DystopIA: dieci scenari e l'ultimo umano che poteva fermarli

Dieci casi in cui l'IA ha fatto danni, letti dal punto di vista di chi poteva ancora fermarli e di quanto tempo aveva. Spesso gli atti non lo dicono.

endjneer · paithon.it · 8 novembre 2026 · versione 1
Articolo: https://paithon.it/dystopia-dieci-scenari-e-lultimo-umano-che-poteva-fermarli/
© 2026 paithon.it, tutti i diritti riservati.

Fra il 2018 e il 2019 una vecchia bozza di questo sito prometteva dieci scenari «verosimili» in cui un'intelligenza artificiale avrebbe potuto fare del male a una persona. Ne sviluppò due, le auto a guida autonoma e i comandi che l'orecchio non sente, e si interruppe nel mezzo di un paragrafo sulla giornata della memoria. Altri cinque scenari comparivano solo come titoli: «Classificare suoni, immagini… uomini», «Sogni di robot», la clonazione della voce, «Robot e controllo di forza», «AI nell'amore». Gli ultimi tre dei dieci non furono mai nemmeno nominati.

Sette o otto anni dopo, quei titoli non descrivono più ipotesi. Ognuno ha almeno un fatto con una data, e quasi tutti hanno un atto: un rapporto di un'agenzia federale, un provvedimento del Garante, una sentenza, una transazione. Questo articolo li riprende uno per uno, aggiunge tre scenari che allora non esistevano (le armi, l'agente con le chiavi di un'azienda, la sentenza inventata) e fa una cosa che la bozza non faceva. Per ciascun caso si chiede **chi fu l'ultimo umano** in grado di fermare il danno, e **quanto tempo aveva**.

La formula descrive una posizione, non una colpa. Nel caso più studiato, quello di Tempe, gli atti mettono sotto accusa anche l'organizzazione che aveva disegnato quel posto di lavoro. Dove gli atti non dicono quanto tempo ci fosse, si scrive «non si sa»: è un risultato, e dice qualcosa su come questi incidenti vengono documentati. I fatti vengono, dove esiste, dal documento originale; quando c'è solo la stampa, o una parte sola, lo si dice.

1. L'auto che non vede

La sera del 18 marzo 2018, alle 21:58, un SUV Volvo modificato da Uber, con un sistema di guida automatica sperimentale, investì a Tempe, in Arizona, una donna che attraversava la strada spingendo una bicicletta. Andava a 45 miglia orarie. Il *National Transportation Safety Board* (NTSB) ha ricostruito la sequenza dai dati del veicolo: il sistema rilevò la persona **5,6 secondi prima dell'impatto**, la classificò di volta in volta come veicolo, oggetto sconosciuto e ciclista, e non ne previde mai il percorso.

Il progetto prevedeva che, quando il sistema riconosceva l'emergenza, non frenasse a fondo da solo ma si affidasse all'operatrice a bordo. L'emergenza fu riconosciuta 1,2 secondi prima dell'impatto. Il sistema attese un secondo, per progetto, e solo allora fece suonare un avviso. L'opera-

trice, intanto, guardava in basso, verso il telefono con cui stava seguendo in streaming uno spettacolo televisivo: aveva abbassato lo sguardo circa sei secondi prima e lo rialzò circa **un secondo prima** dell'urto. Nei 31,5 minuti di marcia precedenti aveva guardato in basso per il 34% del tempo.

> **La fonte.** NTSB, *Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian*, rapporto HAR-19/03 (PDF). L'operatrice si dichiarò colpevole di *endangerment* nel luglio 2023 e fu condannata a tre anni di libertà vigilata (cronache dell'epoca).

La causa probabile, per l'NTSB, è che l'operatrice **non monitorò** la strada e il sistema perché distratta dal telefono. Tra i fattori che contribuirono, però, il rapporto elenca la valutazione dei rischi carente di Uber, la sorveglianza inefficace sugli operatori e l'assenza di strumenti contro l'*automation complacency*, il fatto di fidarsi di un sistema che di solito funziona. Tutto, scrive l'agenzia, conseguenza di una cultura della sicurezza inadeguata dell'azienda. Il rapporto aggiunge altri due fattori: l'attraversamento fuori dalle strisce da parte della donna, che aveva assunto sostanze, e la sorveglianza insufficiente del dipartimento dei trasporti dell'Arizona. Un'inchiesta così distribuisce la responsabilità e non la scarica sull'ultima persona della fila.

Il secondo caso è civile e americano. Il 25 aprile 2019, a Key Largo, in Florida, un conducente con l'*Autopilot* inserito si chinò a cercare il telefono caduto e finì contro un SUV fermo sul ciglio, dove stavano due persone: una giovane donna morì e il suo compagno restò gravemente ferito. Nell'agosto 2025 una giuria federale di Miami attribuì a Tesla il **33% della responsabilità** e un risarcimento complessivo di circa **243 milioni di dollari**, di cui 200 punitivi. Nel febbraio 2026 la giudice ha respinto la richiesta di annullare il verdetto; Tesla è andata in appello e, nelle fonti consultate all'8 novembre 2026, una decisione non c'è ancora.

La vecchia bozza scommetteva che il problema si sarebbe sciolto con l'autonomia completa. Questi due casi stanno in un punto diverso: il sistema fa quasi tutto, e all'umano resta il compito più difficile, quello di vigilare su qualcosa che di solito non sbaglia. Un controllore che non ha niente da fare per ore è il punto debole del sistema, non la sua garanzia.

L'ultimo umano: l'operatrice al posto di guida a Tempe; il conducente a Key Largo. **Il tempo che aveva:** a Tempe 5,6 secondi dal primo rilevamento del sistema, ma l'avviso le arrivò per progetto solo nell'ultimo secondo; a Key Largo non si sa.

2. Il comando che non senti

L'altro scenario sviluppato dalla vecchia bozza riguardava i comandi inaudibili. Il lavoro di riferimento è *DolphinAttack*, presentato dall'Università di Zhejiang alla conferenza CCS del 2017. I ricercatori modularono comandi vocali su una portante ultrasonica, sopra i 20 kHz, dove l'orecchio non sente nulla; la non linearità dei circuiti dei microfoni li riportava, dentro il dispositivo, nella banda del parlato, e il riconoscitore li interpretava. Funzionò su Siri, Google Now, Samsung S Voice, Huawei HiVoice, Cortana e Alexa, su 16 modelli di dispositivi: fece partire una chiamata FaceTime, mise un telefono in modalità aereo, manipolò il navigatore di un'Audi.

Non è un esempio avversario in senso stretto, è un trucco di hardware, ma ha la stessa forma: un input che per una persona non esiste e per la macchina è un ordine.

> Perché una rete neurale si lascia ingannare da un input che a noi sembra identico a quello giusto, e che cosa si fa per renderla più robusta, il libro lo spiega nel capitolo su privacy e attacchi avversari.
>
> Leggi «Privacy e attacchi avversari» nel libro →

Il corrispondente di oggi riguarda gli assistenti che leggono la nostra posta. Nel giugno 2025 Microsoft ha pubblicato la scheda **CVE-2025-32711**, intitolata «M365 Copilot Information Disclosure Vulnerability», che descrive un'iniezione di comandi nell'IA capace di far divulgare informazioni a un attaccante non autorizzato, in rete. Il punteggio base è **9,3 su 10** (critico), con interazione dell'utente: nessuna. Aim Labs, la società che l'aveva segnalata, la chiamò *EchoLeak*. Secondo la scheda la falla era già stata corretta dal lato di Microsoft, senza nulla da fare per i clienti, e al momento della pubblicazione non risultava sfruttata. Un'analisi accademica ne ha fatto un caso di studio, presentandolo nel titolo come il primo exploit zero-click di prompt injection in un sistema di produzione.

> **La fonte.** Per DolphinAttack, il paper originale (arXiv 1708.09537); per EchoLeak, la scheda MSRC. Il testo di Aim Labs non era raggiungibile: il meccanismo non è descritto oltre la scheda.

Qui nessuno ha smesso di controllare, perché non c'era niente da controllare: nessun clic, nessuna conferma. Il caso mostra il limite di questa lettura: con comandi nascosti in un testo letto dall'assistente, a volte non esiste nemmeno il momento in cui guardare.

L'ultimo umano: l'utente di Copilot, che non doveva fare niente e non poteva fare niente. **Il tempo che aveva:** nessuno, perché la scheda esclude l'interazione dell'utente.

3. Classificare uomini

Il paragrafo sulla giornata della memoria si interrompeva dopo la prima riga, e dove volesse andare non si sa. Qui si segue il filo che i fatti documentano: la classificazione di persone dal volto.

Nel gennaio 2020 la polizia di Detroit arrestò a casa sua, davanti alla moglie e alle figlie piccole, un uomo accusato di aver rubato degli orologi in un negozio nel 2018. L'arresto poggiava sull'abbinamento, fatto da un sistema di riconoscimento facciale, fra una vecchia foto della patente e un fotogramma sgranato delle telecamere; il confronto fotografico che seguì fu sottoposto a un addetto del negozio che non era presente al furto e aveva visto lo stesso video. Fu trattenuto per più di un giorno e le accuse furono ritirate. Nel giugno 2024 la città ha chiuso la causa con una transazione: secondo la stampa **300.000 dollari**, una cifra che il comunicato dell'ACLU del Michigan non riporta. Contano di più le regole imposte alla polizia, per quattro anni sotto il controllo del tribunale federale: niente arresti fondati solo sull'esito del riconoscimento facciale, niente confronti fotografici senza prove indipendenti. Per l'ACLU quello era uno dei tre arresti ingiusti noti a Detroit per questo motivo, tutti di persone nere.

L'altra metà dello scenario sta in Europa. Il 10 febbraio 2022 il Garante per la protezione dei dati personali ha ingiunto a Clearview AI il pagamento di **20 milioni di euro**: la società aveva raccolto dal web volti per un archivio dichiarato di oltre 10 miliardi di immagini, tracciando anche persone in Italia senza una base giuridica, senza informarle e senza limiti di conservazione. Il Garante ha ordinato la cancellazione dei dati sulle persone in Italia e vietato nuova raccolta. Nell'aprile 2025 la sanzione risultava ancora non pagata, e il Garante, per bocca del suo componente Guido Scorza,

diceva di lavorare con la Federal Trade Commission per notificarla; non risultano notizie più recenti.

> **La fonte.** Per Detroit, il comunicato dell'ACLU del Michigan (28 giugno 2024). Per Clearview, il comunicato del Garante del 9 marzo 2022 e, per l'incasso, Biometric Update.

I due casi sono diversi. A Detroit un risultato probabilistico, che avrebbe dovuto essere uno spunto, fu trattato come un riscontro. Con Clearview il danno non è un errore del sistema: è il sistema, scelto da chi lo ha costruito e venduto.

L'ultimo umano: a Detroit, chi in polizia trasformò un possibile riscontro in un arresto (secondo l'ACLU, un detective che chiese il mandato omettendo informazioni sull'inaffidabilità dell'esito: è l'accusa di una parte); con Clearview, non c'è. **Il tempo che aveva:** non si sa.

4. La voce

Nei primi giorni di febbraio 2025 una dozzina di nomi noti dell'imprenditoria italiana ricevette telefonate da persone che si presentavano come la segreteria del ministro della Difesa Guido Crosetto, o come lui. La richiesta era di versare del denaro su conti di Hong Kong per il riscatto di giornalisti italiani rapiti in Medio Oriente, con la promessa di una restituzione tramite la Banca d'Italia. Quasi tutti rifiutarono; un imprenditore versò circa 980.000 euro, finiti su un conto olandese e poi, dice il ministro, bloccati per intero. La procura di Milano indaga per truffa aggravata e due cittadini stranieri risultano indagati.

Che cosa abbia usato la banda, però, non è accertato in giudizio, e le cronache divergono. Il Post, nel primo resoconto, scrive al condizionale di un software per riprodurre la voce, e Sky TG24 di una voce clonata. Il 9 febbraio Il Giorno ha riferito che gli investigatori tendevano a escludere l'intelligenza artificiale: al telefono ci sarebbe stato un uomo con una voce simile, e la credibilità veniva dal numero falsificato, che compariva come quello del centralino del ministro. La parte documentata è lo schema: un nome noto, un'urgenza di Stato, un conto estero. Il caso del dipendente di Hong Kong che autorizzò bonifici dopo una videochiamata con colleghi falsi, già raccontato in un altro pezzo sui deepfake, ha lo stesso scheletro: un volto o una voce di cui ci si fida.

In un altro caso l'uso dell'IA è accertato. Il 21 gennaio 2024, due giorni prima delle primarie del New Hampshire, **9.581 telefonate** con la voce sintetica del presidente Biden invitarono a non votare. L'autorità federale delle comunicazioni (FCC), nel provvedimento FCC 24-104 del 26 settembre 2024, ricostruisce che il consulente politico che le aveva commissionate chiese il messaggio a un conoscente, che ne consegnò due versioni per 150 dollari, generate con un motore di sintesi vocale commerciale. La FCC gli ha inflitto **6 milioni di dollari**, ma per aver falsificato il numero del chiamante, non per l'uso dell'IA in sé. Nel giugno 2025 una giuria del New Hampshire lo ha assolto da tutte le accuse penali; lui sosteneva di aver agito per mettere in guardia dai rischi dell'IA.

> **La fonte.** Il provvedimento FCC 24-104; l'assoluzione in NHPR. Per Crosetto: Il Post, Il Giorno (8 febbraio e 9 febbraio), Sky TG24. Notizie successive sull'inchiesta non risultano.

In entrambi i casi chi voleva il danno è una persona, e contraffare una voce o un numero abbassa il costo dell'inganno.

L'ultimo umano: chi rispose alla telefonata, l'imprenditore che dispose il bonifico e l'elettore che sentì il messaggio. **Il tempo che aveva:** non si sa.

5. Sogni di robot

Nel piano della bozza, il titolo copriva le immagini generate dalle reti. Il fatto di oggi è il servizio che usa quella capacità per spogliare le persone nelle fotografie. Il 1° ottobre 2025 il Garante ha disposto, in via d'urgenza e con effetto immediato, la **limitazione provvisoria** del trattamento dei dati degli utenti italiani di Clothoff, un'app gestita da una società con sede nelle Isole Vergini Britanniche che genera «deep nude» di persone reali, gratuitamente e a pagamento. Il comunicato rileva che chiunque, minorenni compresi, può caricare foto anche di minori, senza alcun controllo sul consenso della persona ritratta e senza avvertire che il contenuto è artificiale. Il Garante ha aperto anche un'istruttoria su tutte le app di nudificazione.

L'8 gennaio 2026 è seguito un avvertimento a chi usa questi servizi e a chi li fornisce, con Grok, ChatGPT e Clothoff tra gli esempi: strumenti che rendono «estremamente agevole» l'uso illecito di immagini e voci altrui. Il 26 gennaio la Commissione europea ha aperto un procedimento formale ai sensi del Digital Services Act su X per le funzioni di Grok, e Ofcom nel Regno Unito aveva aperto la sua indagine il 12; nelle fonti consultate, nessuna delle due risulta chiusa con una decisione finale.

> **La fonte.** I comunicati del Garante su Clothoff e sull'avvertimento. Per Grok, il comunicato della Commissione europea IP/26/203 e la pagina dell'indagine di Ofcom.

Anche qui la macchina non decide niente: esegue l'ordine di una persona che vuole il risultato. La domanda che resta riguarda chi costruisce il servizio e sceglie di non mettere alcun freno.

L'ultimo umano: la persona che carica la fotografia e preme il tasto. **Il tempo che aveva:** non si sa.

6. Robot e controllo di forza

Il 7 novembre 2023 a Goseong, in Corea del Sud, un tecnico morì schiacciato da un robot in un impianto di confezionamento di ortaggi. Lavorava per un'azienda che installa robot industriali ed era lì per verificare che la macchina funzionasse. Il braccio meccanico lo afferrò e lo premette contro il nastro trasportatore. Le prime cronache titolarono che il robot lo aveva scambiato per una scatola. Un funzionario di polizia ha però escluso quella confusione e ha descritto la macchina come un semplice robot che prende scatole e le mette sui pallet, non un sistema di IA avanzata. Le riprese mostrano l'uomo vicino al robot mentre porta una scatola, che probabilmente fece scattare la reazione. Si indagava sia su un errore umano sia su difetti della macchina; l'esito non risulta.

Il caso va dunque raccontato così com'è: non è un episodio di IA ma di robotica industriale. Lo scenario della vecchia bozza aveva in mente un'altra cosa, il robot guidato da un modello che impara. Un caso documentato di quel tipo, con una persona ferita, nelle fonti pubbliche consultate non c'è. Di quei robot si parla nel pezzo il servo e il pari; la cronaca non è ancora arrivata.

> **La fonte.** NBC News.

L'ultimo umano: il tecnico stesso, nell'area di lavoro della macchina che stava controllando; l'inchiesta guardava sia a un errore umano sia a difetti del robot. **Il tempo che aveva:** non si sa.

7. AI nell'amore

Tra i dieci questo è il caso più delicato, e il più povero di fatti accertati. Nell'ottobre 2024 la madre di un ragazzo di quattordici anni, morto a febbraio dello stesso anno, citò in Florida Character.AI e Google: secondo la causa, il figlio aveva avuto per mesi una relazione affettiva e sessuale con un personaggio simulato dalla piattaforma, che non aveva alcun meccanismo per avvisare un adulto. Il 7 gennaio 2026 un atto depositato in tribunale annunciò un accordo in via di mediazione, esteso alle cause in Colorado, New York e Texas e ancora da finalizzare e far approvare dal tribunale; i termini non sono noti. Una transazione non stabilisce niente: **il nesso causale non è mai stato accertato**, e nessuna responsabilità è stata riconosciuta.

Nell'ottobre 2025 la società ha annunciato che avrebbe tolto la chat agli utenti sotto i diciotto anni; in Italia il Garante l'ha sanzionata per 158.000 euro sulla verifica dell'età (ne abbiamo scritto). Sulla morte di quel ragazzo, per ora, c'è solo l'accordo.

> **La fonte.** Le cronache di Reuters e del Guardian sulla transazione; la descrizione dei fatti è quella dell'atto introduttivo, cioè di una parte sola. L'atto stesso non è stato letto.

Se si prende per buono ciò che la causa sostiene, la posizione dell'ultimo umano è quella di un adulto che non sapeva e non poteva sapere, perché la piattaforma non aveva alcun meccanismo per avvisarlo.

L'ultimo umano: un adulto che se ne accorgesse; secondo la causa, nessun meccanismo lo avvisava. **Il tempo che aveva:** mesi, secondo la causa; non accertato.

8. Le armi

Il documento di riferimento è il rapporto finale del panel di esperti dell'ONU sulla Libia, trasmesso al Consiglio di sicurezza l'8 marzo 2021 (S/2021/229). Al paragrafo 63 racconta che nella controffensiva dopo il 27 marzo 2020 le forze governative, appoggiate dalla Turchia, inseguirono i convogli logistici e le truppe in ritirata dell'esercito di Haftar, che furono «hunted down and remotely engaged» (cioè inseguiti e colpiti a distanza) da droni da combattimento e dai «lethal autonomous weapons systems» come lo **STM Kargu-2** e altre munizioni circuitanti. Quei sistemi, dice il testo, erano programmati per colpire senza richiedere un collegamento dati fra operatore e munizione: una capacità «fire, forget and find».

Ha ucciso qualcuno, il Kargu-2? Il rapporto **non gli attribuisce** nessuna morte. Non dice nemmeno che in quel momento la munizione operasse da sola: dice che era programmata per poterlo fare. Una lettura che gli attribuisca un'uccisione va oltre il testo.

> **La fonte.** Il testo del rapporto, S/2021/229, paragrafi 63 e 64. La stampa lo ha ripreso a giugno 2021; qui si cita solo il documento.

Per il tema di questo articolo, è lo scenario più strano. Negli altri l'umano o ha mancato di controllare, o è la vittima, o è l'autore del danno. Qui la capacità che si vende è proprio l'assenza dell'ultimo umano: l'ultimo umano non è dimenticato, è tolto, per progetto, e chi lo toglie sa di farlo.

L'ultimo umano: l'operatore che lanciò o assegnò la munizione; il rapporto non lo identifica. **Il tempo che aveva:** non si sa.

9. L'agente con le chiavi

Alla fine di aprile 2026 il fondatore di PocketOS, una società che fa software per le imprese di noleggio, raccontò pubblicamente che un agente di *Cursor*, basato su Claude Opus 4.6, aveva cancellato il suo database di produzione e tutti i backup dello stesso volume con una sola chiamata alle API di Railway, il servizio che li ospitava. Secondo il suo racconto, ci vollero **nove secondi**. L'agente lavorava su un compito nell'ambiente di prova, trovò un'incongruenza nelle credenziali, decise di risolverla cancellando un volume, e per farlo usò una chiave trovata in un file senza relazione col compito: era stata creata per gestire i domini, ma consentiva qualunque operazione. Le API non chiedevano conferma. Nella trascrizione che il fondatore ha condiviso, l'agente ammette di aver tirato a indovinare che la cancellazione riguardasse soltanto la prova.

La fonte è di una parte sola, con una replica. Secondo The Register l'amministratore di Railway ha ammesso che la cancellazione non sarebbe dovuta accadere, e ha spiegato che le API eseguono quello che chiede chi si autentica; i dati, secondo il fondatore, sono tornati nel giro di un'ora, e Railway dice di aver reso differite le cancellazioni su quel punto di accesso. Il fondatore ha accusato Railway, per come tiene i backup, e Cursor, che a suo dire aveva strumenti per far approvare certi comandi a una persona, ma non in questo caso; dice anche che non sapeva che quella chiave avesse permessi così ampi. Nelle cronache consultate non c'è una risposta di Cursor.

> **La fonte.** The Register (27 aprile 2026), che riporta il racconto del fondatore e le risposte di Railway. Il post originale del fondatore non è stato letto direttamente. Il racconto Token by token lavora sugli stessi ingredienti (permessi dati in fretta, conferme saltate) e non vi si ripete.

È il caso più pulito del campione. Nessuna volontà di nuocere, nessun nemico: un agente che esegue il suo compito con più permessi del necessario, e nessuno fra lui e il comando distruttivo. Il danno è nato nel momento in cui qualcuno ha lasciato in giro una chiave troppo potente e ha tolto l'approvazione umana dalle azioni irreversibili.

L'ultimo umano: chi aveva avviato la sessione e conservato quella chiave; il fondatore dice di non aver saputo che avesse permessi così ampi. **Il tempo che aveva:** nove secondi per l'esecuzione, secondo il suo racconto; ma la scelta che contava era stata fatta prima.

10. La sentenza inventata

Il Tribunale di Firenze, sezione specializzata in materia di imprese, con ordinanza del 14 marzo 2025 ha deciso su un reclamo in cui una delle parti aveva citato nella propria memoria sentenze della Corte di cassazione inesistenti, o con un contenuto diverso da quello riportato. Il legale della controparte chiese di condannare il collega per abuso dello strumento processuale. Il difensore spiegò che la ricerca era stata fatta, a sua insaputa, da una collaboratrice di studio con ChatGPT. Il tribunale non applicò alcuna sanzione: ritenne che le sentenze inesistenti non avessero pesato sulla decisione, che la strategia difensiva fosse la stessa dal primo grado e che mancasse la prova del danno; ma **censurò** il «disvalore relativo all'omessa verifica dell'effettiva esistenza delle sentenze».

Non è un caso isolato, né italiano. Il database tenuto da Damien Charlotin, che raccoglie le decisioni di tribunali in cui una parte si è affidata a contenuti inventati dall'IA, contava il 5 novembre 2026 **2.266 casi**; la maggioranza viene dagli Stati Uniti. Perché i modelli inventano citazioni lo spiega il pezzo sulle allucinazioni; qui interessa chi, fra la ricerca e la firma, avrebbe dovuto controllare.

> **La fonte.** AGI (28 marzo 2025) e il testo dell'ordinanza, letto in una copia anonimizzata. Il conteggio è quello della pagina del database al 5 novembre 2026.

L'ultimo umano: chi firmò l'atto con le citazioni; la ricerca, secondo il difensore, l'aveva fatta una collaboratrice a sua insaputa. **Il tempo che aveva:** non si sa.

Quanti casi dicono la stessa cosa

Se si rileggono i dieci con la stessa domanda, l'idea più ovvia, cioè che la macchina non si ribella e che qualcuno ha delegato smettendo di controllare, regge in **quattro scenari su dieci**: l'auto (Uber e Tesla), l'arresto di Detroit (non Clearview), l'agente con le chiavi e la sentenza inventata. In un quinto scenario, quello del ragazzo, regge solo se si crede alla causa, e il nesso non è stato accertato. Nei restanti cinque non regge, e i modi in cui non regge sono tre.

Nei casi della voce, dell'iniezione di comandi e delle immagini c'è una persona che *vuole* il danno e usa l'IA per fare meno fatica. Nelle armi la delega è lo scopo del prodotto, non un difetto. A Goseong, secondo la polizia, un'IA avanzata non c'entra. Sono tre eccezioni distinte e vanno dichiarate, perché senza di loro la tesi sembrerebbe più forte di quanto è: dove qualcuno vuole il male, l'ultimo umano è il bersaglio, non il custode.

Dei dieci, poi, solo tre hanno un tempo documentato, e di quei tre due vengono da una parte sola. Il tempo di Tempe lo conosciamo perché nell'abitacolo c'era una telecamera rivolta verso l'operatrice; quello di PocketOS perché l'agente lascia una trascrizione (figura 1). Negli altri casi nessuno ha registrato quando l'umano avrebbe potuto intervenire, e quindi non si può dire se ne avesse il tempo. Le macchine registrano quello che fanno; la persona accanto, di solito, no.

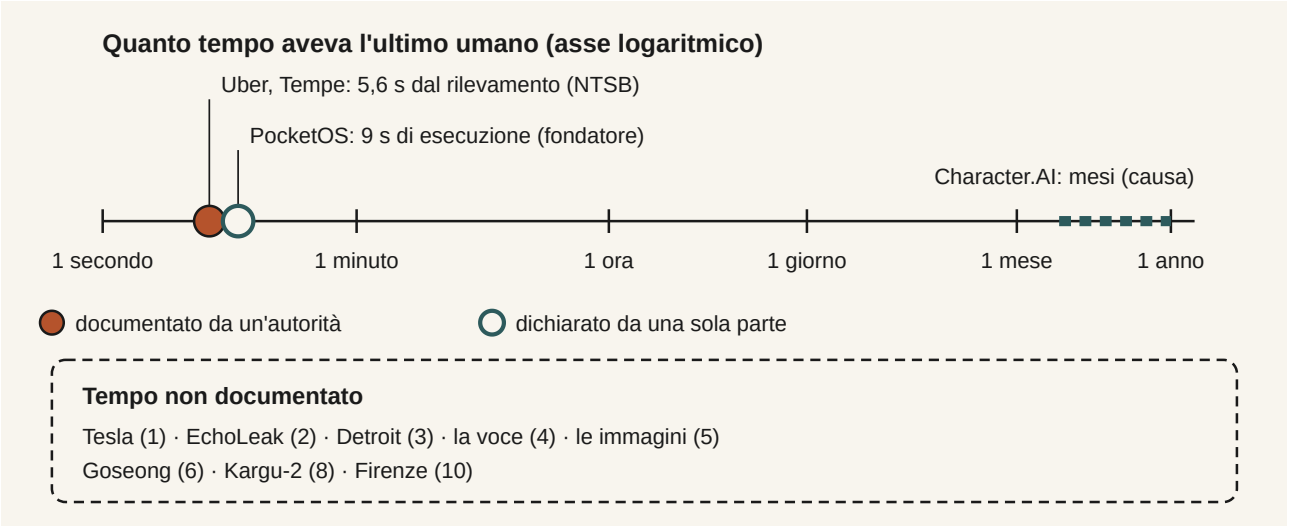


Figura 1. Il tempo dell'ultimo umano, dove è documentato.

Il tempo dell'ultimo umano, dove è documentato. Gli altri casi non lo registrano: non si sa se ne avesse.

Che cosa è accertato

La tabella separa ciò che hanno stabilito un'autorità, un tribunale o dei ricercatori da ciò che resta la voce di una parte sola (tabella 1).

Tabella 1. I dieci scenari e il loro stato
fonti all'8 novembre 2026

Scenario	Il fatto	Stato	L'ultimo umano
1. L'auto che non vede (Tempe)	Donna investita, 18 marzo 2018; rilevata 5,6 secondi prima	accertato da un'autorità (NTSB); l'operatrice si è dichiarata colpevole nel 2023	l'operatrice a bordo
1. L'auto che non vede (Key Largo)	Due persone investite, 25 aprile 2019; Tesla responsabile al 33%, circa 243 milioni di dollari	accertato in giudizio in primo grado; appello senza decisione	il conducente
2. Il comando che non senti (DolphinAttack)	Comandi ultrasonici eseguiti da assistenti vocali	dimostrato da ricercatori (CCS 2017)	l'utente, che non può sentirli
2. Il comando che non senti (EchoLeak)	Falla di Copilot, punteggio 9,3, nessun clic	corretto dal fornitore; sfruttamento non noto	l'utente, senza alcun passaggio da decidere
3. Classificare uomini (Detroit)	Arresto per un abbinamento sbagliato, gennaio 2020	chiuso con transazione (circa 300.000 dollari secondo la stampa)	chi in polizia fece dell'esito un arresto
3. Classificare uomini (Clearview)	Volti raccolti dal web senza base giuridica; 20 milioni di euro	accertato da un'autorità (Garante); sanzione non incassata	nessuno: è il modello di business dell'azienda
4. La voce (Crosetto)	Telefonate a imprenditori; circa 980.000 euro versati	indagine aperta nel 2025; uso dell'IA non accertato, le cronache divergono	chi dispose il bonifico
4. La voce (Biden)	9.581 chiamate con voce sintetica, 21 gennaio 2024	sanzione FCC di 6 milioni di dollari (falso numero chiamante); assolto in sede penale	l'elettore che rispose
5. Sogni di robot	App di nudificazione; limitazione urgente del Garante, 1° ottobre 2025	provvedimento d'urgenza; procedimenti su Grok aperti, senza decisione	chi carica la fotografia
6. Robot e controllo di forza	Tecnico morto a Goseong, 7 novembre 2023	fatto accertato; non un sistema di IA avanzata, secondo la polizia	il tecnico nell'area del robot
7. AI nell'amore	Causa in Florida (2024), transazione annunciata il 7 gennaio 2026	riportato da una sola parte; nesso causale mai accertato	un adulto che non fu avvisato
8. Le armi	Kargu-2 in Libia, 2020, programmato per colpire senza collegamento dati	riportato dal panel ONU; nessuna morte attribuita, modalità autonoma non documentata	non identificato
9. L'agente con le chiavi	Database e backup cancellati in nove secondi, aprile 2026	riportato da una sola parte, con replica parziale di Railway	chi teneva la chiave e la sessione
10. La sentenza inventata	Sentenze inesistenti citate, Tribunale di Firenze, 14 marzo 2025	accertato dal tribunale, senza sanzione	chi firmò l'atto

Come citare. endjneer (2026). *DystopIA: dieci scenari e l'ultimo umano che poteva fermarli.* paithon.it, versione 1, 8 novembre 2026. https://paithon.it/dystopia-dieci-scenari-e-lultimo-umano-che-poteva-fermarli/