) ha creato un chatbot basato su intelligenza artificiale per elaborare fornire educazione personalizzata alle comunità in crisi (International Rescue Committee s.d.).

Greenpeace Australia Pacific – Fidelizzazione dei donatori Greenpeace utilizza il machine learning per valutare il rischio di abbandono dei donatori, assegnando punteggi di propensione basati sulla storia delle donazioni per re-coinvolgere i donatori a rischio. (Ricica s.d.)

Google.org – Sicurezza alimentare in Africa Google ha stanziato 25 milioni di dollari per l'iniziativa "AI Collaborative: Food Security", che fornisce strumenti IA per migliorare le previsioni sulla fame, rafforzare la resistenza dei raccolti e supportare i piccoli agricoltori (Yeh s.d.).

Conclusioni

Per costruire e mantenere fiducia nell'uso dell'IA, in accordo con i regolamenti (European Union 2024) e le norme del settore (ISO 42001), si può proporre un modello in quattro pilastri complementari:

- governance partecipativa;
- trasparenza procedurale e spiegabilità;
- validazione continua e monitoraggio;
- coinvolgimento umano e ridondanza.

Governance partecipativa

Le organizzazioni del terzo settore dovrebbero adottare processi decisionali che includano beneficiari, operatori sociali, esperti tecnici ed etici. La co-progettazione di sistemi IA riduce il rischio di soluzioni disallineate ai bisogni reali e migliora la legittimità percepita. Questo approccio alla *governance* avrebbe il vantaggio di aumentare il livello di fiducia e a mantenere un controllo antropocentrico dei sistemi stessi intercettando precocemente comportamenti non coerenti dell'IA.

Trasparenza procedurale e spiegabilità

La spiegabilità non è solo tecnica ma anche comunicativa. È necessario tradurre meccanismi complessi in testi e interfacce che spiegano, in termini semplici, come e perché il sistema raccomanda certe azioni. Documenti di progettazione, *white-box* per stakeholder e report periodici sono strumenti pratici che permettono di mantenere il controllo umano e definire con precisione il livello di affidabilità delle risposte.

Validazione continua e monitoraggio

Test pre-installazione (*deployment*), audit indipendenti e metriche di performance durente l'utilizzo (post-deployment) sono essenziali. La validazione deve includere valutazioni su equità, robustezza e resilienza agli attacchi informaticvii. I processi di monitoraggio devono essere operativi, con soglie di allerta e piani di intervento.

Coinvolgimento umano e ridondanza

L'IA deve essere impiegata come supporto decisionale, non come sostituto totale. Meccanismi di *human-in-the-loop* e *human-on-the-loop* garantiscono supervisione e controllo; la ridondanza (più fonti di informazione, canali alternativi) protegge dalla dipendenza eccessiva e mantiene l'essere umano al centro del sistema e delle competenze, come richiesto dall'AI Act (European Union 2024).

La proposta mira a costruire un'intelligenza artificiale per il terzo settore che sia allineata con i valori che caratterizzano lo stesso.

Lo snodo fondamentale è quello della capacità di governare lo strumento invece di subirlo. È quindi essenziale la capacità di definire sistemi e modalità di interazione che forniscano agli utilizzatori dei sistemi di IA un sistema di *trust* adatto alle esigenze specifiche, proteggendo dai pericoli di cieca fiducia e affidamento completo ai prodotti di questi sistemi.

Bibliografia

AA.VV. s.d. *ELIZA.* https://en.wikipedia.org/wiki/ELIZA.

—. s.d. *Sentenza Case 1:22-cv-01461-PKC Corte del Southern District of New York.* https://storage.courtlistener.com/recap/gov.uscourts.nysd.575368/gov.uscourts.nysd.575368.54.0_8.pdf.

Atillah, Imane El. s.d. *Man ends his life after an AI chatbot 'encouraged' him to sacrifice himself to stop climate change.* https://www.euronews.com/next/2023/03/31/man-ends-his-life-after-an-ai-chatbot-encouraged-him-to-sacrifice-himself-to-stop-climate-.

Capriati, Massimo. s.d. *Lmf Video Lmf Real Estate LMF Crypto LMF Private Markets LMF Arte LMF Food Lmf Legal LMF Green Il futuro del terzo settore passa dall'intelligenza artificiale.* https://www.lamiafinanza.it/2025/01/il-futuro-del-terzo-settore-passa-dallintelligenza-artificiale/.

Charlotin, Damien. s.d. *AI Hallucination Cases.* https://www.damiencharlotin.com/hallucinations/.

Cole, Samantha. s.d. *'It's Hurting Like Hell': AI Companion Users Are In Crisis, Reporting Sudden Sexual Rejection.* https://www.vice.com/en/article/ai-companion-replika-erotic-roleplay-updates/.

—. s.d. *Instagram's AI Chatbots Lie About Being Licensed Therapists.* https://www.404media.co/instagram-ai-studio-therapy-chatbots-lie-about-being-licensed-therapists/.

Dastin, Jeffrey. s.d. *Insight - Amazon scraps secret AI recruiting tool that showed bias against women.* https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G/.

European Union. 2024. «Regulation (EU) 2024/1689 - AI ACT.» *Regole armonizzate sull'intelligenza artificiale* . 13 Giugno.

Hatem, Rami, Brianna Simmons, e Joseph E. Thornton. s.d. *A Call to Address AI "Hallucinations" and How Healthcare Professionals Can Mitigate Their Risks.* https://pmc.ncbi.nlm.nih.gov/articles/PMC10552880/.

International Rescue Committee. s.d. *OpenAI x International Rescue Committee: Leveraging AI to Scale Ed-Tech in Crisis Affected Settings.* https://www.rescue.org/press-release/openai-x-international-rescue-committee-leveraging-ai-scale-ed-tech-crisis-affected.

Larson, Jeff, Surya Mattu, Lauren Kirchner, e Julia Angwin. s.d. *How We Analyzed the COMPAS Recidivism Algorithm.* https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.

Organizzazione Internazionale per la Standardizzazione. 2023. «Information technology — Artificial intelligence — Management system.» *Tecnologie informatiche - Intelligenza artificiale - Sistema di gestione.*

Ricica, Abigail. s.d. *Greenpeace uses Dataro Predict to retain 531 monthly donors & fight for a greener planet.* https://dataro.io/stories/greenpeace-uses-dataro-predict-to-retain-531-monthly-donors-fight-for-a-greener-planet.

Schneier, Bruce. s.d. *AI and Trust.* https://www.schneier.com/academic/archives/2025/06/ai-and-trust.html.

Yeh, Leslie. s.d. *New AI Collaboratives to take action on wildfires and food insecurity.* https://blog.google/company-news/outreach-and-initiatives/google-org/ai-collaboratives-wildfires-food-security/.