

Intelligenza Artificiale: Usi e Temi per l'UX



Dr. ANDREA CAPUTO, PhD
Politecnico di Torino

andrea.caputo@unito.it

Cosa vedremo

1. Intelligenza umana e intelligenza artificiale
2. The bitter lesson
3. Il problema della explainability



Intelligenza Umana e Artificiale si assomigliano?

Gli umani sono limitati nell'uso di:

- Comunicazione
- Computazione/calcolo
- Risorse cognitive

Trends in Cognitive Sciences

Series: Machine Behavior

Opinion

Understanding Human Intelligence through Human Limitations

Thomas L. Griffiths^{1,2,*}

Recent progress in artificial intelligence provides the opportunity to ask the question of what is unique about human intelligence, but with a new comparison class. I argue that we can understand human intelligence, and the ways in which it may differ from artificial intelligence, by considering the characteristics of the kind of computational problems that human minds have to solve. I claim that these problems acquire their structure from three fundamental limitations that apply to human beings: limited time, limited computation, and limited communication. From these limitations we can derive many of the properties we associate with human intelligence, such as rapid learning, the ability to break down problems into parts, and the capacity for cumulative cultural evolution.

CellPress
REVIEWS

Highlights

Humans are limited in time, computation, and communication, defining a set of computational problems that human intelligence has to solve.

Considering the structure of these computational problems can help us to understand why human minds have some of the characteristics that they do.

The solutions to these problems involve mathematical formalisms such as Bayesian inference and meta-learning, rational meta-reasoning, and distributed algorithms, which may be particularly relevant to cognitive science.

A cosa serve l'IA?



Ottenere un modello da pochi dati

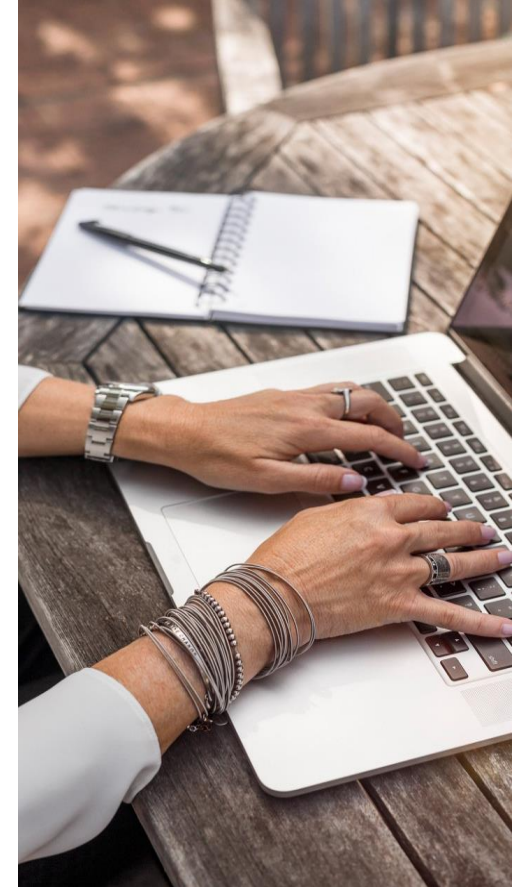
- Gli umani usano il ragionamento induttivo (e talvolta sbagliano)
- Le AI non sono progettate per imparare da pochi dati (sbagliano più degli umani)
- Le attuali possibilità di computazione ci permettono di avere grandi risorse computazionali
- “Secondo voi, in cosa l'AI assomiglia all'intelligenza umana? E in cosa se ne differenzia?”



Human vs AI



- L'umano ragiona con poche risorse cognitive (poche informazioni, poca memoria) → ragionamento induttivo.
- L'AI usa tantissimi dati e potenza computazionale (ma non è sostenibile all'infinito).
- Esempio pratico: L'umano riconosce un volto visto una sola volta (pochi dati). L'AI ha bisogno di milioni di immagini per fare lo stesso.



La “bitter lesson”



«Immaginate di imparare a cucinare: io studio le ricette della nonna, l’AI invece prova 10 milioni di combinazioni di ingredienti in un secondo. Chi troverà prima il piatto perfetto?»

Gioco – Trovate i numeri primi, da 1 a 200.

Umano: pensa a regole per capire quali numeri sono primi
→ scorciatoie, strategie, memoria.

IA (ChatGPT): “Dammi tutti i numeri primi fino a 200.” Te li fornisce subito, perché sfrutta calcolo e dati, non strategie “umane”, ma una grande mole di dati.

👉 Messaggio: noi ragioniamo con regole; l’IA vince col calcolo massivo.



La “bitter lesson”



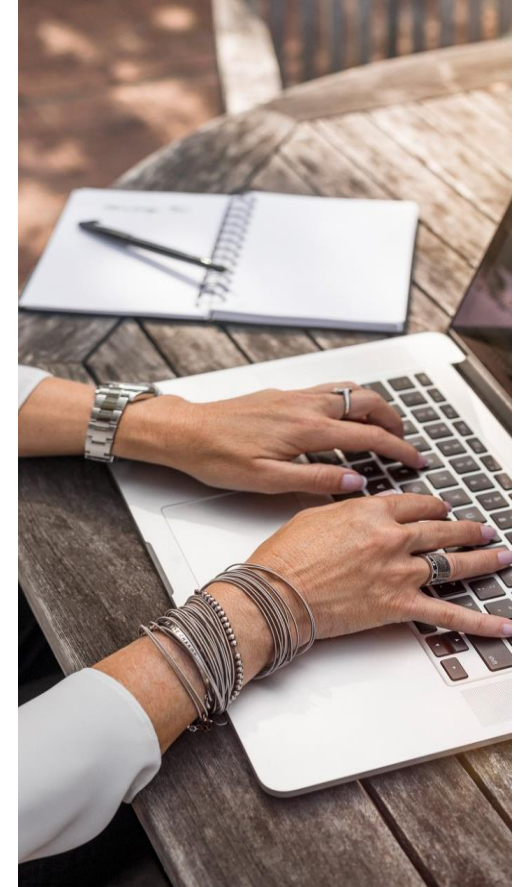
- Gli umani riescono a imparare da pochi esempi grazie al ragionamento **induttivo**, ma possono commettere errori (stereotipi, bias). Pochi dati → regola generale (partendo da casi piccoli). Ma...
- ...Le AI (machine learning, deep learning) hanno bisogno di tantissimi dati e molta potenza computazionale, ma non imparano facilmente da pochi casi. Macinano tantissimi dati. Quando chiedi di fare un modello da pochi dati, sbagliano tantissimo (es. anche citazioni bibliografiche ex novo).
- Tema chiave: le risorse. L'AI sembra avere risorse illimitate (più dati, più calcolo), ma in realtà ci sono **limiti energetici e di sostenibilità**.



La “bitter lesson”



- Nella storia dell'AI, ottenere più dati e più potenza di calcolo ha sempre vinto sugli approcci che cercano di inserire bias “intelligenti” progettati da umani.
- Conseguenza: i sistemi diventano **black box**, difficili da spiegare e comprendere.
- Ma la **fiducia** (specialmente in campi critici come la medicina) dipende anche dalla **explainability**: **capire come un modello è arrivato a una decisione.**
- Anche il GDPR (art. 22) riconosce il “diritto alla spiegazione” delle decisioni automatizzate.



Modi di risolvere un problema



Human: Dividere un problema in sotto-task.

Gli umani risolvono problemi complessi (es. «arrivare in tempo a scuola») spezzandoli in sotto-task → questo li rende più comprensibili e insegnabili.

Le AI invece non sempre seguono questa logica: cercano soluzioni “dirette”, senza una struttura intermedia facilmente leggibile.

Bias: gli umani hanno **bias cognitivi**, ma anche i modelli di AI ereditano **bias dai dati**. *Es. caso di Amazon (vedi slide successive)*

👉 Esercizio: come affrontereste un compito complesso come **progettare un percorso (cioè i passaggi da eseguire) in un'app?**

Confrontiamo al risposta con chatgpt

Modi di risolvere un problema



Es. Cubo di Rubick.

Pensare a tante strategie, una alla volta.

IA: tanti processi **paralleli**, non **sequenziali**. Vantaggio, ma problema: non è spiegabile.

IA 20 mosse. Human: circa 70 mosse.

Metodo Human: meno efficace, ma insegnabile

Metodo IA: non insegnabile.

Gli algoritmi di deep learning e machine learning non devono necessariamente suddividere gli obiettivi **in sotto-task**.

Questo aspetto però peggiora l'explainability (cioè: *come ha fatto? Come è arrivato a questa soluzione?*)

SIAM J. DISCRETE MATH.
Vol. 27, No. 2, pp. 1082–1105

© 2013 Society for Industrial and Applied Mathematics

THE DIAMETER OF THE RUBIK'S CUBE GROUP IS TWENTY*

TOMAS ROKICKI[†], HERBERT KOCIEMBA[‡], MORLEY DAVIDSON[§], AND JOHN
DETHRIDGE[¶]

Abstract. We give an expository account of our computational proof that every position of Rubik's Cube can be solved in 20 moves or less, where a move is defined as any twist of any face. The roughly 4.3×10^{19} positions are partitioned into about two billion cosets of a specially chosen subgroup, and the count of cosets required to be treated is reduced by considering symmetry. The reduced space is searched with a program capable of solving one billion positions per second, using about one billion seconds of CPU time donated by Google. As a byproduct of determining that the diameter is 20, we also find the exact count of cube positions at distance 15.

Key words. group theory, algorithm performance, Rubik's Cube

AMS subject classifications. 20-04, 05C12, 20B40

DOI. 10.1137/120867366

Modi di risolvere un problema



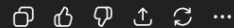
ChatGPT ▾

con quante mosse risolveresti il cubo di rubick?

Se parliamo del **cubo di Rubik standard $3 \times 3 \times 3$** , la risposta dipende da cosa intendi:

1. **In generale per chi lo risolve "a mano"**: la maggior parte dei metodi classici (CFOP, Roux, ZZ, ecc.) richiede tra **50 e 60 mosse** in media per risolvere un cubo partendo da una posizione casuale, se si conta ogni rotazione di faccia come una mossa. I speedcuber esperti possono arrivare anche a **40–50 mosse** mediamente.
2. **Ottimale matematicamente (God's Number)**: la soluzione più corta possibile per qualsiasi configurazione del cubo è stata calcolata dai matematici e dai computer: **al massimo 20 mosse**. Questo numero è noto come **God's Number**, e significa che qualunque configurazione del cubo può essere risolta in 20 mosse o meno, se usi la sequenza perfetta.

Se vuoi, posso spiegarti **come calcolare una soluzione vicino al numero minimo di mosse** usando algoritmi conosciuti, senza dover fare tentativi a mano. Vuoi che lo faccia?



Black box

come fai a farlo in 20 mosse?

Farlo in 20 mosse **non è una questione di intuizione o di metodi classici da manuale**, ma di usare algoritmi di ottimizzazione estremamente avanzati. Ecco come funziona:

1. **God's Number** è stato calcolato con l'aiuto di computer potentissimi che hanno analizzato **tutte le 43 quintilioni di configurazioni possibili** del cubo $3 \times 3 \times 3$. Non è un trucco manuale: è pura forza bruta combinata con ottimizzazioni matematiche.
2. **Come funziona l'algoritmo**:
 - Ogni stato del cubo viene rappresentato in modo matematico (spesso come permutazioni dei cubi più piccoli).
 - Si costruisce un **albero di mosse** partendo dallo stato iniziale.
 - Algoritmi come *IDA (Iterative Deepening A)*** cercano sequenze di mosse sempre più corte fino a trovare la soluzione minima.
 - Ottimizzazioni come **pattern databases** riducono drasticamente lo spazio di ricerca, evitando di esplorare tutte le mosse possibili una per una.
3. **Per una persona reale**: un speedcuber non riesce a "vedere" la soluzione in 20 mosse a occhio nudo. Anche i migliori cubisti possono risolvere un cubo in pochi secondi, ma di solito usando **40–60 mosse**, perché memorizzano sequenze (algoritmi) efficienti, non soluzioni matematicamente minime.

In pratica, risolvere un cubo in 20 mosse è oggi **possibile solo tramite computer o algoritmi ottimizzati**, non a mano.

Problema dell'explainability



È molto complesso comprendere la struttura dei risultati che si ottengono dal machine learning o da tecniche di deep learning.

Quella che si definisce "Black-box".

È necessario comprendere come funziona un'intelligenza artificiale?

Gli umani hanno bisogno che le cose siano «explainable» → *need for cognition*

*Tu da chi ti faresti operare? Da un umano con 20 anni di esperienza
o con un robot guidato da IA?*

Problema di trust



[Review](#) > [Eur Heart J.](#) 2019 Jul 1;40(25):2058-2073. doi: 10.1093/eurheartj/ehz056.

Deep learning for cardiovascular medicine: a practical primer

Chayakrit Krittanawong ^{1 2}, Kipp W Johnson ³, Robert S Rosenson ², Zhen Wang ^{4 5}, Mehmet Aydar ⁶, Usman Baber ², James K Min ⁷, W H Wilson Tang ^{8 9 10}, Jonathan L Halperin ², Sanjiv M Narayan ¹¹

Affiliations + expand

PMID: 30815669

[Free PMC article](#)

[Review](#) > [Cancer Commun \(Lond\).](#) 2021 Nov;41(11):1100-1115. doi: 10.1002/cac2.12215.

Epub 2021 Oct 6.

Artificial intelligence for assisting cancer diagnosis and treatment in the era of precision medicine

Zi-Hang Chen ^{1 2}, Li Lin ¹, Chen-Fei Wu ¹, Chao-Feng Li ³, Rui-Hua Xu ⁴, Ying Sun ¹

Affiliations + expand

PMID: 34613667 PMCID: [PMC8626610](#) DOI: [10.1002/cac2.12215](#)

[Free PMC article](#)

Problema di trust



AI in medicine must be explainable

[Shinjini Kundu](#) ✉

[Nature Medicine](#) 27, 1328 (2021) | [Cite this article](#)

5912 Accesses | 57 Citations | 130 Altmetric | [Metrics](#)

AI algorithms used for diagnosis and prognosis must be explainable and must not rely on a black box.

In the wake of new medical devices powered by artificial intelligence and machine learning (AI/ML), a debate has erupted about whether AI's underlying logic must be intelligible. To gauge the popular sentiment, a conference posed the following question to its attendees: suppose you have cancer and need surgery to remove the tumor. Which of the two surgeons would you pick if you had to choose between a human surgeon, with a 15% change of dying, or a robot surgeon, with a 2% chance of dying—with the caveat that no one knows how the robot operates and no questions may be asked of it? All but one of the attendees [preferred the human](#).



Tu da chi ti faresti operare per rimuovere un tumore?

- *Da un umano con il 15% di chance di morire o...*
- *un chirurgo-robot guidato da IA con un 2% di morire?*

Sentiment: tutti tranne 1 hanno detto Human → problema di **trust**

Problema di trust



Non ci importa che l'IA possa essere più infallibile.

Ma che l'umano possa capire di più il paziente, capire meglio cosa fare. Siamo disposti ad accettare di più il rischio dato dall'umano, ma di cui possiamo **comprendere** meglio gli stati mentali, di cui abbiamo più **fiducia**.

Nella UX un pezzo del vostro compito è guadagnarvi la *fiducia degli utenti*

Perchè le persone usano una tecnologia?

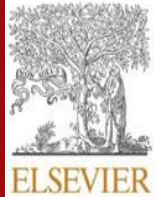


- Perceived value of technology (**costi/benefici**) → ChatGPT è gratis
- Perceived trust of technology (**affidabilità percepita**) → legata alla explainability
- Explainability (**comprensione**)? → Se non siamo in grado di spiegare come funziona e come ci si è arrivati, c'è un problema di trust...

Come ottenere fiducia?



International Journal of Medical Informatics 168 (2022) 104884



Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf



Does AI explainability affect physicians' intention to use AI?

Chung-Feng Liu^a, Zhih-Cherng Chen^b, Szu-Chen Kuo^{c,*}, Tzu-Chi Lin^{d,*}

^a Department of Medical Research, Chi Mei Medical Center, No. 901, Zhonghua Rd., Yongkang Dist., Tainan City 710402, Taiwan

^b Division of Cardiology, Department of Internal Medicine, Chi Mei Medical Center, No. 901, Zhonghua Rd., Yongkang Dist., Tainan City 710402, Taiwan

^c Department of Pediatrics, Chi Mei Medical Center, No. 901, Zhonghua Rd., Yongkang Dist., Tainan City 710402, Taiwan

^d Department of Nursing, Chi Mei Medical Center, Liouying, No. 201, Taikang Taikang Vil., Liouying Dist., Tainan City 73657, Taiwan

ARTICLE INFO

Keywords:

Physician
Artificial intelligence (AI)
AI explainability (XAI)
Behavioral intention
Technology trust
Perceived value

ABSTRACT

Background: Artificial Intelligence (AI) is increasingly being developed to support clinical decisions for better health service quality, but the adoption of AI in hospitals is not as popular as expected. A possible reason is that the unclear AI explainability (XAI) affects the physicians' consideration of adopting the model.

Purpose: To propose and validate an innovative conceptual model aimed at exploring physicians' intention to use AI with XAI as an antecedent variable of technology trust (TT) and perceived value (PV).

Methods: A questionnaire survey was conducted to collect data from physicians of three hospitals in Taiwan. Structural equation modeling (SEM) was used to validate the proposed model and test the hypotheses.

Results: A total of 295 valid questionnaires were collected. The research results showed that physicians expressed a high intention to use AI. The XAI was found to be of great importance and had a significant impact both on AI TT and PV. We also observed that TT is AI had a significant impact on PV. Moreover, physicians' PV and TT in AI

- Migliorando l'explainability
- Abbiamo confermato che la trust è un fattore significativo che influenza la intenzione dei medici nei confronti dell'uso dell'IA. Pertanto, gli sviluppatori di IA medica potrebbero concentrarsi sul superamento delle questioni di fiducia che preoccupano i medici, in particolare l'interpretabilità degli algoritmi di Machine learning.

L'explainability è necessaria



*Article 22 of the General Data Protection Regulation (GDPR) sets out the rights and obligations of the use of automated decision making. Noticeably, it introduces the right of explanation by giving individuals **the right to obtain an explanation of the inference(s) automatically produced by a model, confront and challenge an associated recommendation, particularly when it might negatively affect an individual legally, financially, mentally or physically.** By approving this GDPR article, the European Parliament attempted to tackle the problem related to the propagation of potentially biased inferences to society, that a computational model might have learnt from biased and unbalanced data.*

Caso Amazon dello screening CV

Un algoritmo per scegliere il miglior candidato si basa su informazioni che immettono gli esseri umani (primo bias) e che si riferiscono a situazioni presenti o passate --> cerchiamo i migliori candidati sulla base dei candidati più performanti già presenti in azienda. Essendoci molti uomini, l'algoritmo tendeva ad escludere le donne → Bias e discriminazione di genere.

Esempi di explainability



Immagina un'azienda che usa un algoritmo per valutare i CV dei candidati e decidere chi invitare a un colloquio.

L'algoritmo analizza variabili come: anni di esperienza, titoli di studio, competenze, risultati nei test...

Per ogni candidato, genera un punteggio da 0 a 100 che indica la “probabilità di successo nel ruolo”.

Senza explainability → Il sistema dice solo: «Il candidato Andrea ha un punteggio di 82». Ma non spiega perché. → L'HR non può capire su quali basi è stata presa la decisione (rischio di discriminazioni, sfiducia nel modello, difficoltà di *audit*).

Con explainability → Il sistema aggiunge una spiegazione: «Il punteggio di 82 deriva da: +15 punti per esperienza > 5 anni nel settore, +10 per test di problem solving sopra la media, -5 per scarsa conoscenza di programmazione Python». Oppure mostra una visualizzazione grafica (es. un diagramma a barre) che evidenzia il contributo di ogni fattore alla decisione finale. → L'HR ora capisce come e perché il modello ha deciso così, e può verificare se il processo è **equo** e **coerente**.

Esempi di exaplainability



Un'app di fitness
ti suggerisce di dormire di più.

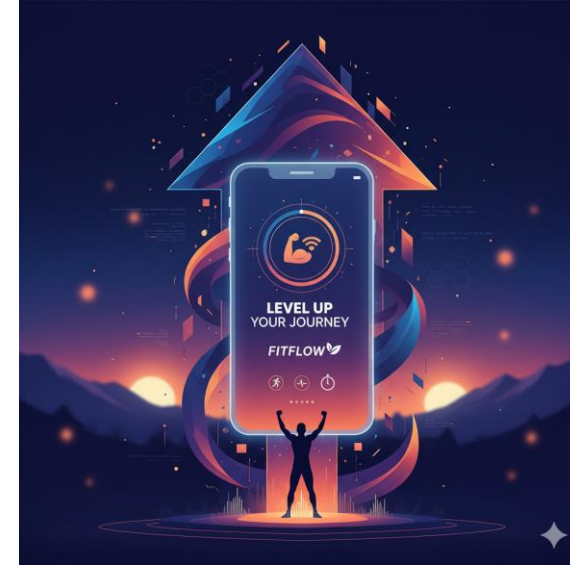


Immagine creata ad hoc per questo documento
con la AI Google Gemini

Senza explainability → «Ti consigliamo di dormire 8 ore stanotte»... why?!

Con explainability → «Ti consigliamo di dormire 8 ore perché ieri hai dormito 6h20 e il tuo battito medio è aumentato del 10%» → L'utente capisce la logica del consiglio e si fida di più dell'app.

SCENARIO: mondo organizzativo

L'Intelligenza Artificiale (IA) è stata introdotta nell'ultimo decennio all'interno dei processi di Reclutamento e Selezione del personale per supportarli e velocizzarli

ESEMPI

- Big Data
- Face Recognition
- Voice Interaction



ATTIVITÀ

- estrazione dai curricula tramite analisi del profilo professionale
- assunzione del candidato (teoria)
- colloquio di lavoro (teoria)
- proposta di contratto (teoria)

ACCETTAZIONE

Le persone (in generale) trascorrono sempre più tempo online. Quindi è necessario intercettare i potenziali candidati attraverso i giusti canali.

Studio di van Esch e Black (2019)

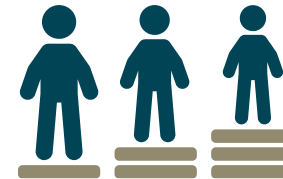
Valutare quali fattori influenzano i candidati di nuova generazione (18-35 anni) ad intraprendere una processo di selezione completamente gestito da IA (atteggiamento di accettazione delle nuove tecnologie)

USO DEI SOCIAL



RICOMPENSE
INTRINSECHE

PERCEZIONE DI
EQUITÀ



MODA

CANDIDATE EXPERIENCE

Queste tecnologie permettono, ad esempio, di fissare qualsiasi orario per svolgere un'intervista, facilitando la candidate experience che permette di favorire un atteggiamento positivo nei confronti dell'azienda (Employer Branding).

Infatti, il report Job Vite (2017) mostra che solo l'8,52% dei candidati che applicano per una posizione completano tutti gli step necessari e molti di loro, **se non ricevono un feedback** che segue la loro candidatura, non si candideranno per future posizioni aperte della stessa azienda.



CANDIDATE EXPERIENCE: UN ESEMPIO

VIDEO-INTERVISTA REGISTRATA

Alcune aziende chiedono ai candidati di registrare un video in cui rispondono a delle domande che compaiono sullo schermo, all'interno di un limite di tempo.

Ansia: comparsa improvvisa delle domande, poco tempo per pensare e rispondere

Studio (Tiltson et al. 2023): aumentare il **tempo di preparazione** o la **possibilità di registrare una seconda volta** favorisce reazioni positive.



ETICA E BIAS

General Data Protection Regulation 2016/679 (GDPR) è un regolamento del diritto dell'UE sulla protezione dei dati e della privacy. Es. Articolo 9: proibisce il processamento di dati su appartenenza etnica, opinioni politiche, credenze religiose e filosofiche, iscrizione a sindacati, dati biometrici, dati sulla salute e sull'orientamento sessuale

Bias: scorciatoie cognitive che portano a conclusioni veloci. Sono euristiche inefficaci che si basano su pre-giudizi

Il caso Amazon: Bias discriminazione di genere



STUDIO ESPLORATIVO ITALIANO

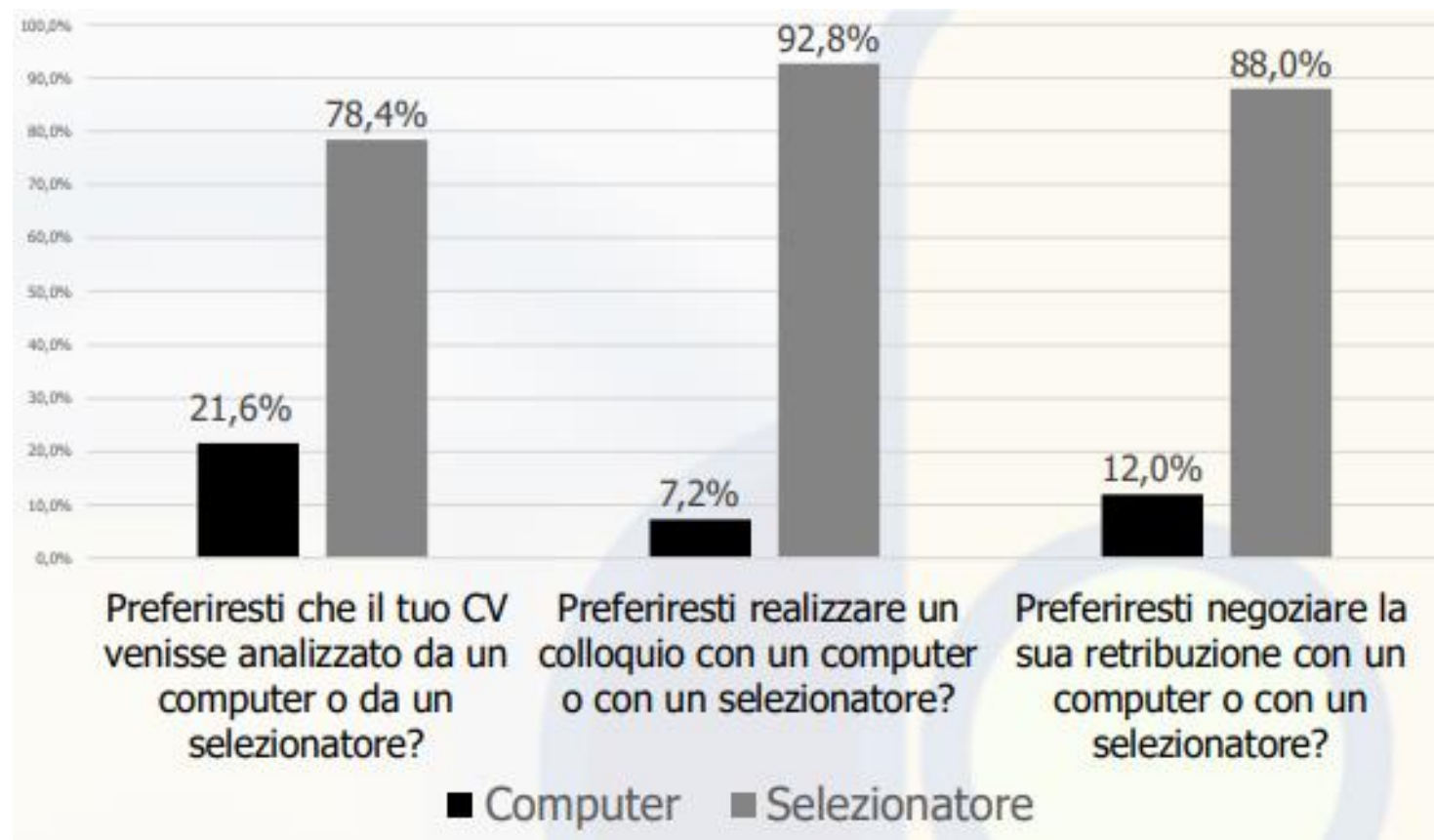
2020/2021

Questionario

125 rispondenti

Risultati:

Preferenza per una
interazione con
Selezionatore Umano



IN CONCLUSIONE

L'uso dell'IA nei processi di reclutamento e selezione può apportare un **valore aggiunto quantitativo** in una fase iniziale, ad es. semplificare e velocizzare attività quali lo **screening** dei CV e **l'analisi delle interazioni social** delle persone in cerca di lavoro.

Tuttavia, l'uso di questi strumenti, ad oggi, non può sostituire il valore qualitativo che una relazione umana può consentire, in particolare le conoscenze che le persone in cerca di lavoro possono trovare in un contatto bidirezionale, uno scambio a livello umano, chiedendo al selezionatore (durante il colloquio = “cum loqui”) alcune informazioni sulla **cultura**, i **valori** e il **clima organizzativi**.

Gestendo la comunicazione, ad esempio, il selezionatore è anche in grado di **comunicare l'immagine dell'azienda**, focalizzandosi sui valori (EVP) da evidenziare (= Employer Branding).

Esercitazione



Needfinding: valutare la **necessità di explainability** e stabilire il **livello di trust**

Background

Scegliere un prodotto/servizio digitale qualsiasi, con o senza AI

Strategie

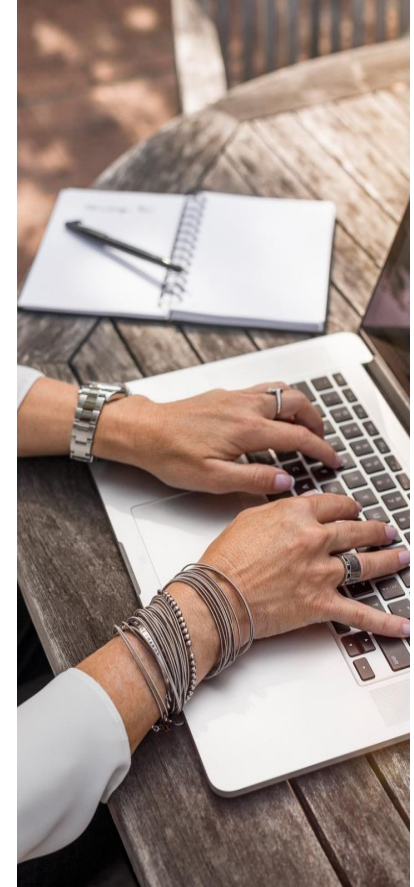
Guardate come questi prodotti o servizi cercano di aumentare il loro livello di trust e/o explainability

Needfinding

Immaginate un percorso di needfinding per chiedere agli utenti informazioni su trust ed explainability di questo prodotto che avete scelto o immaginato. (Potete poi usare questi spunti per il vostro assignment).

*Come fareste per chiedere alle persone se il prodotto è **trustable** ed **explainable**?*

Cosa migliorerebbe l'affidabilità del prodotto?



**Grazie per
l'attenzione**

Dr. ANDREA CAPUTO, PhD
Politecnico di Torino

andrea.caputo@unito.it

