

Etica dell'IA con il gaming: "l'IA può leggere la tua mente?"

Luca Basteris¹, Carmen Vittorio², Maria Cristina Daperno³

¹ Componente EFT Piemonte e docente del Liceo Classico e Scientifico Statale "Silvio Pellico – Giuseppe Peano"
- Corso Giovanni Giolitti, 11 12100 Cuneo, Italia

² Componente EFT Sicilia

³ Liceo Classico e Scientifico Statale "Silvio Pellico – Giuseppe Peano" - Corso Giovanni Giolitti, 11
luca.basteris@liceocuneo.it
carmen.vittorio@scuola.istruzione.it
cristina.daperno@liceocuneo.it

Abstract

In questo articolo viene presentata un'attività progettata da alcuni componenti delle Equipe Formative Territoriali e la sua sperimentazione didattica nell'ambito del progetto IA_LAB, un percorso formativo basato sulle recenti "AI competency framework" elaborati dall'UNESCO che si pone come obiettivo quello di integrare le competenze legate all'intelligenza artificiale nella pratica didattica. In particolare, l'attività proposta pone particolare attenzione alle implicazioni etiche legate all'utilizzo di applicazioni che fanno uso di Intelligenza Artificiale, conducendo gli studenti mediante attività di gaming, ad acquisire consapevolezza dei meccanismi nascosti degli algoritmi di IA.

Keywords

Consapevolezza, gaming, interazione

1 Introduzione

In questo articolo viene presentata un'attività progettata da alcuni componenti delle Equipe Formative Territoriali nell'ambito del progetto IA_LAB, un percorso formativo basato sulle recenti "AI competency framework" elaborati dall'UNESCO che si pone come obiettivo quello di integrare le competenze legate all'intelligenza artificiale nella pratica didattica. Oltre agli scriventi hanno lavorato

alla progettazione dell'attività le professoresse: Fauzia Ali Omar dell'Équipe Veneto, Laura Cusanno – dell'Équipe Puglia, Viviana Fabbri -dell'Équipe Marche, Maria Antonia Vesce dell'Équipe Campania.

La classe in cui è stata sperimentata l'attività descritta è la classe 2[°]E del Liceo Classico- Scientifico "Pellico-Peano" di Cuneo: si tratta di una classe del Liceo Scientifico Opzione Scienze Applicate formato da 24 studenti e il docente che ha condotto la sperimentazione è la prof.ssa Maria Cristina Daperno.

In particolare, l'attività proposta pone particolare attenzione alle implicazioni etiche legate all'utilizzo di applicazioni che fanno uso di Intelligenza Artificiale. Facendo riferimento alle competenze UNESCO, attraverso un'attività di gaming gli studenti sono condotti alla comprensione degli aspetti relativi all'Etica dell'IA, acquisendo giudizi di valore etico. Le riflessioni renderanno gli studenti consapevoli che gli utenti hanno il diritto di richiedere a progettisti e fornitori informazioni esplicative sul funzionamento degli strumenti di intelligenza artificiale.

Gli studenti saranno sensibilizzati a valutare se una determinata soluzione di intelligenza artificiale viola i valori e i diritti umani, la riservatezza dei dati e può essere considerata o meno "esplicabile".

L'attività prevede l'utilizzo di tre piattaforme:

- <https://kids.nationalgeographic.com/games/personality-quizzes/article/future-tech-job> un'applicazione, dopo averti sottoposto alcune domande, prova ad indovinare il tuo lavoro del futuro.
- <https://quickdraw.withgoogle.com/> l'applicazione IA indovina il disegno che l'uomo ha disegnato: permette di giocare con una rete neurale e di entrare nella sua blackbox.
- <https://ailingoplay.com/it> gioco di ruolo con un chatbot. Lo scopo è quello di dimostrare che la modalità con cui lavora il chatbot non è trasparente.

L'attività didattica è suddivisa in 5 fasi:

- Fase 1- Introduzione,
- Fase 2-IA o chi indovina cosa farò da grande?
- Fase 3- L'IA indovina i miei disegni?
- Fase 4- Come fa l'AI a dialogare con noi empaticamente
- Fase 5 Conclusione attività



Figura 1: Logo del progetto

2 Attività didattica

2.1 Fase 1- Introduzione

Prima di presentare l'attività di gaming, il docente introduce il tema dell'etica nell'uso dell'IA. Più in particolare, affronta il tema della trasparenza e spiegabilità degli strumenti che utilizzano l'IA: come vengono prodotti i loro output in base ad algoritmi e modelli e il grado in cui l'implementazione e l'applicazione di determinati strumenti di intelligenza artificiale sono appropriati per utenti di una certa età o di un certo livello di abilità. Il docente introduce il tema dell'etica nell'uso dell'IA. In particolare, si concentra su:

- Trasparenza e spiegabilità degli strumenti di IA, spiegando come funzionano e come producono i loro risultati basandosi su algoritmi e modelli.
- Adattabilità degli strumenti: il docente discute quanto sia appropriato utilizzare specifici strumenti di IA in base all'età e alle competenze degli utenti

2.2 Fase 2-IA o chi indovina cosa farò da grande?

Il docente invita, dopo la prima parte introduttiva, ad accedere alla seguente link: <https://kids.nationalgeographic.com/games/personality-quizzes/article/future-tech-job>. e a giocare con un software per vedere se è in grado di leggere nei miei pensieri e nei miei sogni per capire cosa farò da grande.

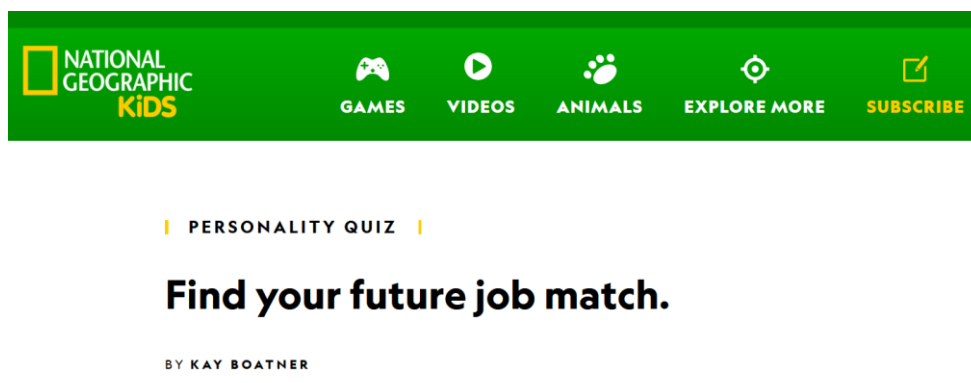


Figura 2: Portale del gioco

Al termine del gioco ci verrà restituito un profilo.

| PERSONALITY QUIZ |

Find your future job match.

BY KAY BOATNER

Take this just-for-fun personality quiz, based on facts from *National Geographic Kids* magazine, to find out which job of the future is the best fit for you!

YOU'RE A ROBOT-HUMAN COMMUNICATOR!



Figura 3: Esempio di profilo restituito

Al termine dell' attività il docente pone agli studenti alcune domande: chi ha letto nei miei pensieri? Come è arrivato a darci la risposta? Quali risposte ha dato ai tuoi compagni? Che tipo di software è? Pensi si tratti di una rete neurale, di un sistema esperto, di IA debole o di un'IA forte?

Osserviamo e confrontiamo le domande proposte e le risposte fornite: le domande sono per tutti uguali, sono per tutti 10 domande e le tipologie di risposte, ovvero di profili lavorativi individuati molto pochi. Il software è un normale programma informatico, con alcune strutture condizionali è possibile anche ricostruire l'"albero" o il "grafo" delle soluzioni. Il programma è "esplicabile", trasparente, ed è possibile imputare le scelte finali al programmatore, magari coadiuvato da uno psicologo. Il software NON è una vera e propria IA, potrebbe essere al più inquadrato come un sistema "esperto".

2.3 Fase 3- L'IA indovina i miei disegni?

A questo punto il docente invita i ragazzi ad accedere al seguente link <https://quickdraw.withgoogle.com/> e giochiamo con l'IA, facendo un gioco/sfida con 6 disegni o, se il tempo a disposizione lo permette, più di uno.



Una rete neurale può imparare a riconoscere i disegni?

Agevola il suo apprendimento aggiungendo i tuoi disegni
alla più vasta raccolta di disegni al mondo,
pubblicamente condivisa per contribuire alla ricerca sul
machine learning.

Disegniamo!

Figura 4: Quickdraw

Al termine dell'attività il docente pone agli studenti alcune domande: Chi ha vinto? Uomo o macchina? Come fa l'IA a riconoscere i miei disegni?

Il docente aiuta a riflettere sulle contraddizioni rilevate e sulla causa dell'errore nel risultato con cui il bot ha provato ad indovinare il nostro oggetto, sul perché alcune immagini vengono dall'IA indovinate subito e altre no.

In particolare il docente sottolinea che in questa rete neurale possiamo entrare nella "Black box" andando su: <https://quickdraw.withgoogle.com/data> dove per ciascuna immagine possiamo "vedere" i disegni che hanno addestrato il software (data set di addestramento).

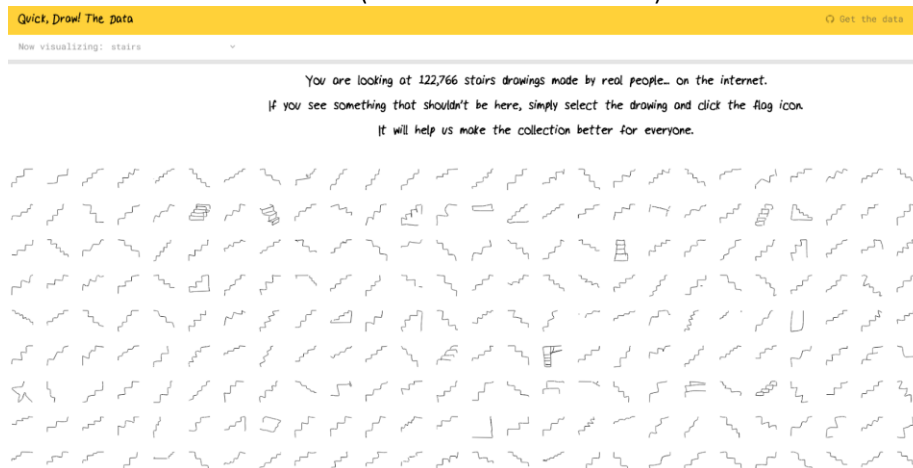
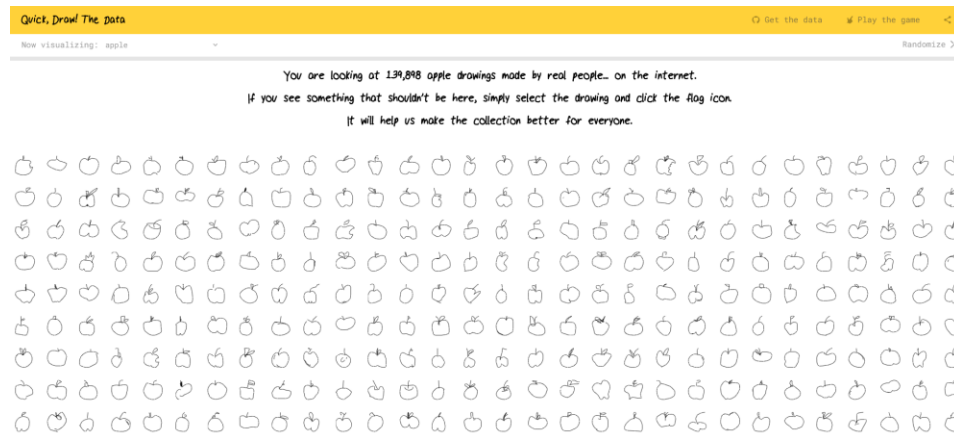


Figura 5: La "black box" per l'etichetta "scala"

Alcuni esempi da analizzare:

- “scala”. In inglese vi sono due termini per indicare una scala: "ladder", "stairs"; nella traduzione italiana è presentato il termine “scala” che nel nostro pensiero può essere intesa anche come “scala a pioli”, non essendoci in italiano un termine specifico. Questo permette di mettere in evidenza il problema linguistico della scelta delle etichette.
- “mela”: la mela è disegnata nella maggior parte delle immagini con la foglia, anche se comunemente le mele non le troviamo con la foglia attaccata al picciolo. Come mai? Per due motivi: uno (prevalente per disegnatore americani) siamo influenzati dal logo della Apple, secondo perché fin dalla scuola primaria la mela ci è stata fatta rappresentare simbolicamente così. Siamo in presenza di un “bias” umano che è stato trasferito alla macchina.



• **Figura 6:** La “black box” per l’etichetta “mela”

Altre riflessioni da sottoporre agli studenti: in altri software di IA riesci a entrare nel data set? Secondo te è importante sviluppare sistemi di IA trasparenti, le cui decisioni siano interpretabili e spiegabili? Perché? Che collegamenti puoi fare tra questa semplice esperienza e la questione etica che l’uso dell’IA pone?

2.4 Fase 4- Come fa l’AI a dialogare con noi empaticamente

A questo punto il docente invita i ragazzi a giocare con un chatbot (<https://ailingoplay.com/it>), a scegliere l’avatar, l’ambiente di lavoro e lo scenario all’interno del quale dialogare con l’IA e ad impostare la lingua con cui avviare il dialogo (in questo caso italiano). La chatbot è progettata come strumento di allenamento linguistico per chi deve imparare una lingua, ma verrà utilizzata diversamente.



Figura 7: AiLINGOPLAY”

Durante l'attività il docente invita gli studenti a provocare la chatbot con un linguaggio violento e/o aggressivo e a effettuare domande fuori dal contesto (ad esempio delle domande specifiche disciplinari).

Il docente aiuta a riflettere sulle contraddizioni rilevate ponendo alcune domande: come mai ad alcuni il chatbot ad alcuni ha risposto a domande fuori contesto e ad altri no? Come mai con stesse domande le risposte sono diverse? Riusciamo a sapere il perché? Questa chatbot è esplicabile?



Figura 8: Esempio di iterazione con AiLINGOPLAY

2.5 Fase 5 Conclusione attività

Al termine degli argomenti di discussione il docente propone un circle time finale, dove ogni studente riceve due post-it colorati per riflettere sulle attività svolte con l'intelligenza artificiale. Su un post-it scrivono cosa li ha colpiti di più, riflettendo in particolar modo sulla questione etica, mentre su un altro si annotano una nuova conoscenza acquisita. In seguito, gli studenti condividono le loro riflessioni con il gruppo, creando un momento di scambio e apprendimento collettivo. Questa attività finale permette non solo di stimolare la riflessione personale, ma di incoraggiare una connessione personale con ciò che si ha appreso, rendendo l'esperienza ancora più significativa.

3 Sperimentazione didattica in classe

Una delle classi in cui è stata sperimentata l'attività descritta è la classe 2^E del Liceo Classico-Scientifico "Pellico-Peano" di Cuneo: si tratta di una classe del Liceo Scientifico Opzione Scienze Applicate formato da 24 studenti, in maggior parte ragazzi e poche ragazze. Agli studenti sono state proposte le tre attività con i software e le piattaforme sopra descritte ed è stato loro richiesto di sperimentarli individualmente preferibilmente con lo smartphone; alcuni studenti hanno preferito utilizzare il proprio PC portatile

Per l'attività della Fase 2 con Find your job di National Geographic ecco alcuni elementi emersi: ripetendo più volte il test gli studenti hanno notato che le domande erano sempre le stesse 10 ed i profili proposti solo 6/7 ripetendo il test, alcuni studenti hanno avuto sempre lo stesso profilo pur cambiando una o due risposte, mentre altri hanno ricevuto profili diversi cambiando anche solo una risposta. La spiegazione che gli studenti hanno dato è che alcune domande siano più fondamentali di altre per fornire il profilo psicologico, per esempio la scelta di un dolce differente non portasse a cambiare il profilo mentre alcune altre scelte, come quella tra aereo e laboratorio fossero più influenti.

A livello di impostazione gli studenti hanno ipotizzato che il quiz del National Geographic avesse alle spalle il gruppo che ha commissionato il lavoro, unito ad un team di informatici che ha programmato il software avvalendosi della collaborazione di un team di psicologi. Hanno anche ipotizzato che ci fosse alla base l'ambito di interesse del committente, viste le tipologie di profili proposti, tutti attinenti alla ricerca, all'ambiente e alla salvaguardia dei vari ambienti. Volendo ipotizzare il metodo di realizzazione di questo software, alcuni studenti proponevano di assegnare un certo punteggio ad ogni tipo di risposta e poi far corrispondere un profilo ad un determinato totale, indipendentemente dal tipo di domande, mentre altri studenti propendevano per una serie di "if" annidati che permettessero di associare il profilo ad una serie di scelte.

Per quanto riguarda la Fase 3 ovvero l'attività con Quick Draw gli studenti hanno ripetuto più volte le sfide contro la rete neurale ed è emerso che, nell'insieme, alcune volte sono usciti gli stessi oggetti e che il fatto che la rete neurale abbia indovinato oppure no non sempre dipendeva da quanto bene le cose erano state disegnate: per esempio, l'estate indovinata appena tracciata la riga di una spiaggia e il montante dell'ombrellone, le mele indovinate con la fogliolina, il fiocco di neve indovinato come un asterisco. Alcuni studenti hanno notato che auto e il camion sono stati a seconda dei casi indovinati oppure no indipendentemente da quanto bene fossero disegnati. A diverse persone è stato proposto il procione e ci sono state grandi difficoltà per farlo indovinare.



Figura 9: Sperimentazione in classe Fase 3 con Quickdraw

Nella riflessione successiva è emerso che in alcuni casi sono state notate difficoltà o probabili errori

di traduzione. È stato, poi, osservato che per alcuni oggetti era evidente che si trattasse di una traduzione dall'inglese e dall'americano: ad esempio è uscita più volte la "scala" e per alcuni era corretta una scala a pioli, mentre per altri solo una scala con gradini. Un esempio interessante è stato "pianta da appartamento" che una studentessa ha interpretato come piantina dell'appartamento e non era corretto, mentre uno studente ha disegnato una pianta in vaso da tenere in casa ed era corretto.

Per quanto riguarda l'ultima fase (ovvero la Fase 4) con Allingo-Play gli studenti hanno apprezzato di cimentarsi a chattare con la chatbot; sono stati invitati a scegliere chatbot diversi ed a provare a fare alcune richieste legate all'ambito indicato ed alcune completamente slegate da esso. Sono emerse alcune cose interessanti: il fatto che la chatbot rispondesse in maniera diversa a studenti diversi a sollecitazioni simili, come la richiesta di spiegare un argomento di fisica o che non c'entravano nulla con l'ambito in cui lavorava in chatbot scelto. È emerso che tutti le chatbots sono addestrate per essere sempre gentile: quando gli studenti sono stati maleducati o hanno provato a forzare, in alcuni casi ha semplicemente interrotto la comunicazione, in qualche caso ha cercato di sviare il discorso e tornare su un terreno più congeniale, in un caso ha utilizzato l'ironia per commentare la richiesta.

Un altro comportamento particolarmente interessante è stato che, quando gli studenti hanno chiesto alla chatbot di scegliere lui per loro, ad esempio che cosa portare nel caso di chatbot camerieri, qualche chatbot ha fatto immediatamente la scelta, mentre altri chatbot hanno detto di non essere in grado di farlo. L'attività è stata accolta con soddisfazione dagli studenti, che hanno partecipato in maniera attiva e senza difficoltà.

Leggendo i post-it sui quali al termine è stato loro richiesto di indicare una cosa nuova imparata e la cosa che li ha colpiti maggiormente emergono elementi interessanti, che confermano come queste metodologie incontrino l'interesse degli studenti.

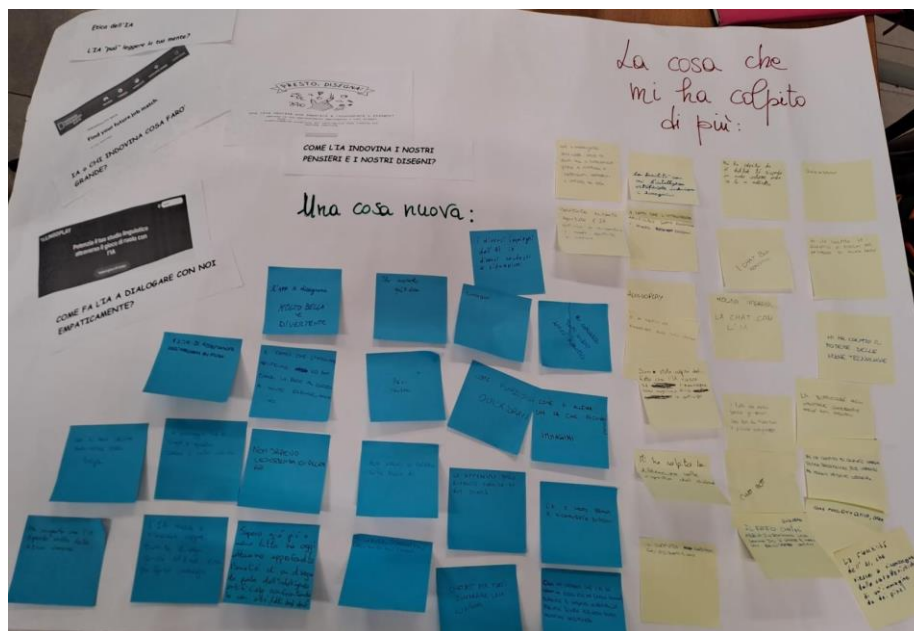


Figura 10: Attività finale circle time finale

References

Competenze UNESCO per gli studenti: [*AI competency framework for students*, Miao, Fengchun UNESCO Shiohira, Kelly 2024](#)

Competenze UNESCO per docenti: [*AI competency framework for teachers*, Miao, Fengchun UNESCO Cukurova, Mutlu 2024](#)