QUESTIONI ETICHE EMERGENTI CORRELATE ALL'USO DELL'IA GENERATIVA

A cura della Commissione Etica AIP:
Antonietta Curci, Annalisa Guarini, Gian Mauro Manzoni, Silvio Carlo Ripamonti, Elena Trifiletti

Un sentito ringraziamento va a Fiorella Battaglia (Università del Salento, Fabio Paglieri (CNR), e Luca Pietrantoni (Università di Bologna) che, in qualità di esperti, hanno fornito un importante contributo alla rielaborazione di questo documento.

Indice

PREMESSA 1
1. QUESTIONE DELLA PRIVACY E DELLA SICUREZZA DEI DATI 2
Sfide. *2*
Linee guida. *2*
2. QUESTIONE DELLA RESPONSABILITÀ E DELL'OBBLIGO DI RENDERE CONTO 2
Sfide. *2*
Linee guida *2*
3. QUESTIONE DEL CONSENSO INFORMATO AL TRATTAMENTO DEI DATI 3
Sfide. *3*
Linee guida. *3*
4. QUESTIONE DEL CONSENSO ALLA PARTECIPAZIONE ALLA RICERCA 3
Sfide. *3*
Linee guida. *3*
5. QUESTIONE DELLA QUALITÀ DEL DATO RACCOLTO E DELLA OPACITÀ EPISTEMICA 4
Sfide. *4*
Linee guida. *4*
6. QUESTIONE DELL'EQUITÀ NELLA RICERCA 4
Sfide. *4*
Linee guida. *5*
7. QUESTIONE DELLA PROGETTAZIONE DI STRUMENTI DI RICERCA 5
Sfide. *5*
Linee guida. *5*
8. SFIDA DELLA PROPRIETÀ INTELLETTUALE 6
Rischi. *6*
Linee guida. *6*
9. QUESTIONE DELL'USO DEI MODELLLI LLM 6
Sfide. *6*
Linee guida. *6*
10. QUESTIONE DELLA TUTELA DELLE COMPETENZE DI RICERCA 7
Sfide. *7*
Linee guida. *7*

Premessa

La comunità scientifica è concorde nel ritenere che l'uso dell'Intelligenza Artificiale generativa (IA) nella ricerca possa ampliare la capacità di generare e diffondere la conoscenza scientifica. L'IA può ampliare la capacità di generare conoscenza, ma anche aumentare la necessità di normative e linee guida per garantire l'uso etico e responsabile della tecnologia nelle pratiche di ricerca.

Questo documento si propone di offrire una riflessione sulle questioni etiche correlate all'uso dell'IA (specificatamente, IA generativa) nella ricerca psicologica e di suggerire buone pratiche da adottare.

Per ciascuna delle questioni etiche elencate in questo documento, si evidenziano le sfide legate all'uso dell'IA nella ricerca psicologica e si suggeriscono buone pratiche da implementare nelle attività di ricerca. Il documento è articolato in 9 questioni, per ciascuno dei quali vengono introdotte le rispettive sfide e linee guida.

1. Questione della privacy e della sicurezza dei dati

Sfide.

Le principali sfide sulla privacy legate all'IA Generativa (IA) riguardano il rischio di violazioni dei dati e l'accesso non autorizzato alle informazioni personali. Date le enormi quantità di dati raccolti e processati, esiste il rischio che questi possano essere utilizzati con fini secondari non consoni né al codice etico né rispettosi del GDPR. È possibile, inoltre, che aziende private che offrono servizi di IA possano utilizzare i dati inseriti dagli utenti per addestrare direttamente i loro modelli con la conseguente possibilità di violazioni dei dati.

Linee guida.

- I ricercatori e le ricercatrici devono garantire la sicurezza dei dati adottando misure efficaci per proteggere le informazioni sensibili, rispettando la normativa vigente sulla privacy (Regolamento UE 2016/679, General Data Protection Regulation, GDPR). In un contesto in cui il trattamento e la conservazione sicura dei dati emergono come questioni primarie nell'uso dell'IA nella ricerca, è necessaria una gestione attenta e responsabile, che assicuri l'integrità e la riservatezza delle informazioni.

2. Questione della responsabilità e dell'obbligo di rendere conto

Sfide.

L'integrazione IA nella ricerca solleva questioni cruciali riguardanti la responsabilità e l'obbligo di rendere conto. Poiché l'IA influenza le decisioni e i risultati della ricerca attraverso algoritmi complessi, è essenziale comprendere chi sia moralmente e/o legalmente responsabile per le decisioni prese dall'IA e come rendere conto dei processi decisionali automatizzati che entrano in gioco nelle diverse fasi della ricerca
È essenziale definire in modo preciso e trasparente le responsabilità di tutti gli attori coinvolti, inclusi i ricercatori, gli sviluppatori degli strumenti di IA e gli utenti, al fine di garantire un utilizzo etico e responsabile dell'IA.

Linee guida.

- Nei progetti e nelle ricerche deve essere chiaramente identificabile il ricercatore o il gruppo di lavoro che sono gli unici a cui ci può attribuire la responsabilità delle decisioni prese dall'IA
- Devono essere identificabili anche gli sviluppatori degli strumenti di IA rendendo palese quale IA è stata utilizzata

3. Questione del consenso informato al trattamento dei dati

Sfide.

Come indicato dal Codice Etico AIP, il consenso al trattamento dei dati è distinto dal consenso alla partecipazione alla ricerca.
La sfida legata al consenso al trattamento dati in un processo di ricerca che utilizzi l'IA è legata alla esplicitazione dei rischi aggiuntivi legati all'uso delle informazioni personali raccolte. Deve essere garantito che tutte le informazioni personali raccolte durante la ricerca siano trattate con la massima riservatezza e che i dati saranno protetti contro accessi non autorizzati. Il rischio è che i partecipanti non comprendano al meglio i temi legati all'suo dell'IA in relazione al trattamento dati.

Linee guida.

- Quando si utilizzano algoritmi di IA, i ricercatori/le ricercatrici devono chiarire in modo comprensibile ai partecipanti le modalità di trattamento dei dati. I ricercatori/le ricercatrici devono comunicare chiaramente i dettagli sull'uso dell'IA, assicurando che i partecipanti comprendano appieno le implicazioni. È fondamentale discutere i rischi potenziali, che potrebbero includere l'uso improprio dei dati, la violazione della privacy o altri rischi etici associati all'uso dell'IA.
- È fondamentale proteggere i partecipanti e i loro dati, rispettando le normative sulla privacy e i principi etici. Garantire la riservatezza e la sicurezza dei dati personali è essenziale per un processo di ricerca eticamente corretto e scientificamente valido
- I ricercatori e le ricercatrici che usano l'IA sono tenuti a fornire, nei limiti del possibile, informazioni chiare e trasparenti ai partecipanti all'interno del modulo informativo del consenso al trattamento dati.

4. Questione del consenso alla partecipazione alla ricerca

Sfide.

Il consenso informato deve essere richiesto e ottenuto da tutti i partecipanti a qualsiasi progetto di ricerca, come stabilito dal Codice Etico AIP, che all'articolo 1 indica il tipo di informazioni da fornire ai partecipanti alla ricerca.
Nei casi in cui si utilizzi l'IA è necessario che il modulo informativo del consenso informato fornisca una spiegazione approfondita di vari aspetti etici e tecnici legati alla ricerca. Questo processo si basa su una comunicazione chiara e trasparente con i partecipanti riguardo alla natura specifica dell'uso dell'IA nella ricerca.
La sfida riguardante il consenso informato che utilizzi l'IA è relativa alla possibilità che i partecipanti non siano pienamente informati sulle procedure, sui rischi e sui benefici potenziali dell'uso dell'IA legati alla loro partecipazione, o non comprendano pienamente gli obiettivi dell'uso dell'IA e le metodologie impiegate. Questa possibilità è legata all'opacità dei sistemi di IA, i cui processi di elaborazione delle informazioni non sono del tutto chiare agli utilizzatori e nemmeno agli esperti di IA.

Linee guida.

- I partecipanti devono essere informati dettagliatamente sugli obiettivi della ricerca e su come l'IA verrà utilizzata per raggiungere questi obiettivi. Questo include una spiegazione chiara delle metodologie dell'IA impiegate, come l'analisi dei dati, il machine learning o l'elaborazione del linguaggio naturale, e del motivo per cui queste tecnologie sono necessarie per la ricerca. I partecipanti devono comprendere non solo il contesto generale

della ricerca, ma anche il ruolo specifico che l'IA gioca all'interno di essa. Esempi di tali informazioni riguardano il tipo e il nome del sistema di IA utilizzato, le finalità per le quali viene utilizzato, come verranno usati i dati o le risposte fornite dai partecipanti, quali sono i possibili rischi e i possibili vantaggi per i partecipanti, esperienza del ricercatore/ricercatrice con il sistema di IA utilizzato, compiti e responsabilità dell'IA e dei ricercatori. Tali informazioni devono essere comunicate in un linguaggio comprensibile ai non esperti.

- Devono essere esplicitati i possibili vantaggi e rischi che i partecipanti, e più in generale gli stakeholder, potrebbero sperimentare grazie alla ricerca che utilizza l'IA.

5. Questione della qualità del dato raccolto e della opacità epistemica

Sfide.

L'uso di IA spesso non mette in grado la ricercatrice/ricercatore di comprendere come gli algoritmi sintetizzano il materiale su cui lavorano per generare nuova informazione. Molte tecnologie di IA rendono difficile comprendere appieno come lavorano sui dati presenti in rete. Questo fenomeno è noto con il termine di "opacità epistemica". Senza un approccio strutturato e metodico non saremo in grado di valutare le informazioni fornite dalle sequenze di istruzioni del programma che potrebbero essere fonte di malintesi. Gli algoritmi di IA spesso agiscono come "scatole nere", rendendo complesso capire come si giunge a specifiche conclusioni. Il rischio è che l'uso dell'IA non permetta la piena condivisione con i partecipanti delle modalità di interpretazione dei dati raccolti e delle conclusioni alle quali si arriva.

Un tema importante nell'uso dell'IA è che il ricercatore, facendo eccessivo affidamento sulla mole di dati processati su cui basa le sue conclusioni, non metta in evidenza il tema dell'umiltà epistemica, che deve riguardare ogni processo conoscitivo. Ogni processo di ricerca deve evidenziare il riconoscimento e l'accettazione dei limiti della propria conoscenza. Il ricercatore deve essere consapevole che le proprie credenze e conoscenze possono essere incomplete, errate o influenzate da pregiudizi. L'uso di IA può far nascere l'illusione di fare affidamento su una mole di dati e processi di sintesi degli stessi che possono portare a conclusioni certe e affidabili.

Linee guida.

- È importante che il ricercatore, consapevole della propria "umiltà epistemica", evidenzi i limiti e i possibili sviluppi delle conclusioni tratte anche con l'uso di IA. La validazione dei risultati è cruciale e deve rimanere a carico del ricercatore, che deve mantenere il controllo dell'analisi dei dati e attenuare bias algoritmici. È fondamentale adottare tutte le misure per evitare errori o distorsioni nei risultati e nelle loro interpretazioni. Seguendo le linee guida della Commissione Europea (ALTAI) sull'affidabilità dei sistemi AI, i ricercatori devono garantire che i modelli di IA siano trasparenti e spiegabili, documentando e giustificando le decisioni degli algoritmi, specialmente nell'interpretazione dei dati psicologici.

6. Questione dell'equità nella ricerca

Sfide.

L'utilizzo degli algoritmi di intelligenza artificiale può accentuare le disparità tra i partecipanti e non tener conto delle specifiche categorie di individui, poiché tali algoritmi spesso si basano su

dati storici che possono contenere bias preesistenti. Queste distorsioni possono perpetuare ingiustizie e diseguaglianze, influenzando negativamente gruppi già svantaggiati.
Gli algoritmi di intelligenza artificiale possono altresì essere soggetti a bias sui dati derivanti dall'addestramento o dall'analisi dei risultati e portare a conclusioni di ricerca non rappresentative e/o discriminatorie.
Il coinvolgimento attivo dei ricercatori/le ricercatrici è essenziale per identificare e mitigare eventuali distorsioni incorporate nell'IA. Solo attraverso un'integrazione sinergica tra intelligenza artificiale e competenze umane può essere garantito un risultato di ricerca valido, etico e socialmente responsabile.

Linee guida.

- Le ricercatrici/ricercatore devono garantire equità nella selezione dei partecipanti garantendo a tutti la possibilità di essere rappresentati nei processi di ricerca.
- È fondamentale valutare gli effetti sociali delle tecnologie impiegate nella ricerca, affinché non contribuiscano all'emarginazione di gruppi e minoranze nei vari contesti sociali e alla diffusione e rafforzamento di stereotipi, pregiudizi e discriminazioni.
- Le ricercatrici/ricercatore devono monitorare i bias di campionamento eventualmente emergenti nell'uso dell'IA assicurandosi che le soluzioni proposte siano eque e rappresentative di tutte le categorie di individui.
- Le ricercatrici/ricercatore devono essere consapevoli dei bias generati dall'IA e lavorare attivamente per identificarli e mitigarli in tutte le fasi della ricerca, dall'ideazione del progetto alla raccolta dei dati, fino all'interpretazione dei risultati e alla pubblicazione delle conclusioni. Questa attenzione richiede un monitoraggio continuo e una valutazione critica dei metodi utilizzati per assicurare che le distorsioni non influenzino negativamente gli esiti della ricerca.

7. Questione della progettazione di strumenti di ricerca

Sfide.

L'intelligenza artificiale può svolgere un ruolo significativo nella progettazione e nell'implementazione di strumenti di ricerca, come i questionari per la raccolta di dati, la creazione di vignette, prove sperimentali e video. Può, però, fornire strumenti di ricerca che non recepiscano le conoscenze pragmatiche e generate dall'esperienza sul campo che costituiscono le comunità di pratiche dei ricercatori. Il rischio nell'uso dell'IA è che l'esperienza dei ricercatori non sia la base sulla quale vengono progettati e proposti alla comunità scientifica gli strumenti di ricerca, che di fatto possono essere esito soltanto di algoritmi. Di qui la possibilità perdita di "unicità sostanziale" degli strumenti di ricerca.

Linee guida.

- La responsabilità della progettazione accurata degli strumenti di ricerca deve rimanere saldamente nelle mani delle ricercatrici/ricercatori. L'IA può aiutare nella creazione di questionari o test più efficienti attraverso l'analisi dei dati storici e la generazione di domande pertinenti. Tuttavia, la comprensione del contesto della ricerca, la definizione degli obiettivi dello studio e la garanzia della correttezza procedurale sono competenze umane indispensabili, che devono guidare tutto il processo di progettazione della ricerca.
- Ogni fase del processo di ricerca deve essere caratterizzata da Human Oversight. Il contributo distintivo e creativo del ricercatore deve essere evidente dalla definizione domanda di ricerca, alla raccolta e analisi dei dati fino all' interpretazione dei dati e alla disseminazione dei risultati.

8. Sfida della proprietà intellettuale

Rischi.

L'IA generativa, in particolare i Large Language Models (LLM) possono generare del tutto casualmente informazioni non veritiere e citazioni di fonti inesistenti. Questo crea un rischio per la validità della conoscenza scientifica diffusa attraverso le pubblicazioni realizzate, anche solo in parte, attraverso i LLM.
I sistemi AI possono essere anche vulnerabili a manipolazioni e attacchi, chiamati "attacchi adversarial", che rischiano di rendere discutibile l'affidabilità dei dati.

Linee guida.

- Ogni testo prodotto dall'intelligenza artificiale deve essere attentamente verificato per assicurare la conformità alle leggi sul copyright e per garantire che sia stato ottenuti i necessari permessi o licenze.
- Quando si fa uso di informazioni specifiche, citazioni o idee tratte da un'opera soggetta a copyright, è fondamentale attribuire correttamente la fonte e l'autore. Nei resoconti di ricerca è essenziale poter risalire correttamente alle fonti e attribuire l'autorialità in modo corretto a tutte le informazioni presenti in un report. Inoltre, è essenziale ricordare che i ricercatori e le ricercatrici sono responsabili della validità e correttezza di tutte le informazioni contenute in una pubblicazione scientifica. Tale responsabilità non può essere demandata all'IA.
- Il ricercatore deve essere consapevole delle possibili manipolazioni dei dati che possono avvenire nel WEB. Il ricercatore deve, quindi, porre in campo tutte le possibili cautele affinché non vengano tratte conclusioni indebite dall'analisi di dati distorti.

9. Questione dell'uso dei modellli LLM

Sfide.

Il tema dei modelli LLM (LLM - Large Language Models) pone in primo piano l'esplicitazione da parte del ricercatore del loro utilizzo nel processo di scrittura. Il loro utilizzo trasparente permette al lettore di sapere che il contenuto è, totalmente o parzialmente, prodotto da una macchina. Dichiarare l'uso dei LLM aiuta a mantenere la credibilità del ricercatore e può essere necessario per conformarsi a linee guida etiche emergenti.
L'uso palese degli LLM deve implicare maggiore attenzione alla qualità e alla revisione del contenuto, garantendo accuratezza e pertinenza dei contenuti da parte del ricercatore. Esplicitare l'uso dei LLM nella scrittura è un passo essenziale verso la trasparenza e l'assunzione di responsabilità da parte del ricercatore. Tale atteggiamento può contribuire ad incrementare la fiducia dei lettori e contribuendo ad un uso più etico di questi strumenti.
L'utilizzo non esplicito dei sistemi LLM introduce una frattura profonda tra i destinatari dei risultati della ricerca e può determinare una compromissione del rapporto di fiducia tra ricercatori, editori e stakeholder.

Linee guida.

- È necessario esplicitare sempre l'uso degli LLM nella scrittura di un paper, nei documenti che rendicontano una ricerca e, in generale, nei processi di divulgazione scientifica.
- È importante specificare quale LLM è stato utilizzato.

- È fondamentale revisionare e modificare i risultati di un prodotto LLM, assumendosi la piena responsabilità dei contenuti pubblicati.

10. Questione della tutela delle competenze di ricerca

Sfide.

La crescente dipendenza dall'IA generativa potrebbe influenzare negativamente lo sviluppo delle competenze tradizionali dei ricercatori durante la fase di formazione. Mentre l'IA offre strumenti avanzati per l'analisi dei dati e la generazione di nuove conoscenze, c'è il rischio che i ricercatori non sviluppino pienamente competenze cruciali come la capacità critica di analizzare manualmente i dati e di formulare ipotesi indipendenti. Questo può limitare la loro capacità di affrontare sfide complesse e impreviste nel processo di ricerca, compromettendo la profondità del loro expertise nei rispettivi campi di studio. Al contempo, l'uso maturo ed eticamente adeguato dei sistemi di IA, così come la capacità di valutare gli usi fatti da altri (ad es., pazienti in un setting clinico o persone in formazione) richiede nuove competenze professionali, con un processo di costante aggiornamento della formazione professionale in psicologia.
È fondamentale promuovere le competenze di ricerca nella formazione dei nuovi ricercatori. A tal fine è importante concentrarsi sulla capacità di contestualizzare i dati, di valorizzare le culture organizzative locali, di comprendere le conoscenze tacite che caratterizzano una comunità scientifica. Le limitazioni dell'IA derivano dalla sua incapacità di comprendere appieno i contesti culturali di riferimento e dalla difficoltà di percepire le sfumature culturali, emotive o contestuali che possono invece essere sviluppate ed allenate nei giovani ricercatori.

Linee guida.

- È essenziale che l'utilizzo dell'IA garantisca la tutela ed il miglioramento continuo delle competenze e delle capacità di progettazione, ricerca ed analisi dei dati dei ricercatori, specialmente di coloro che si trovano ancora in fase di apprendimento. L'uso delle tecnologie emergenti non deve compromettere lo sviluppo delle competenze umane necessarie per proporre ricerche di qualità.
- La formazione all'uso dell'IA in psicologia deve diventare parte integrante del percorso educativo dei ricercatori, come anche dell'aggiornamento professionale degli psicologi in attività, prestando particolare attenzione all'attualità degli apprendimenti, data la rapidità con cui le tecnologie si evolvono in questo settore.
- I ricercatori e le ricercatrici con maggiore esperienza devono assumersi la responsabilità di monitorare la formazione dei giovani ricercatori e degli studenti di psicologia garantendo che l'IA non comprometta lo sviluppo delle loro competenze di ricerca.