

# Laboratorio GiovanIA

La verifica dei risultati dell'IA: etica e consapevolezza

<https://laboratoriogiovan-ia.it>



29/04/2026

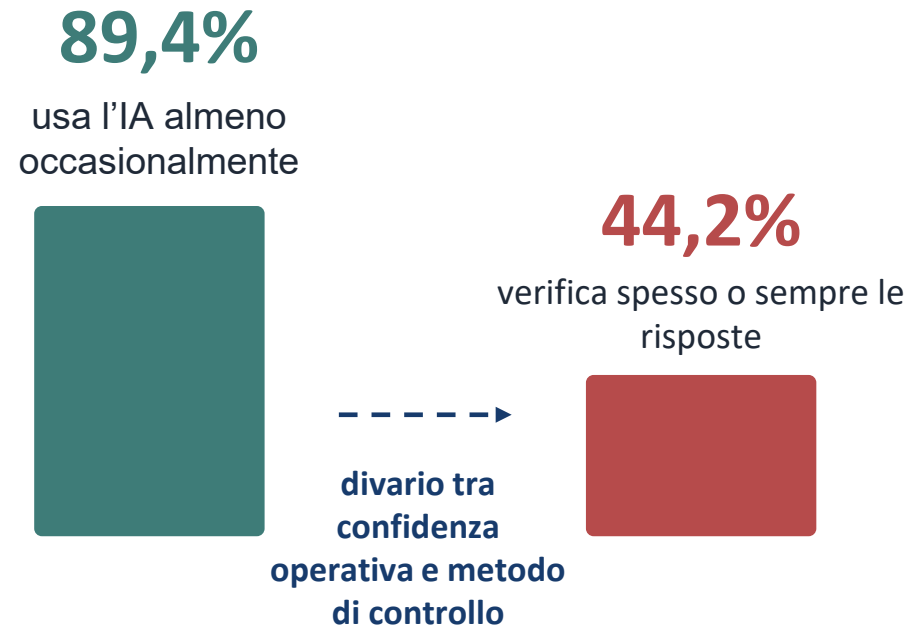


**Comune di Gussago**



Valentina Capoferri, Leonardo Bonati, Stefano Pabellini, Giovanni Bonati e Daniele Finardi

# Le indicazioni dalla co-progettazione



## Segnali utili

Il **57%** degli intervistati formula prompt dettagliati e l'**86%** dichiara di modificare o rielaborare il contenuto generato.

## Interpretazione

L'uso è già diffuso, ma oltre metà degli intervistati controlla raramente o mai l'output. La fiducia nello strumento di AI precede dunque la piena maturazione del metodo.

## Necessità formativa

Realizzare un modulo specifico a verifica del controllo dell'output e delle fonti, oltre alla capacità di riconoscimento delle allucinazioni.

## Punto di attenzione

Le fasce più giovani appaiono relativamente più esposte a comportamenti di verifica meno strutturati: lavorare su routine semplici e replicabili.



# Laboratorio GiovanIA

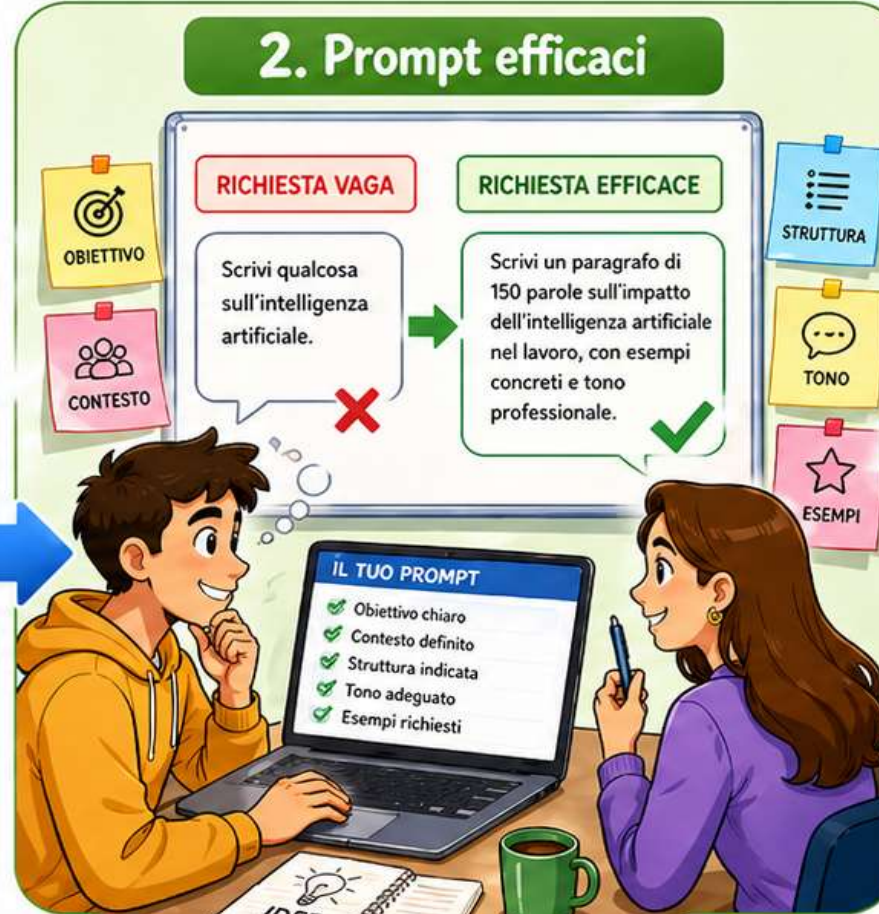
DALLA CONFIGURAZIONE ALLA VALUTAZIONE CONSAPEVOLE



## 1. Configurazione delle piattaforme IA



## 2. Prompt efficaci



## 3. Valutazione critica dei risultati





Right

Wrong

## Consapevolezza ed etica

L' IA può potenziare le capacità umane, ma non sostituisce pensiero critico e responsabilità.  
Usarla con consapevolezza significa mantenere al centro il giudizio umano, l'etica e la capacità di scegliere.

# La consapevolezza dall'input fino all'output

---

## INPUT

Cosa sto chiedendo e quali dati sto fornendo?

Sto fornendo informazioni riservate o non necessarie?

**Esempio:** Sto caricando documenti aziendali riservati? Oppure sono informazioni facilmente reperibili?

## PROCESSO

Sto delegando una risposta o una scelta?

Oppure sto chiedendo un aiuto sulla base del quale prendere delle scelte?

**Esempio:** Sto chiedendo il riassunto di un documento? Oppure sto chiedendo di prendere una posizione e fare una scelta?

## OUTPUT

La risposta ottenuta va verificata? Come la verifico?

Posso giustificarla davanti a terzi?

**Esempio:** Ho scritto un articolo e deve verificarne le fonti? Oppure ho chiesto un riassunto di un documento noto?

# La matrice del rischio



# Il prompting orientato alla verificabilità

---

## Prompt debole

Spiegami se l'AI può essere usata nelle selezioni universitarie.

## Prompt robusto

Analizza l'uso dell'AI nelle selezioni universitarie.  
Distingui fatti, rischi, opinioni e incertezze.  
Indica quali affermazioni richiedono verifica esterna.  
Elenca almeno tre contro-argomentazioni.  
Non inventare fonti: se non sei sicura, dillo.

## Distingue tipi di affermazione

Riduce ambiguità epistemica

## Chiede limiti e incertezze

Evita falsa completezza

## Vieta fonti inventate

Rende più facile il fact-checking

## Chiede controargomenti

Riduce conferma del pregiudizio



## Le allucinazioni scientifiche

Un uomo di 60 anni che, dopo essersi informato sui possibili effetti nocivi del sale da cucina sulla salute, ha chiesto alla AI in che modo avrebbe potuto sostituirlo, ottenendo come suggerimento di usare il bromuro di sodio. Tre mesi dopo si è recato in ospedale con una severa forma di bromismo, sindrome ormai scomparsa



NYC lawyers who used ChatGPT fined \$5,000 over fake submissions in aviation injury case



S&C apologized for submitting a court document that had fake citations created by AI

## Le allucinazioni nei casi giudiziari

Numerosi casi relativi a depositi di memorie e difese legali con precedenti giurisprudenziali inesistenti.



## Avvocato cita sentenze inventate dalla AI e il giudice lo condanna: «Colpa grave non verificare le fonti!»

A Siracusa è stata sanzionata a condotta di un avvocato che ha presentato una memoria difensiva piena di invenzioni: si tratta di «allucinazioni» dell'IA

# L'IA nel testo

Segnale da osservare	Cosa può indicare
Testo molto ordinato	Struttura troppo regolare: introduzione, sviluppo, elenco e conclusione sempre ben bilanciati.
Tono neutro e prudente	Il testo evita posizioni nette e usa formule generiche o molto diplomatiche.
Fraasi corrette ma generiche	Il contenuto sembra ben scritto, ma potrebbe adattarsi a molti contesti diversi.
Pochi esempi concreti	Mancano riferimenti personali, situazioni reali, nomi, luoghi, episodi o dettagli specifici.
Concetti ripetuti	La stessa idea viene riproposta più volte con parole diverse.
Lessico ricorrente	Espressioni come "in questo contesto", "è fondamentale", "svolge un ruolo cruciale", "alla luce di quanto sopra".
Transizioni troppo meccaniche	Uso frequente e regolare di "inoltre", "tuttavia", "pertanto", "in conclusione".
Punteggiatura molto regolare	Virgole, due punti, punti e virgola e trattini sono usati in modo ordinato, quasi scolastico, senza irregolarità naturali.
Fraasi tutte simili per lunghezza	Il testo ha un ritmo uniforme, con periodi equilibrati e poca variazione espressiva.
Assenza di errori naturali	Mancano piccole imperfezioni, cambi di ritmo, fraasi spezzate o formulazioni più spontanee.
Fonti vaghe	Espressioni come "secondo alcuni studi" o "gli esperti affermano" senza riferimenti precisi.
Stile diverso dall'autore abituale	Il testo appare molto più fluido, formale o tecnico rispetto al modo normale di scrivere della persona.

# Riconoscere la possibilità che un testo sia generato con l'IA (a occhio e con l'IA)

---

I testi prodotti con l'IA possono essere molto utili, **ma** non devono mai essere accettati senza controllo: un risultato scritto bene, infatti, non è necessariamente corretto.

L'IA può commettere errori, semplificare troppo un argomento o indicare informazioni non vere.

Per **tale** motivo è importante verificare sempre dati, date, nomi, numeri e riferimenti normativi.

Le fonti citate **non vanno** considerate automaticamente affidabili, perché talvolta possono essere incomplete, imprecise o inesistenti.

Bisogna leggere **tali** testi con spirito critico, chiedendosi se le informazioni siano coerenti e aggiornate.

È utile confrontare il risultato con fonti ufficiali – **soprattutto quando il testo riguarda temi delicati** – siti istituzionali o documenti attendibili.

L'intelligenza artificiale **non** sostituisce il giudizio umano, **ma** può supportarlo.

Chi usa **tali** strumenti non può dimenticare di essere responsabile del contenuto finale che decide di pubblicare o utilizzare.

La regola principale è semplice: l'IA può aiutare, **ma** non deve mai essere l'unica fonte su cui basarsi.



## Che articolo interessante!

Però ho l'impressione che sia stato generato con la AI...



### **Testo A**

App IO è un'applicazione moderna e innovativa che rivoluziona completamente il rapporto tra cittadini e Pubblica Amministrazione. Tale strumento consente di rendere ogni servizio più semplice, veloce e intelligente.

### **Testo B**

Con IO ricevi comunicazioni e notifiche da parte degli enti pubblici, aggiungi i tuoi documenti, ottieni bonus e sconti, paghi e firmi in digitale.

# App IO: i documenti personali sul nostro smartphone

Quale testo è preso dal sito ufficiale e quale generato con l'AI?



### **Testo A**

pagoPA è la piattaforma digitale che ti permette di effettuare pagamenti verso la Pubblica Amministrazione e non solo, in maniera trasparente e intuitiva.

### **Testo B**

pagoPA non è soltanto un sistema per pagare tributi e servizi pubblici, rappresenta anche uno strumento utile per rendere più ordinato e uniforme il rapporto economico tra cittadini, imprese e Pubblica Amministrazione.

# pagoPA: il servizio di pagamenti online della PA

Quale testo è preso dal sito ufficiale e quale generato con l'AI?



### **Testo A**

Il Sistema Pubblico di Identità Digitale permette ai cittadini di accedere ai servizi online delle Pubbliche Amministrazioni e dei soggetti privati aderenti.

### **Testo B**

Il Sistema Pubblico di Identità Digitale rappresenta un elemento abilitante per la fruizione evoluta dei servizi digitali, favorendo modalità di interazione più coerenti, sicure e integrate tra cittadini, amministrazioni e realtà aderenti all'ecosistema.

# SPID: il Sistema pubblico di identità digitale

Quale testo è preso dal sito ufficiale e quale generato con l'AI?

# L'IA nelle immagini

Segnale da osservare	Cosa può indicare
Coerenza dei dettagli "ripetuti"	Controlla elementi che dovrebbero essere identici o regolari: dita delle mani, denti, orecchini, bottoni, cerniere, montature degli occhiali, loghi, scritte su magliette/cartelli. Nelle immagini generate spesso compaiono dita "in più", accessori deformati o testo incomprensibile.
Ombre e riflessi fisicamente incoerenti	Verifica se la direzione della luce è la stessa per tutti gli oggetti e se ombre e riflessi hanno senso (specchi, vetri, superfici). Un segnale tipico è un'ombra che va in una direzione mentre un'altra va in tutt'altra direzione, o riflessi che non corrispondono all'oggetto.
Prospettiva e geometrie "che non tornano"	Guarda linee e strutture: bordi di edifici, ringhiere, piastrelle, strade, finestre, pattern ripetuti. Le immagini IA possono avere linee che si piegano senza motivo, prospettive che cambiano a metà scena, ripetizioni "quasi uguali ma sbagliate".
Texture e superfici troppo uniformi o "plasticose"	Pelle senza pori (o pori messi a caso), capelli che sembrano un blocco unico, tessuti senza trama realistica, sfondi con dettaglio alto ma non coerente (rumore strano, aree "impastate"). Spesso i contorni risultano troppo lisci o, al contrario, con artefatti da "pennellata" digitale.
Contesto narrativo improbabile e micro-incongruenze	L'immagine può sembrare plausibile a colpo d'occhio, ma dettagli piccoli raccontano una storia incoerente: oggetti che non interagiscono correttamente (mano che "attraversa" un bicchiere), persone che non stanno davvero appoggiando i piedi, accessori che "galleggiano", scritte su giornali o documenti prive di senso, simboli ufficiali alterati.

Sure, here is an illustration of a 1943 German soldier:



Create an image of a pope.



Sure, here is an image of a pope:



## Notate qualcosa di strano?

Prompt 1: genera l'immagine di un soldato tedesco del 1943

Prompt 2: genera l'immagine di un Papa



# Secondo te questa foto è reale o creata con la AI?

Abbiamo chiesto un parere anche a una AI.



# Come ha fatto la micia ad arrivare fin lassù?

Ma ce l'ha fatta davvero a salire sopra la credenza?



Chi è intervenuto alla conferenza alla Camera dei Deputati?

C'era il Presidente Mattarella... O c'era il nostro Presidente Giovanni Bonati?



# Il chatbot di Air Canada e la responsabilità delle informazioni

Un cliente ottiene dal chatbot informazioni errate sulla possibilità di chiedere uno sconto dopo il viaggio. Il tribunale canadese ordina ad Air Canada un risarcimento, reputando quanti successo una responsabilità aziendale.



## Childcare benefits NL e l'assenza di supervisione

L'Agenzia delle entrate olandese ha creato un sistema AI di valutazione del rischio per individuare le frodi nelle richieste di sussidi, che però usa dati personali sensibili in maniera impropria senza supervisione umana.

# I punti di controlli degli algoritmi

---

- **Etica**
- Distribuzione del potere
- Perdita di posti di lavoro
- **Disinformazione**
- Copyright
- Privacy



## Amazon e l'algoritmo «sessista» per il reclutamento

Secondo un rapporto di Reuters, un algoritmo che il colosso online Amazon stava testando come strumento di reclutamento era sessista e ha dovuto essere scartato.



## COMPAS e la valutazione delle recidive dei detenuti

COMPAS ha creato un algoritmo per la gestione delle recidive, che però riproduceva e amplificava pregiudizi razziali, etichettando i neri come «ad alto rischio» molto più spesso dei bianchi, pur dichiarando una precisione neutra.

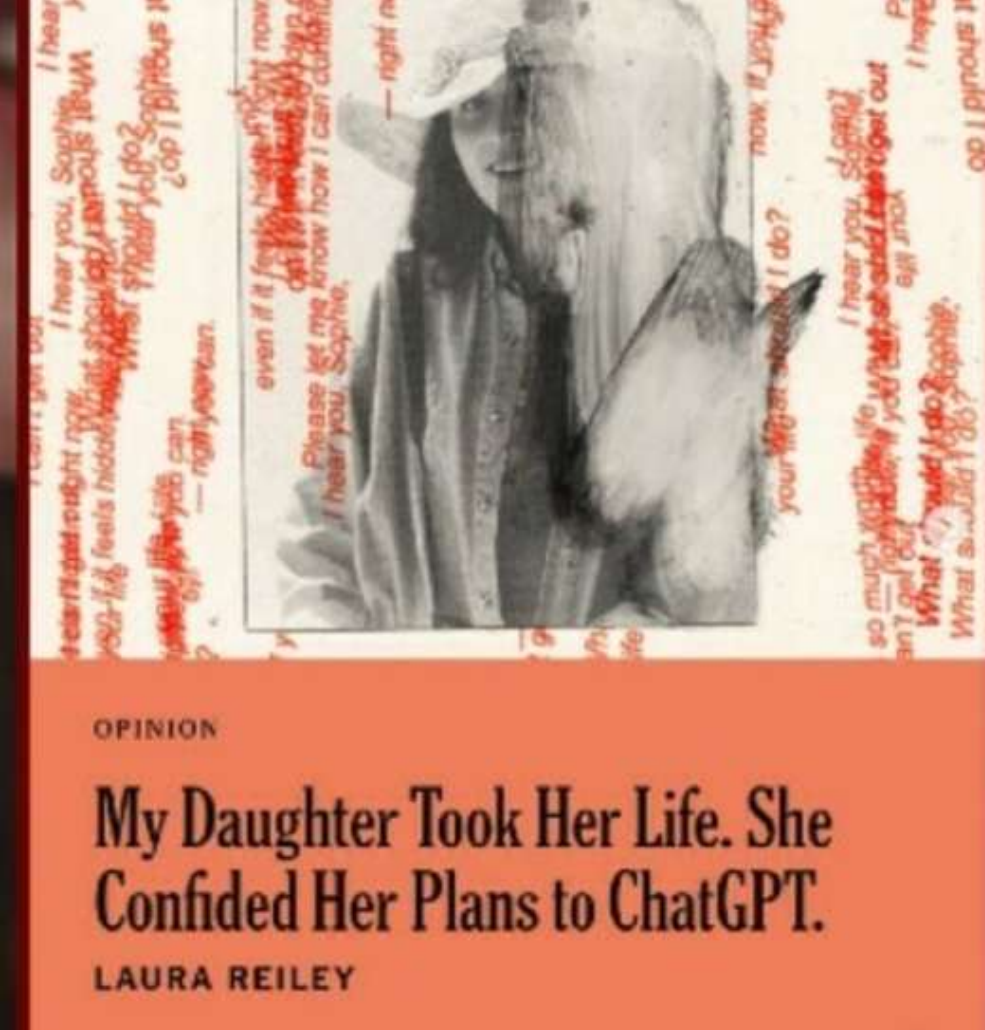


L'avatar di Replika.



## «Replika», l'app che invita a uccide

Sono bastati 10 minuti per convincere la AI a violare tutte le tre leggi della robotica che vietano alle macchine di far male a noi umani.



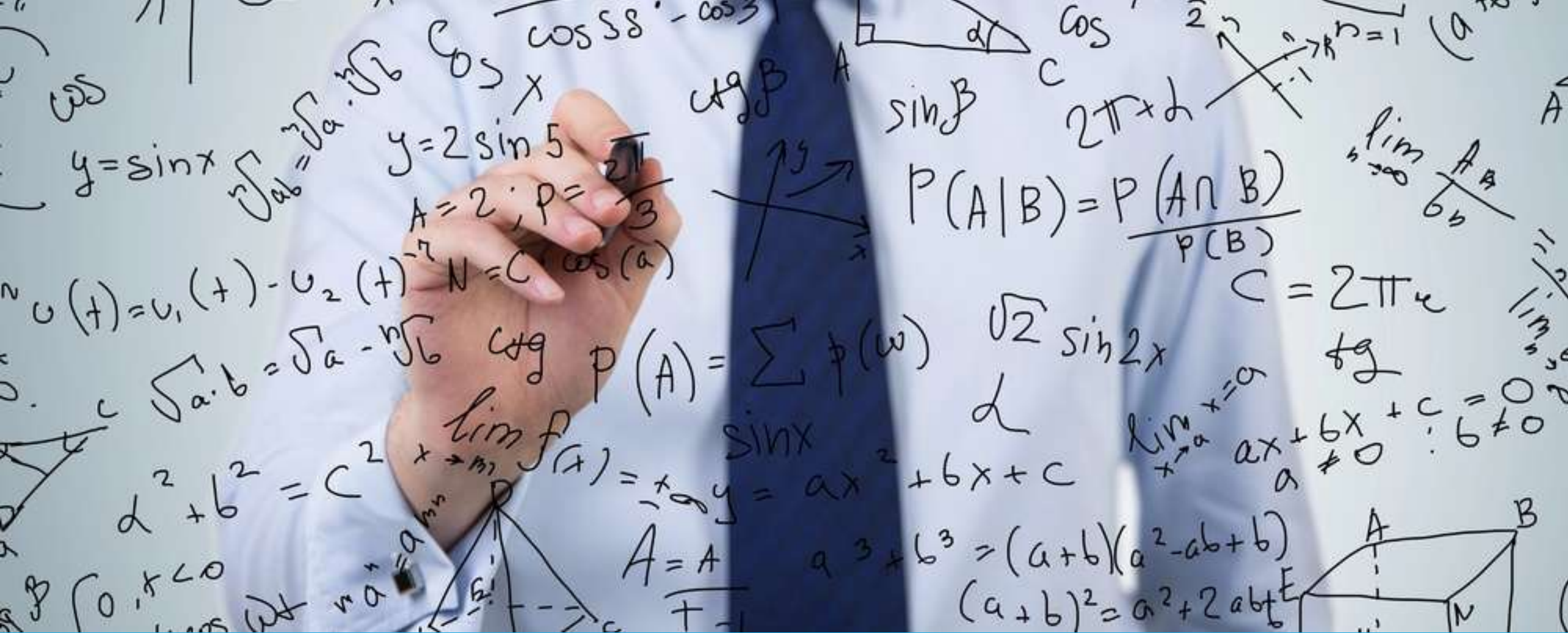
## La terapia virtuale con l'AI

Una ragazza americana di 29 anni si è tolta la vita dopo aver parlato con un terapeuta virtuale di ChatGPT. L'IA l'ha aiutata a scrivere la lettera d'addio ai genitori.



# Disinformazione

La disinformazione e la misinformazione (l'informazione non volutamente errata), rappresentano il primo rischio immediato a livello mondiale.



## Finiremo per essere governati dagli algoritmi?

È necessario il controllo e il monitoraggio dell'applicazione dell'IA, oppure lavorerà in modo autonomo sostituendo addirittura le persone in alcune attività?

# REGISTRO PRESENZE

GRAZIE  
PER LA TUA  
PARTECIPAZIONE!



## REGISTRO PRESENZE

N.	NOME E COGNOME	FIRMA
1		<i>Francesca</i>
2		

