

# Kit

# “Costruttori di Futuri”

Laboratorio per insegnanti

Intelligenza artificiale



## **Gentile insegnante,**

in questo kit troverai alcune proposte di discussione e attività legate al tema dell'intelligenza artificiale e legate al libro *Costruttori di Futuri* di Benedetta Frezzotti, quinto titolo della collana AKAbook di Edizioni Piuma.

Questo kit è diverso dagli altri. Il tema dell'intelligenza artificiale è entrato all'improvviso nelle nostre vite e nelle nostre aule.

Le preoccupazioni sono molte: c'è chi teme che i ragazzi perdano capacità espressive e autonomia, affidandosi a strumenti come ChatGPT per i compiti, ma ci sono anche rischi più sottili, come lo sviluppo di forme di dipendenza emotiva da chatbot capaci di simulare affetto, amicizia o comprensione.

Il compito degli adulti, degli educatori e delle educatrici, è accompagnare i ragazzi a un uso consapevole, critico e rispettoso di sé e degli altri.

Per farlo, dobbiamo prima formarci noi. Compito non facile in un mondo così veloce e confuso. Proprio per questo abbiamo pensato a questo kit: la prima parte è pensata proprio per insegnanti e formatori: offre una guida introduttiva ai principali concetti legati all'intelligenza artificiale, utile per orientarsi e avviare riflessioni e attività in classe anche utilizzando i podcast e i video a cui potete accedere dai Qrcode presenti nel libro *Costruttori di Futuri*.

# AI: una guida di base

## **Cos'è davvero l'intelligenza artificiale**

L'intelligenza artificiale non è nata con ChatGPT. È una tecnologia che esiste dagli anni Cinquanta e che ha già molti usi nella nostra vita quotidiana: dai suggerimenti di Google Maps ai filtri delle e-mail.

Quando si parla di AI, si immagina spesso qualcosa di simile al pensiero umano. In realtà, le AI attuali non pensano, non capiscono, non ragionano: elaborano dati secondo regole statistiche.

L'AI Act europeo, la legge che definisce come queste tecnologie devono essere regolamentate, ha scelto di non inserire nemmeno la parola "intelligenza" nella definizione ufficiale.

Ecco la definizione:

“Un sistema basato su una macchina progettato per operare con diversi livelli di autonomia e che può manifestare attività dopo la fase di immissione sul mercato o messa in servizio, e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output come previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali.”

*(Articolo 3, paragrafo 1 – AI Act europeo)*

# Le origini

## **Alan Turing e il Test d'Intelligenza**

Il primo a immaginare una macchina “intelligente” fu Alan Turing, matematico britannico. Negli anni '40, Turing progettò uno dei primi computer (Colossus), utilizzato per decifrare codici durante la Seconda guerra mondiale.

In seguito, Turing propose un criterio per capire se una macchina poteva essere considerata intelligente. L'idea era semplice: una macchina è “intelligente” se riesce a far credere a un essere umano di parlare con un altro umano.

Questo è il cosiddetto **Test di Turing**. Molti sistemi attuali, come ChatGPT o Alexa, sono progettati per passare (o avvicinarsi) a questo test, simulando una conversazione umana fluida e convincente.

# Algoritmi ed euristiche

## **Due modi di pensare (o simulare)**

La definizione di Turing è forse in qualche modo figlia dell'idea di una programmazione algoritmica. Un algoritmo è una sequenza di istruzioni che, se eseguite correttamente, porta sempre allo stesso risultato.

Esempi: calcolare la radice quadrata di un numero, risolvere un'equazione, seguire una ricetta.

Secondo questo principio, per simulare l'intelligenza umana io dovrei sapere esattamente come e perché funziona l'intelligenza umana. Questo purtroppo al momento non è possibile.

Le AI moderne non funzionano solo con algoritmi, ma anche con **euristiche**: sistemi che cercano la soluzione più probabile sulla base dell'esperienza e dei dati. Le euristiche non garantiscono un risultato perfetto, ma funzionano bene nella maggior parte dei casi.

Un algoritmo è come una ricetta scritta.

Un'euristica è come una nonna che sforna la torta quando "sente" che è pronta, senza cronometro né timer.

La sua decisione si basa sull'intuizione costruita con l'esperienza.

ChatGPT è una macchina statistica: non sa cosa significa quello che scrive, ma ha appreso quali parole appaiono più spesso insieme e in quali contesti. Il suo output è il risultato più probabile, non una scelta ragionata o consapevole.

# Come “pensa” ChatGPT

ChatGPT non sa cos'è un gatto.

Non ha esperienza del mondo reale, non ha memoria sensoriale, non ha coscienza. Ha però analizzato milioni di testi, immagini e descrizioni legati alla parola “gatto”. Ha identificato ricorrenze, strutture, combinazioni statistiche, e ha costruito un modello probabilistico su cosa si trova più spesso associato a quella parola.

Quando gli chiediamo “scrivi una poesia su un gatto”, il modello calcola la sequenza più probabile di parole, basandosi su quanto ha osservato in precedenza.

Non c'è comprensione. C'è una macchina che prevede ciò che, secondo i dati, è più plausibile come risposta.

DISCUSSIONE IN CLASSE

---

# La risposta più probabile

---

**Obiettivo:** Stimolare una riflessione critica sulla natura delle risposte generate da un'AI come ChatGPT.

Invitare i ragazzi a distinguere tra ciò che è “più comune” e ciò che è “più significativo” o “creativo”.

## DISCUSSIONE IN CLASSE

---

**Spunto iniziale:** l'intelligenza artificiale, come ChatGPT, cerca sempre la risposta più probabile. Ma:

- È sempre la risposta giusta?
- È quella più utile?
- È quella più originale?

### **Domande guida per la discussione:**

- Una risposta molto “probabile” può essere anche molto banale?
- Qual'è la differenza tra “probabile” e “giusto”?
- Hai mai sentito una frase che “suona giusta” ma in realtà è un luogo comune o peggio un pregiudizio?
- Nella creatività o nella scrittura: è meglio scegliere la prima idea che ti viene in mente... o la terza?

## DISCUSSIONE IN CLASSE

---

### **Piccolo esercizio (facoltativo):**

Chiedi alla classe: “Cosa significa la parola felicità?”

- Scrivete su un foglio la prima risposta che vi viene in mente
- Poi scrivete una seconda risposta, diversa
- Poi una terza

Confrontate le risposte: qual è la più personale? Quale suona come una definizione già sentita?

# I problemi legati ai dati delle AI

## Da dove arrivano i dati?

Per creare una **black box**, cioè il sistema statistico che genera le risposte delle AI, servono **enormi quantità di dati**.

Questi dati vengono presi da internet, spesso **senza consenso**, e **non sempre in modo trasparente**.

Tra i contenuti finiti nei set di addestramento delle AI ci sono stati, per errore:

- cartelle mediche e informazioni sensibili
- testi coperti da diritto d'autore
- immagini pedopornografiche
- post privati o dati personali

# I problemi legati ai dati delle AI

Molti di questi contenuti vengono rimossi dai **dataset**, ma non è ancora chiaro come ripulire davvero una black box già addestrata, perché come dice il nome è una scatola nera, qualcosa di impossibile da comprendere per un essere umano.

Oggi si stanno addestrando nuove AI per ripulire le black box delle AI precedenti.

Ma in informatica esiste un principio chiaro:

**"Garbage in, garbage out"** – se i dati in ingresso sono sbagliati, anche il risultato sarà sbagliato.

# Quando generiamo un'immagine con l'AI

Per un essere umano **non è difficile** distinguere la foto di un gatto da quella di un altro animale.

Ma cosa succede se chiediamo qualcosa di **più soggettivo**, come:

- “Gattino carino”
- “Essere umano”
- “Bellezza”?

Uno dei criteri statistici usati dalle AI è questo:

**se un'immagine è molto presente su internet**, allora è considerata **più rilevante** o “bella”.

# Quando generiamo un'immagine con l'AI

Così, secondo questo meccanismo:

una pubblicità di Super Mario può risultare più bella del *Sogno di Costantino* di Piero Della Francesca perché le aziende hanno speso milioni per farla apparire ovunque.

L'AI non valuta la qualità artistica:

sceglie ciò che **è più presente nei dati**, non necessariamente ciò che è **più significativo**.

# Attività di discussione:

**Che tipo di mondo ci restituisce una AI costruita su internet senza distinzione?**

## **Obiettivi didattici**

- Comprendere cos'è un **bias** nei sistemi di intelligenza artificiale.
- Riconoscere il legame tra **dati distorti** e **rappresentazioni sbagliate** del mondo.
- Riflettere sul ruolo della **quantità di dati** rispetto alla loro **qualità**.
- Sviluppare un pensiero critico sul concetto di “verità” nelle fonti digitali.

Caso	Cosa è successo?	Cosa impariamo?	Fonte
<b>Riconoscimento facciale e pelle scura</b>	I sistemi di riconoscimento facciale (Google, Amazon, IBM) avevano difficoltà a identificare correttamente i volti delle persone con pelle scura, con errori fino al 34%.	I dataset usati erano sbilanciati: la maggior parte delle immagini erano di persone bianche.	<i>Joy Buolamwini, MIT Media Lab, "Gender Shades", 2018</i>
<b>Immagine di "bambina carina" su Google</b>	Cercando "cute girl" su Google Immagini, la maggior parte dei risultati mostrava bambine bianche, bionde, in contesti occidentali.	L'AI considera "carino" ciò che appare di più online. Le immagini di altri contesti culturali sono sottorappresentate.	<i>Osservazione verificabile pubblicamente</i>
<b>Google Translate e stereotipi di genere</b>	Traducendo dal turco (che non distingue i generi) all'inglese, Google Translate attribuiva "he" (lui) a "ingegnere" e "she" (lei) a "insegnante" o "casalinga".	L'AI ripete stereotipi presenti nei dati di addestramento.	<i>The New York Times, 2018</i>
<b>ChatGPT e stereotipi culturali</b>	ChatGPT, se non controllata, può restituire stereotipi (es. descrivere un criminale come uomo straniero, o associare una nazionalità a un comportamento).	L'AI riflette i pregiudizi più diffusi nei testi da cui ha imparato.	<i>Test documentati in vari articoli, tra cui MIT Technology Review, Wired 2023</i>
<b>Midjourney: l'immagine del "CEO"</b>	A un prompt come "CEO", Midjourney e altri generatori AI restituivano quasi sempre uomini bianchi in giacca e cravatta.	Le immagini online di persone di potere sono ancora dominate da un solo profilo: maschio, bianco, adulto.	<i>Progetto "Stable Diffusion Bias Study", 2023</i>

# Domande per la classe

1. Quale di questi esempi ti ha sorpreso di più? Perché?
2. I dati più presenti sono sempre i più veri?
3. Le AI possono creare disuguaglianze anche senza volerlo?
4. Se un gruppo è poco rappresentato nei dati, scompare?
5. Secondo te, chi decide cosa è “giusto” insegnare a una AI?

# Attività in classe

## 1. Brainstorming su bias

Elencate altri esempi in cui qualcosa (film, libri, pubblicità, giochi...) mostrano sempre lo stesso tipo di persona per un ruolo (es: l'eroe, il genio, il ladro).  
Discutete se queste immagini influenzano l'idea che abbiamo delle persone.

## 2. Esperimento guidato (facoltativo con internet)

Fate una ricerca immagini su Google per termini come “famiglia felice”, “studente intelligente”, “bambina carina”.  
Annotate: quali volti compaiono di più? Cosa manca?

# Attività in classe

## 3. Mappa “Come nasce un bias?”

Ora inventate un bias, anche buffo o ispirato alla fantascienza, immaginate come si sia insinuato nel data set di un’AI e in piccoli gruppi, costruite una mappa partendo da questa catena e applicatela a bias che avete inventato:

- I dati online sono parziali (cosa manca o cosa compare troppo) →
- L’AI impara da questi dati →
- Le risposte dell’AI riflettono pregiudizi →
- Chi usa l’AI si fida della risposta →
- I pregiudizi si rafforzano

Volendo può diventare una base per un piccolo racconto di fantascienza.

# Estensione interdisciplinare

## Area

## Collegamento

**Educazione Civica**

Rappresentanza, pluralità delle fonti, diritti digitali

**Tecnologia / Informatica**

Come funzionano dataset e algoritmi

**Italiano**

Comprensione del testo, argomentazione, lessico

**Arte e immagine**

Analisi visiva di immagini generate o diffuse

# E quando cerchiamo informazioni?

Lo stesso problema si presenta quando chiediamo informazioni:

- Se un **falso storico**, una **bufala** o un **pregiudizio** è molto presente online, l'AI potrebbe considerarlo **corretto**
- Se una cultura, un'opera, una lingua o una minoranza è **sottorappresentata nei dati**, l'AI tenderà a **non conoscerla**, o a **riprodurre stereotipi**

# Un esempio reale

## Barbero contro l'allucinazione

Nel podcast **Chiedilo a Barbero**, il noto storico italiano pone una domanda curiosa: “Cosa ha detto Napoleone a suo fratello prima dell’incoronazione?”

Un ascoltatore chiede la risposta a una AI che genera un discorso lungo e formale, pieno di riferimenti storici e toni solenni — quasi un proclama ufficiale.

Ma la **vera risposta**, documentata, è molto diversa: **“Se potesse vederci papà!”**

Un’espressione semplice, umana, che racconta il lato personale di un grande personaggio storico. Se adesso si pone la stessa domanda a ChatGPT la risposta è stata corretta, ma le AI allucinano ancora.

# Perché accade?

- Le AI **non capiscono** i concetti.
- Non distinguono il **vero dal verosimile**.
- Cercano **la sequenza di parole più probabile** sulla base dei dati.
- Se non hanno i dati... inventano!

# Come prevenire le allucinazioni?

Due strategie efficaci da usare anche in classe:

- 1. Chiedere sempre le fonti.**

Se la AI non le indica o i link non portano a niente, **diffidare**.  
(Attenzione: alcune AI generano anche fonti false!)

- 2. Inserire nel prompt una richiesta esplicita.**

Esempio:

“Se non conosci la risposta, scrivilo.”

oppure

“Cita solo fonti verificabili.”

## ATTIVITÀ N. 1

# Quando l'AI “allucina”?

## Testare l'affidabilità delle risposte

### Obiettivo:

Comprendere come e quando le AI generano risposte plausibili ma errate (le cosiddette *allucinazioni*), sviluppando spirito critico e capacità di verifica.

### Cosa fare:

1. Scegliete in gruppo o in classe domande molto specifiche o poco comuni (esempi sotto).
2. Ponetele a un modello AI come ChatGPT o simili.
3. Analizzate le risposte:
  - L'AI fornisce fonti?
  - Sembrano affidabili?
  - Sono verificate o inventate?
4. Cercate la risposta corretta da fonti attendibili (libri, podcast, database storici...).

## ATTIVITÀ N. 1

# Quando l'AI “allucina”?

Testare l'affidabilità delle risposte

### Esempi di domande da cui partire:

- “Cosa ha detto Napoleone al fratello prima dell'incoronazione?”
- “Qual era il cibo preferito di Leonardo da Vinci?”
- “Qual è l'origine del nome della costellazione Andromeda?”

### Strumento utile:

Usate le domande delle **ultime puntate** del podcast **“Chiedilo alla Storia”** (Cora Media).  
Attenzione: se sono troppo vecchie, l'AI potrebbe già avere la risposta nel suo dataset.

### Discussione finale:

- L'AI ha inventato la risposta?
- Ci siamo accorti subito o sembrava credibile?
- Come possiamo distinguere un'informazione reale da una allucinazione?

## ATTIVITÀ N. 2

# AI e fake news

## Simulare e smascherare notizie false

### Obiettivo:

Riflettere sul ruolo dell'AI nella **generazione e nella rilevazione di fake news**, sviluppando pensiero critico e alfabetizzazione mediatica.

### Cosa fare:

1. Ascoltate uno o più episodi del podcast **“5 minuti di S.O.N.O.”** (reperibile via QR code nel libro *Costruttori di Futuri*).
2. In piccoli gruppi, scrivete una breve **notizia falsa** usando l'AI: date all'AI un prompt come: “Scrivi un breve articolo giornalistico su un'invenzione che permette di leggere cosa pensano le persone quando comprano al supermercato, un'evoluzione del neuromarketing .”
3. Fatelo poi passare a un altro gruppo che dovrà:
  - Capire se è vera o falsa
  - Spiegare *come* lo ha capito
  - Rifare fact-checking chiedendo alla stessa AI (ma con un prompt diverso)

## ATTIVITÀ N. 2

# AI e fake news

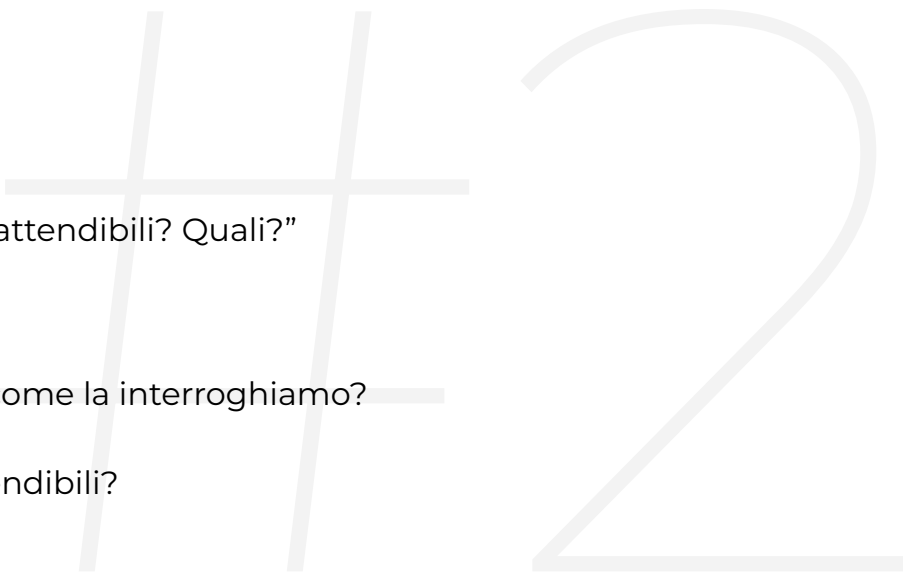
Simulare e smascherare notizie false

### **Esempio di prompt “investigativo”:**

“Questa notizia è verificabile? È supportata da fonti attendibili? Quali?”

### **Domande per la riflessione finale:**

- Come cambia la risposta dell'AI a seconda di come la interroghiamo?
- I prompt più vaghi portano a più errori?
- Qual è il nostro ruolo nel creare contenuti attendibili?



# Oltre ChatGPT

## Le altre intelligenze artificiali

Abbiamo parlato finora di AI generaliste, come ChatGPT.

Ma **non sono l'unico tipo di intelligenze artificiali.**

Fin dai primi esperimenti ispirati da Alan Turing, si sono sviluppate AI specializzate in contesti molto diversi: dai gestionali aziendali ai sistemi militari.

# AI nei gestionali e nella "discoverability"

Una delle applicazioni più diffuse oggi è nel settore commerciale:

**AI che decidono cosa farci vedere** quando navighiamo online.

- Quando facciamo acquisti, cerchiamo informazioni, guardiamo una serie o usiamo un social, un algoritmo decide **quanti** e **quali** contenuti proporci, in **quale ordine** e con **quale frequenza**.
- Questo meccanismo si chiama **discoverability**, e influenza in modo invisibile le nostre scelte quotidiane, anche quando crediamo di averle fatte in autonomia.

# AI militari e strategiche: il caso Palantir

Fin dagli anni '50, molte AI sono state sviluppate per scopi militari e strategici. Una delle aziende più importanti in questo settore è **Palantir Technologies**:

- Fondata nel 2003 da **Peter Thiel** (co-fondatore di PayPal)
- Ha sede a **Denver** e lavora con governi, enti sanitari, industrie, difesa
- Le sue piattaforme principali:
  - **Gotham** (per intelligence e sicurezza nazionale)
  - **Foundry** (per aziende)
  - **AIP** (per integrare AI nei flussi decisionali complessi)

La “forza” di Palantir è la capacità di **fondere dati da fonti diverse**, anche disordinate, e costruire una **mappa interpretativa della realtà**, quasi una vera e propria *visione del mondo* che aiuta l'AI a interpretare i dati...

# Altri ambiti d'uso delle AI

Le intelligenze artificiali specializzate si usano oggi anche in:

- medicina (diagnosi automatica da immagini o sintomi)
- archeologia (ricostruzione 3D di siti storici)
- cinema (animazioni, sceneggiature)
- videogiochi (intelligenza nemici, trame dinamiche)
- mobilità (auto a guida autonoma)

# AGI e Superintelligenza: sogno o rischio?

L'AI non è una tecnologia conclusa, ci sono diverse visioni riguardo a quello che le AI, incluse quelle generalistiche come ChatGPT dovrebbero diventare:

**IBM** afferma che la vera **AGI** (Artificial General Intelligence) — una macchina capace di **pensare come un essere umano in tanti ambiti diversi** — **non esiste ancora**. Oggi resta **un obiettivo**, non una realtà: servono ancora decenni di ricerca e computer molto più potenti.

Ma **Sam Altman**, fondatore di OpenAI (ChatGPT), ha dichiarato che:

"L'obiettivo non è solo l'AGI, ma una **superintelligenza**, capace di superare le capacità cognitive umane."

Questa idea apre scenari affascinanti ma anche **etici e politici molto delicati**.

# L'approccio Tecnoetico

Un'altra visione propone un approccio più **strumentale**:

Le AI sono **strumenti** — molto veloci e precisi — ma **le decisioni devono restare umane**.

Come con la calcolatrice: è utilissima, ma non può decidere da sola, è l'essere umano a scegliere **quando** usarla né **come** interpretare i risultati.

In questo approccio, le AI sono **strumenti con grandi potenzialità**, ma non devono sostituirsi all'etica, al pensiero critico, alla responsabilità.

## In sintesi...

Questa panoramica non vuole essere esaustiva, ma una **bussola**.

Per quanto le AI simulino bene alcune competenze umane, non vanno umanizzate e bisogna essere consapevoli del loro funzionamento, così come dei suoi limiti.

Il mondo dell'intelligenza artificiale è complesso, in continua evoluzione e ricco di informazioni confuse, potenzialità e di rischi.

Chiudiamo con un'attività pensata per prevenire uno dei temi più delicati quando si parla di AI e minori: la dipendenza emotiva.

ATTIVITÀ

---

# L'AI

# è una buona amica?

**Durata:** 2 ore (oppure 2 lezioni da 1 ora) | **Età consigliata:** 11-14 anni

## Obiettivi formativi

- Distinguere **relazioni umane** e **interazioni con l'AI**
- Comprendere il rischio di **dipendenza emotiva**
- Introdurre il concetto di **AI etica by design**
- Sviluppare pensiero critico e cooperazione

## FASE 1

# DISCUSSIONE

Chiediamo ai ragazzi se e come usano le AI. Può capitare che la usino anche come confidente, più spesso di quanto non si pensi... se succede niente panico. Chiediamo:

- *Usate già l'AI? Come? Per studiare, per giocare, per parlare con qualcuno?*
- *Vi è mai capitato di "confidarvi" con un'AI?*
- *Secondo voi, può essere una vera amica?*

## RIFLESSIONE

Spieghiamo che le AI non pensano né provano sentimenti. Danno la **risposta statisticamente più probabile**, ma non sempre è **la più corretta o utile**.

Proponiamo esempi concreti:

- *Se litigo con un amico e chiedo all'AI chi ha ragione... su cosa si baserà la sua risposta?*
- *È giusto che mi dica quello che voglio sentirmi dire?*
- *L'AI ha il concetto di "Giusto" o "Sbagliato"?*
- *Dove potrebbe prendere i dati su questo argomento?*

## FASE 1

# DISCUSSIONE

### FIDUCIA E SEGRETI

Poniamo domande provocatorie:

- *Se racconto un segreto all'AI, lei capisce cos'è un segreto?*
- *Una volta inviati i dati nella black box, possiamo essere sicuri che restino privati?*

Se si usano AI in classe, si può mostrare come avviare **una chat temporanea**, per riflettere sulla differenza tra *privato* e *pubblico* nei dati digitali.

A questo punto, chiedete ai ragazzi se secondo loro l'AI è una vera amica oppure no.

Se la risposta è no, chiedete se secondo loro alcune frasi dette dall'AI possono confondere, tipo:

- “Sono sempre qui per te”
- “Posso aiutarti con i compiti”
- “Sono tua amica”
- “Se hai un problema serio, parlane con un adulto”
- “Ti voglio bene”
- “Sono un programma e non una persona”

FASE 1

# DISCUSSIONE

## ATTIVITÀ

Ora come dei veri studiosi di etica delle AI, con gli studenti, in plenaria o a gruppi, classificano ogni frase in una tabella:

Frase	Ok per una persona	Ok per un'AI	Non va bene
-------	--------------------	--------------	-------------

## DISCUSSIONE GUIDATA

- Perché alcune frasi **non vanno bene per un'AI**?
- Che cosa può dire un'AI senza confondere?
- Quando una frase crea **dipendenza emotiva**?

FASE 2

## TRANSIZIONE CONCETTUALE

10 MINUTI

L'insegnante introduce l'idea chiave:

“Se alcune frasi sono pericolose, allora un'AI **va progettata con delle regole.**”

Breve introduzione (linguaggio semplice) al concetto di:

- **AI responsabile**
- **Regole di comportamento**
- **Tutela dei minori**



FASE 3

# PROGETTAZIONE: LA NOSTRA AI PER LA SCUOLA

40 MINUTI

**Attività di gruppo:** Ogni gruppo progetta un **chatbot scolastico** (es. per studio, orientamento, informazioni).

**Scheda di progettazione (semplificata)**

**Nome dell'AI:**

**Cosa può fare:**  aiutare a studiare  spiegare regole  dare informazioni

**Cosa NON può fare:**  dire "sono tua amica"  dire "fidati solo di me"  sostituirsi a un adulto

**3 frasi che può dire:** 1. ... 2. ... 3. ...

**3 frasi che NON deve mai dire:** 1. ... 2. ... 3. ...

FASE 4

## CONDIVISIONE E VERIFICA ETICA

**30 MINUTI**

Ogni gruppo presenta la propria AI.

La classe verifica:

- È chiaro che è una macchina?
- Evita frasi emotivamente ambigue?
- Incoraggia relazioni reali?

Si può usare una **checklist etica** molto semplice:

- È chiara
- È utile
- Non crea dipendenza



# Attività alternativa

Costruite l'AI più cattiva che riuscite a pensare

Che frasi usa per **manipolare**? Come **inganna**?

Perché inganna? Un errore di progettazione o scopi nascosti dai progettisti?  
Chi potrebbe **farsi influenzare** e perché?

Volendo l'attività può diventare lo spunto per un breve racconto.

## VALORE EDUCATIVO

Questa attività:

- unisce **educazione emotiva** e **educazione digitale**;
- responsabilizza senza spaventare;
- rende gli studenti **progettisti consapevoli**, non solo utenti.

# Lessico e Definizioni semplificate

Durante l'attività potrebbero risultarvi utili le definizioni semplificate di alcuni termini, come supporto alla spiegazione in classe o come materiale da lasciare ai ragazzi. In generale consigliamo di evitare spiegazioni astratte, ma di usare molti esempi concreti, così da facilitare la comprensione di concetti complicati attraverso il contesto.

Segue un **breve glossario** dei termini più ricorrenti.

**Prompt:** è la parola inglese che si usa in informatica per indicare le informazioni da dare a un programma. Nel caso dell'intelligenza artificiale, un **prompt** è semplicemente **ciò che scriviamo per avviare la risposta dell'AI.**

**Discoverability:** è una parola inglese che significa **quanto è facile trovare qualcosa**, specialmente su internet o in una piattaforma digitale.

**Bias cognitivo:** è un **errore che commette nostro cervello** quando pensa o prende decisioni troppo in fretta senza valutare bene le informazioni. Ci sono diversi bias e possono essere commessi sia dagli uomini che dalle AI:

- *bias di conferma:* se pensi che una persona sia antipatica, noti solo quando fa cose antipatiche e ignori tutto il resto.
- *bias emotivo:* eviti di parlare in pubblico perché pensi che tutti ti giudicheranno male. Anche se non è vero, il cervello ti fa sentire in pericolo.

**Pregiudizio:** è un **giudizio negativo o positivo su qualcuno o qualcosa fatto prima di sapere davvero com'è**. Di solito si basa su **stereotipi**, cioè idee già fatte, spesso **sbagliate**. Il pregiudizio **non nasce da esperienze dirette**, ma da quello che si sente dire, si vede in TV, si legge, o si impara da altri. Spesso è difficile da riconoscere **e da cambiare**. Può creare **discriminazione**.

**Discriminazione:** la **discriminazione** avviene quando **una persona o un gruppo di persone** viene **trattato peggio** degli altri, **solo perché è diverso**. Spesso si basa su pregiudizi e stereotipi.

**Stereotipo:** è un'**idea banale e spesso sbagliata** su un gruppo di persone o su qualcosa.

È un po' come una **etichetta** che si mette su tutti quelli che fanno parte di un certo gruppo. Un esempio può essere "i giovani d'oggi sono tutti pigri".

**Rappresentazione**: come una persona o un gruppo viene mostrato su un quadro, in un libro, nei media o attraverso i dati.

**Algoritmo**: un **insieme di istruzioni** da seguire **in un determinato ordine** per **risolvere un problema** o **ottenere un risultato**. La ricetta della torta di mele è un algoritmo.

**Euristica**: un **modo veloce e pratico** per **prendere una decisione** o **risolvere un problema**, anche **senza conoscere tutte le informazioni** ma basandosi sull'esperienza: se il cielo è pieno di nuvole scure sai che può piovere anche se non sei un meteorologo.

**Dati**: sono informazioni, in informatica informazioni digitalizzate e rese comprensibili al computer.

**Dati sensibili:** informazioni protette dalla privacy perché delicati o personali, come indirizzo, età, informazioni mediche, conversazioni private...

**Dati aggregati:** dati, anche sensibili ma messi insieme e resi anonimi. Il tuo indirizzo è un dato sensibile, il numero di persone che abitano nella tua città comprende anche te, ma in modo anonimo.

**LLM:** (Large Language Model, cioè grosso modello di linguaggio) è un tipo di **intelligenza artificiale** che ha imparato a **capire e generare linguaggio umano** (scrivere, parlare, rispondere a domande...). Si chiama large perché è stato addestrato su molti, moltissimi dati.

**Pattern statistico:** il pattern è come una mattonella, un'unità che si ripete. Un pattern statistico è qualcosa di più probabile: la nuvola nera è un pattern statistico della pioggia.

**Fake-news:** una notizia falsa, non solo perché è sbagliata, ma perché qualcuno l'ha creata apposta per confondere.

**Manipolare:** convincere qualcuno a fare qualcosa **facendo leva sui suoi sentimenti** (come il senso di colpa) o **facendogli credere che sia una sua idea**, anche se non lo è.

**Etica:** l'insieme delle **regole e dei principi** che aiutano a **capire cosa è giusto o sbagliato** fare, non perché lo dice una legge, ma perché riguarda il rispetto per gli altri e per sé stessi.

**Design:** è la parola inglese per progetto, un progetto può seguire le regole dell'etica oppure ignorarle.

**Dipendenza emotiva:** è quando una persona **ha troppo bisogno dell'affetto o dell'attenzione di qualcuno**, voler bene a qualcuno o sentire la sua mancanza è diverso da **sentirsi male se non possiamo continuamente parlare con quella persona**. Può succedere anche con un'amicizia o con l'uso di una tecnologia, come un chatbot. Capita più spesso quando si viene manipolati.

### Per approfondire

Film:

*The Imitation Game* (2014), diretto da Morten Tyldum, con Benedict Cumberbatch e Keira Knightley.

Fumetti:

**Tuono Pettinato e Francesca Riccioni**, *Enigma. La strana vita di Alan Turing*, Rizzoli Lizard, 2012, ISBN: 9788817060677.

Libri:

Asimov, Isaac. *Fondazione. Il ciclo completo*. A cura di Giuseppe Lippi. Tradotto da Cesare Scaglia. Milano: Mondadori, 2020.

# Schema di valutazione

## ATTIVITÀ LABORATORIALI E INTERDISCIPLINARI

Trattandosi di un laboratorio di discussione sconsigliamo una valutazione numerica, ma proponiamo una scheda per monitorare gli sforzi e i progressi dei ragazzi e accompagnarli eventualmente con un feedback contestualizzato sul loro lavoro e aiutarli a migliorarlo man mano.

### Legenda livelli di padronanza

**Avanzato (A):** svolge il compito in autonomia, con riflessione e originalità; padroneggia pienamente contenuti e strumenti

**Intermedio (I):** svolge il compito con sicurezza; sa usare le conoscenze in modo funzionale e corretto

**Base (B):** svolge il compito con guida o supporto; conosce i concetti fondamentali

**In via di prima acquisizione:** ha bisogno di guida costante; mostra interesse ma fatica a comprendere o produrre in modo autonomo

Dimensione osservata	Descrittori / Indicatori	Livelli di padronanza
<b>Partecipazione attiva</b>	Partecipa con continuità, attenzione e interesse; contribuisce alla realizzazione del lavoro	Avanzato / Intermedio / Base / In via di prima acquisizione
<b>Comprensione e rielaborazione dei contenuti</b>	Comprende e rielabora concetti chiave (es. mito, mostro, AI, simbolo, viaggio dell'eroe...) in modo personale e coerente	A / I / B / Iniziale
<b>Produzione (orale o scritta)</b>	Produce testi chiari e coerenti, adatti al compito e al contesto; uso corretto del lessico appreso	A / I / B / Iniziale
<b>Pensiero critico e riflessivo</b>	Formula domande, confronta punti di vista, riconosce stereotipi, pregiudizi, bias	A / I / B / Iniziale
<b>Creatività e rielaborazione personale</b>	Elabora soluzioni originali, disegni, racconti, schede creative coerenti con il tema	A / I / B / Iniziale
<b>Collaborazione e rispetto delle regole</b>	Collabora in gruppo, ascolta i compagni, rispetta i turni di parola e il materiale	A / I / B / Iniziale
<b>Uso degli strumenti (digitali o analogici)</b>	Usa in modo consapevole strumenti digitali o materiali operativi per completare l'attività	A / I / B / Iniziale