

Il ruolo dell'IA nella costruzione di una società rispettosa dei diritti fondamentali. Il caso di studio del filtro Bold Glamour di TikTok

*Fabiana Ciccarella, Lucrezia Fortuna, Elisabetta Lambiase, Mattia Mogetti**

AI'S ROLE IN BUILDING A SOCIETY THAT RESPECTS FUNDAMENTAL RIGHTS. THE CASE STUDY OF THE TIKTOK BOLD GLAMOUR FILTER

ABSTRACT: This article explores the role of Artificial Intelligence (AI) in building a society that respects fundamental rights, using TikTok's "Bold Glamour" filter as a case study. It analyses the ethical and social implications of AI-driven beauty filters, highlighting how they impact self-perception and reinforce unrealistic beauty standards. The research emphasises the need for a balanced approach in regulating AI, protecting consumer rights, and promoting ethical use. It discusses the European Commission guidelines and the importance of transparency, diversity, and non-discrimination in AI systems. The study concludes by proposing legislative, technological, and social interventions to mitigate potential harms

KEYWORDS: Artificial Intelligence; Ethical implications; Beauty filters; User vulnerability; Regulatory framework.

ABSTRACT: L'articolo esplora il ruolo dell'Intelligenza Artificiale (IA) nella costruzione di una società che rispetti i diritti fondamentali, utilizzando il filtro "Bold Glamour" di TikTok come caso di studio. L'articolo analizza le implicazioni etiche e sociali dei filtri di bellezza basati sull'IA, evidenziando il loro impatto sulla percezione di sé e come rafforzino standard di bellezza irrealistici. La ricerca sottolinea la necessità di un approccio equilibrato nella regolamentazione dell'IA, che protegga i diritti dei consumatori e promuova un uso etico. Esamina le linee guida della Commissione europea e l'importanza della trasparenza, della diversità e della non discriminazione nei sistemi di IA. Lo studio si conclude proponendo interventi legislativi, tecnologici e sociali per mitigare i danni potenziali.

PAROLE CHIAVE: Intelligenza artificiale; Implicazioni etiche; Filtri di bellezza; Vulnerabilità dell'utilizzatore; Quadro normativo.

* Dottorand* di ricerca in Gender Studies, Università di Bari. Mail: fabiana.ciccarella@uniba.it, lucrezia.fortuna@uniba.it, elisabetta.lambiase@uniba.it, mattia.mogetti@uniba.it. L. Fortuna ha curato il paragrafo I; M. Mogetti ha curato il paragrafo II; E. Lambiase ha curato il paragrafo III; F. Ciccarella ha curato il paragrafo IV. Contributo sottoposto a doppio referaggio anonimo.



SOMMARIO: 1. Il filtro *Bold Glamour*: un primo inquadramento alla ricerca di una socialità etica – 1.1. Etica e diritto per la costruzione di un'intelligenza artificiale umano centrica - 2. Relazioni scisse e iperrealità di genere - 3. Implicazioni etiche e tutela dei diritti fondamentali nell'utilizzo del filtro *Bold Glamour* - 4. Alcuni cenni conclusivi

1. Il filtro *Bold Glamour*: un primo inquadramento alla ricerca di una socialità etica

La rivoluzione digitale ha profondamente trasformato la vita umana: la parola «*onlife*», coniata dal filosofo Floridi, coglie pienamente un tratto essenziale del nostro vivere, l'ubiquità: per la prima volta nella storia abbiamo la possibilità di essere e agire in due dimensioni contemporaneamente, *online* e *offline*.¹ Su questo *mondo nuovo*, l'Intelligenza artificiale ha – e sempre più avrà – un impatto enorme, con ciò sollevando una moltitudine di questioni dal punto di vista etico, anche con riferimento alle grandi piattaforme digitali come *Meta* e *TikTok*: i *social media* e il crescente spazio che hanno guadagnato nella vita quotidiana plasmano le nostre interazioni sociali e la nostra percezione di noi stessi e degli altri. All'interno di questa nuova esperienza di sé e di ciò che è altro da sé, i filtri rappresentano un fenomeno sempre più diffuso e pervasivo. Si tratta di *maschere* sviluppate in realtà aumentata² che si applicano sul proprio volto o su quello di altre persone ritratte nella foto o nel video. Le piattaforme forniscono una vasta gamma di filtri, tra i quali quelli *bellezza*, che modificano l'aspetto fisico della persona *migliorandolo*. Questa capacità di manipolare l'immagine, non nuova certamente, pone oggi questioni del tutto inedite in ragione della facilità di accesso, della velocità di diffusione e, non da ultimo, del sempre maggiore *realismo* dei filtri: ne è un esempio *Bold Glamour*, oggetto del presente lavoro, un filtro iperrealistico che ha suscitato accese polemiche, dettate anche dal massivo utilizzo fattone dagli utenti. Esso tende ad assottigliare, definire e scolpire i tratti del volto, enfatizzandoli attraverso una pelle più levigata e un trucco professionale ma non marcato e un'illuminazione sofisticata: effetti che si traducono nella realizzazione di un'immagine estremamente verosimile, anche perché restano *attaccati* al volto, non scomparendo neppure se la persona si muove o frappone un oggetto tra sé e lo schermo. L'estremo realismo del filtro, e quindi la difficoltà ad accorgersi del ricorso allo stesso, non sono il solo dato preoccupante: le principali critiche muovono dal fatto che questo tende a schiarire il colore della pelle, ad adattarsi ai lineamenti al punto da rimuovere il trucco laddove riconosce tratti più mascholini e, più in generale, a promuovere un modello estetico occidentale binario, escludente e inarrivabile³.

Sebbene l'uso dei filtri possa apparire innocuo, è importante non sottovalutarne i potenziali effetti negativi, *in primis* sulla percezione di sé e sulla salute mentale degli utenti, profondamente frustrati da canoni estetici resi desiderabili dalla continua esposizione, immediatamente accessibili *online* eppure irraggiungibili *offline*. Ciò appare ancor più vero laddove si consideri che i sistemi di IA sono ad oggi largamente addestrati sulla base di dati che riflettono pregiudizi e disuguaglianze sociali, e dunque in grado di perpetuare stereotipi e discriminazioni accrescendo il rischio di esclusione di tutte quelle

¹ Si v. L. FLORIDI, *Etica dell'intelligenza artificiale, Sviluppi opportunità, sfide*, Milano, 2022.

² Per una individuazione più puntuale della definizione si rinvia al II paragrafo.

³ M. MANFREDI, *ClioMakeUp e le altre imprenditrici beauty contro l'uso dei filtri Instagram*, in *La Repubblica*, disponibile online al https://www.repubblica.it/moda-e-beauty/2021/04/21/news/cliomakeup_contro_filtri_instagram_star_senza_filtri_social-342223446/ (ultima consultazione 10/07/2024).

persone che divergono dal modello *fiutamente* universale proposto. Si pone con evidenza un problema di equità e giustizia sociale che invece potrebbe trovare proprio nella dimensione digitale un'opportunità di riaffermazione, anche alla luce della natura embrionale della normazione in materia.⁴ In tal senso rappresentano uno snodo fondamentale la consapevolezza circa il ruolo trasformativo dei Sistemi di IA e l'autonomia che residua in seno alla persona nei cambiamenti così generati (paragrafo 2).

Sebbene di tutto questo si abbia sempre maggiore contezza – ne sono prova le linee guida implementate dalle piattaforme⁵ – non sembra ancora adeguatamente considerato l'impatto della manipolazione *di massa* delle immagini e le complesse e multifacetiche questioni etiche che esso solleva. Come più diffusamente si dirà in seguito, le prospettive da cui è possibile guardarvi sono innumerevoli e così le soluzioni approntabili, non esauribili in una disciplina incentrata sul riconoscimento in capo alle piattaforme di doveri, e dunque sulla sola responsabilità di queste ultime. L'uso diffuso dei filtri sui *social*, infatti, porta con sé una serie di problematiche etiche e sociali che coinvolgono gli stessi utenti: diventa essenziale adottare un approccio olistico nello sviluppo e nell'uso di queste tecnologie, assicurando che contribuiscano a ridurre, piuttosto che aumentare, le disuguaglianze sociali, a partire dalla promozione di un ambiente *online* sano, inclusivo e rispettoso. Un primo piano di analisi è quello della eticità e affidabilità del filtro e dei sistemi di IA in genere: come si illustrerà nel paragrafo terzo, esistono una serie di criteri alla luce dei quali valutare se un sistema di IA possa dirsi rispondente a quei criteri, così permettendo di elaborare le risposte più opportune.

1.1 Etica e diritto per la costruzione di un'intelligenza artificiale umano centrica

Sebbene possa dirsi che l'etica venga prima della legge – e durante e dopo di questa – e che dunque anche dal punto di vista logico il discorso etico precede quello più strettamente normativo, illustrare – sia pure per brevi cenni – alcune delle soluzioni adottate e gli obiettivi cui esse mirano intervenendo sulla materia, appare utile per coglierne portata e limiti al fine di orientare lo stesso dibattito etico che ne è alla base. Come si vedrà più compiutamente nel prosieguo, gli interventi sembrano limitarsi a *tamponare* gli usi distorti o discriminatori che possono essere fatti dell'IA, senza cogliere il quadro d'insieme: definirne i tratti consente al piano etico di svilupparsi anche laddove la legge non arriva. Non è sufficiente che gli algoritmi seguano le regole, è necessario che la persona e il suo benessere

⁴ A tal proposito si rinvia a S. FERILLI, E. GIRARDI, C. MUSTO, M. PAOLINI, P. POCCIANI, S. POCHETTINO, G. SEMAFORO, *L'Intelligenza Artificiale per lo Sviluppo Sostenibile*, Roma, 2021 disponibile al link: <https://www.cnr.it/sites/default/files/public/media/attivita/editoria/VOLUME%20FULL%2014%20digital%20LIGHT.pdf>.

⁵A esempio: <https://www.facebook.com/help/instagram/477434105621119> (ultima consultazione 10/07/2024). E in specie le regole e i programmi per la creazione di una *comunità solidale*, in cui si fa riferimento esplicito alla lotta contro il bullismo e la pressione per la perfezione, alla promozione dell'accettazione del corpo; colpisce il riferimento alla «promozione dell'empatia per uno spazio Internet migliore e più inclusivo». Ancora, come si illustrerà meglio in seguito, le linee guida di TikTok: «TikTok ha otto pilastri della *community* che scaturiscono dal nostro impegno al rispetto di sicurezza e diritti umani. A tali principi si ispirano le nostre attività giornalieri e il nostro approccio alle decisioni difficili, inerenti all'applicazione di regole. I principi sono incentrati su questi temi: bilanciare l'espressione creativa e la prevenzione dei danni; sostenere la dignità umana; garantire che le nostre azioni siano eque», <https://www.tiktok.com/community-guidelines/it/> (ultima consultazione 10/07/2024).



restino il fine, mentre ad oggi questi sembrano piuttosto relegati a mezzi.⁶ La sfida diventa non solo l'elaborazione di una normativa rispettosa dei diritti della persona, si pensi alle disposizioni sull'utilizzo dei dati e sulla *privacy*, ma anche e soprattutto di una legislazione capace di orientare l'IA verso la promozione attiva di quei diritti.⁷ Si deve, in sostanza, rivendicare la centralità di un approccio all'intelligenza artificiale che sappia rispondere a finalità ulteriori rispetto a quelle meramente economiche e scientifiche, e segnatamente sicurezza, trasparenza e garanzia dei diritti e delle minoranze, in una prospettiva più ampia di tutela della persona che passa anche attraverso la protezione della sua salute mentale per così dire *umano centrica*. Emblematica in tal senso l'approvazione della legge europea sull'intelligenza artificiale (*AI Act*),⁸ volta a «proteggere i diritti fondamentali, la democrazia, lo Stato di diritto e la sostenibilità ambientale dai sistemi di IA ad alto rischio, promuovendo nel contempo l'innovazione e assicurando all'Europa un ruolo guida nel settore». La normativa, non ancora in vigore, si candida a compiere un passo storico nella disciplina dell'IA: è la prima ad affrontare specificamente il tema.⁹ Elemento rilevante è, tra gli altri, la previsione dell'obbligo di etichettare espressamente e chiaramente ogni immagine e contenuto audio o video artificiale o manipolato (*deep fake*) come tale, coerentemente con gli obiettivi di trasparenza che le norme pongono. Il fine ultimo della legge è tenere insieme intelligenza artificiale e valori fondanti dell'Unione e degli stati membri, rispondendo ad una esplicita richiesta della cittadinanza europea: alla base vi sono, infatti, anche le proposte elaborate dalla Conferenza sul futuro dell'Europa (COFE)¹⁰. Non a caso tra i suoi obiettivi c'è

⁶ In questo senso, A. D'Aloia a proposito della direzione intrapresa dai primi atti paranormativi in tema di IA: «Si lavora essenzialmente in chiave di adattamento interpretativo dei principi consolidati del costituzionalismo alle applicazioni di AI: l'obiettivo è quello di configurare un'AI «*human-centric*» and «*trustworthy*», appunto perché connotata da un «*ethical purpose*», e da un obbligo di conformità ai valori fondamentali della convivenza civile (rispetto della dignità umana e dei diritti, eguaglianza e non discriminazione, non interferenza rispetto ai processi democratici, sicurezza, rispetto per la *privacy*), e al benessere delle persone», in *Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale*, in *BioLaw Journal – Rivista di BioDiritto*, 1, 2019, 6.

⁷ Emblematico in questo senso anche G. MOBILIO, *La co-regolazione delle nuove tecnologie, tra rischi e tutela dei diritti fondamentali*, in *Osservatorio sulle fonti*, 1, 2024, 260, disponibile in: <http://www.osservatoriosullefonti.it>.

⁸ Si tratta del Regolamento UE 2024/1689 del Parlamento Europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale. Si rinvia al paragrafo 3 e <https://www.europarl.europa.eu/news/it/press-room/20240308IPR19015/il-parlamento-europeo-approva-la-legge-sull-intelligenza-artificiale> (ultima consultazione 10/07/2024).

⁹ Sebbene già molti paesi siano intervenuti in materia, come sottolineato da T. E. Frosini: «Vale la pena ricordare come più di 30 Paesi nel mondo – fin dal 2017 il Canada, il Giappone, la Cina e la Finlandia – hanno previsto e poi adottato una strategia nazionale per lo sviluppo dei sistemi di IA: a conferma di come la maggior parte delle economie sviluppate attribuisca alla IA un significato e un valore davvero rivoluzionario, che incide significativamente sulla crescita economica, sociale, occupazionale e culturale del Paese», T. E. FROSINI, *L'orizzonte giuridico dell'intelligenza artificiale*, in *BioLaw Journal – Rivista di BioDiritto*, 1, 2022, 158.

¹⁰ La Conferenza sul futuro dell'Europa ha rappresentato un importante rafforzamento del processo democratico europeo: hanno potuto prendere parola e discutere proponendo al termine del percorso apposite raccomandazioni. La conferenza si è svolta dall'aprile 2021 al maggio 2022, quando è stata presentata la relazione finale contenente 49 proposte e 326 misure rivolte alle istituzioni. <https://www.consilium.europa.eu/it/policies/conference-on-the-future-of-europe/> (ultima consultazione 10/07/2024). Si segnalano, riguardo l'AI, la raccomandazione sul rafforzamento della competitività dell'UE nei settori strategici e la promozione dell'innovazione digitale e su una società sicura e affidabile, tra cui la lotta alla disinformazione e la garanzia di un controllo umano di ultima istanza e sull'uso affidabile e responsabile dell'IA, stabilendo salvaguardie e garantendo la trasparenza, e alla proposta sull'utilizzo dell'IA e degli strumenti digitali per migliorare l'accesso dei cittadini

anche la lotta alla discriminazione. Mentre nell'AI Act appare più pacifica la soluzione dell'equazione tecnologia e diritto, la tensione tra dimensione economica e tutela della persona sembra ancora irrisolta nelle prime risposte date ad altre questioni poste dall'utilizzo dei filtri, a partire dalla specifica regolamentazione di cui si sono dotate le piattaforme su cui questi ultimi vengono impiegati, come Meta¹¹ e Tiktok. Un primo piano di intervento ha riguardato il loro impiego in ambito commerciale da parte di *influencer*, *brand* e aziende: si parla della promozione e pubblicizzazione di prodotti e/o servizi. Qui l'alterazione prodotta dall'impiego dei filtri rileva per la potenziale distorsione della concorrenza mediante condotte ingannevoli, in quanto capace di promuovere presso i consumatori prodotti la cui efficacia è potenziata dall'uso dei filtri stessi e, dunque, di trarli in inganno. Ciò ha portato nel Regno Unito all'intervento dell'*Advertising Standard Authority* (ASA), autorità di autodisciplina pubblicitaria: l'ente ha affermato che «l'uso dei filtri nella comunicazione commerciale inerente a determinati prodotti potrebbe aumentare l'effetto del prodotto cosmetico risultando così fuorviante per i consumatori, configurandosi in tal modo una condotta di pubblicità ingannevole» e per questo *influencer* e *advertiser*, che promuovano prodotti cosmetici, sono tenuti a segnalare l'utilizzo di filtri su foto e video direttamente inerenti al prodotto o servizio pubblicizzato e a non applicare a tali contenuti quei filtri laddove questi possano esagerare l'effetto *naturale* del prodotto¹². La Norvegia è invece intervenuta sulla propria legislazione¹³ rendendo illegale la pubblicazione di foto e video ritoccati. La sanzione può persino arrivare alla pena carceraria: ciascun contenuto pubblicitario che contenga alterazioni alla forma o al peso del corpo ovvero al colore della pelle dovrà obbligatoriamente essere segnalato attraverso apposite etichette predisposte dallo stesso legislatore¹⁴.

Alla luce di quanto sinora detto, appare chiaro come gli interventi più puntuali sembrino concentrarsi sulla portata del fenomeno dei filtri in termini di concorrenza e influenza sulle scelte dei consumatori, oltretutto sul solo momento dell'uso: è necessario, tuttavia, riflettere altresì sulla dimensione sociale dell'utilizzo dei filtri e sulla loro fase di elaborazione. Pur se importante, non appare decisivo rendere

alle informazioni, comprese le persone con disabilità.

¹¹ Cfr. nota numero 6.

¹² Si v. per i dettagli: [The \(mis\)use of social media beauty filters when advertising cosmetic products - ASA | CAP](#). Una simile presa di posizione non è stata assunta in Italia da parte dell'Istituto di Autodisciplina Pubblicitaria (IAP), d'altra parte resta applicabile la normativa in materia di pubblicità ingannevole, tra le pratiche commerciali scorrette sanzionate ai sensi dei d.lgs. 145 e 146 del 2007. Per approfondire si rinvia a: <https://www.linkedin.com/pulse/le-foto-ritoccate-su-instagram-sono-legali-norvegia-regno-alba?originalSubdomain=it>, [Filtri Instagram e ingannevolezza: l'autodisciplina inglese mette un freno all'utilizzo dei filtri su Instagram per la promozione di prodotti cosmetici - DGRS - Studio Legale Milano](#) (ultima consultazione 10/07/2024). Si v. però le linee guida emanate dall'Autorità per le Garanzie nelle Comunicazioni (AGCOM) volte ad individuare le disposizioni del Testo Unico sui Servizi di Media Audiovisivi ("TUSMA") che gli *influencer* saranno tenuti a rispettare, «in vista della crescente diffusione delle attività degli influencer e della necessità di garantire una pubblicità trasparente» di recente pubblicazione: [Nuove Linee Guida AGCOM per Influencer Marketing: Trasparenza e Regolamentazione \(dejalex.com\)](#).

¹³ Si tratta di emendamenti al *Marketing Act* del 2009.

¹⁴ Per leggere la normativa in commento si rinvia a: <https://www.stortinget.no/no/Saker-og-publikasjoner/Saker/Sak/?p=84478> (ultima consultazione 10/07/2024) di cui si riporta il passaggio che esplicita le finalità con cui il legislatore ha adottato gli emendamenti: «gli emendamenti alla legge sul marketing mirano a ridurre la pressione del corpo nella società a causa delle persone idealizzate nella pubblicità. Tra le altre cose, viene introdotto l'obbligo di etichettare la pubblicità ritoccata o altrimenti manipolata quando ciò significa che il corpo della persona nelle pubblicità si discosta dalla realtà in termini di forma del corpo, dimensioni e pelle».



conoscibile l'uso del filtro, in quanto se ne trascurano tanto gli effetti quanto i presupposti. Sul primo piano rischiano di essere perpetrati e diffusi stereotipi e *standard* di bellezza dannosi e discriminatori, frutto, del resto, del secondo piano, ossia i pregiudizi, i valori e le finalità di chi crea e diffonde simili filtri.

2. Relazioni scisse e iperrealità di genere

Per inquadrare teoricamente le questioni etiche sollevate dal sistema di IA in esame, è necessario partire dal funzionamento della tecnologia che ne è alla base: *Generative Adversarial Network* (GAN)¹⁵, tecnologia di apprendimento automatico che mette in competizione tra loro diverse reti neurali. *Bold Glamour* è un processo bifasico, in cui gli elementi del volto ripreso dalla fotocamera sono confrontati con un set di immagini e con esso successivamente ricombinati. Due sono le questioni principali sollevate: la sua estrema efficacia e l'adeguamento delle più diverse fisionomie a un unico *standard*¹⁶, che trovano fondamento nell'innovazione apportata dalla tecnologia all'opera. Non abbiamo più infatti elementi sovrapposti più o meno bene al dato reale, come accadeva con i filtri precedenti, ma la produzione di un'immagine completamente nuova. Il filtro tende a ricondurre i volti reali a un unico modello, non solo sotto l'aspetto fisiognomico, ma anche culturale: applica un certo stile di trucco e lo fa soltanto con i volti che decodifica e poi ri-codifica – in questo doppio movimento di confronto e rielaborazione – come femminili. La realtà viene così parametrata, rapportandola a dei termini di riferimento che contengono un giudizio di valore – il filtro è inteso come migliorativo – e con essi armonizzata.

Prendendo le mosse da queste premesse, è ora possibile individuare alcuni punti di riferimento concettuali utili per tracciare un fondamento teorico all'interpretazione e valutazione del caso di studio. Si parla di GAN¹⁷ anche in relazione al fenomeno dei *deep fake*, quel sistema di IA che consiste nel sintetizzare immagini verosimili al punto da essere difficilmente distinguibili dalla realtà¹⁸, spesso associate alla diffusione di *fake news*. E queste tecnologie, come *Bold Glamour*, possono essere lette a partire dalla nozione di iperrealità¹⁹. Il concetto, teorizzato nel 1981 dal sociologo e filosofo francese Baudrillard in *Simulacres et simulation*²⁰, fa riferimento alla creazione di modelli del reale la cui origine non si può più rintracciare nello stesso. Il contenitore non sostituisce, ma prevale sul contenuto, il significativo sul significato, la rappresentazione sul suo oggetto.

Le tecnologie di IA sono addestrate a partire da dati che nel reale hanno di fatto origine, se non altro perché prodotti a partire da un esercizio dell'intelligenza umana e perché tesi a essere riconoscibili e intelligibili dalla stessa: come nel caso di *Bold Glamour*, la persona nell'atto di rappresentarsi attra-

¹⁵ W. PENDERGRASS, *Artificial intelligence and its potential harm through the use of generative adversarial network image filters on Tiktok*, in *Issues in Information Systems*, 24, 2023, 113-127.

¹⁶ *Infra*.

¹⁷ W. PENDERGRASS, *op.cit.*

¹⁸ Cfr. M. WESTERLUND, *The Emergence of Deepfake Technology: A Review*, in *Technology Innovation Management Review*, 9, 2019, 39-52.

¹⁹ Cfr. J. NOSTA, *The 'Scientific Simulacra': When AI and hyperreality Collide*, 2023 <https://johnnosta.medium.com/the-scientific-simulacra-when-ai-and-hyperreality-collide-b676eb265045> (ultima consultazione 07/07/2024).

²⁰ J. BAUDRILLARD, *Simulacres et simulation*, Paris, 1981.

verso fotografie e video, insieme al set di dati di riferimento, prodotto in un momento precedente e ora disponibile per essere trasformato. Tuttavia, l'oggetto iperreale è caratterizzato da un'autonomia dai suoi *input* senza precedenti, finendo per imporsi sul reale stesso: la simulazione si colloca, cioè, sullo stesso piano delle informazioni che elabora, modificando così l'intero sistema con cui l'utente non solo si interfaccia, ma di cui fa parte. Pertanto, abbiamo un soggetto umano che cede a una macchina la sua *agency*, la sua capacità di agire rispetto alle proprie forme di partecipazione al mondo sociale – pratica che richiama ciò che Floridi²¹ definisce «potere di scissione» dell'IA e che descrive come un «divorzio senza precedenti tra l'intelligenza e la capacità di agire».

E in tal senso la dimensione di genere, nei filtri *social*, costituisce un ampio riferimento per la capacità di agire. Poiché il genere è qualcosa che «si fa», che si performa, in un infinito processo di negoziazione e di *feedback* tra l'individuo e il contesto, tra il genere come categoria di assegnazione e il genere come identificazione. I sociologi West e Zimmerman, con lo storico articolo *Doing gender*²², hanno posto una pietra miliare negli studi sul genere, interpretato attraverso la lente della pratica e della reiterazione, contestando l'assunto per cui «*doing gender merely involves making use of discrete, well-defined bundles of behavior that can simply be plugged into interactional situations to produce recognizable enactments of masculinity and femininity*» e affermando che «*doing gender is not so easily regimented [and] to be successful, marking or displaying gender must be finely fitted to situations and modified or transformed as the occasion demands*»²³.

Si potrebbe, dunque, sostenere che l'utilizzo di *Bold Glamour* definisca le condizioni stesse della partecipazione dell'utente, che non ha più bisogno di padroneggiare – e persino incarnare – alcuni codici necessari a inserirsi nella rete sociale di riferimento. Legandoci alla riflessione dei biologi ed epistemologi Maturana e Varela, in relazione alla nozione di sistemi autopoietici²⁴, nonché all'elaborazione portata avanti nell'ambito dell'*Actor-Network Theory* dal filosofo Latour²⁵, potremmo affermare che l'IA favorisca l'accoppiamento dell'utente a un determinato gruppo legato da linguaggi e valori comuni, contribuendo alla riproduzione del sistema sociale stesso.

La nozione di sistema autopoietico, infatti, pone l'accento sui *feedback* che continuamente riproducono e, allo stesso tempo, ridefiniscono il sistema. L'IA, così come trova applicazione nei GAN, vede in azione il medesimo principio: ciò che desideriamo evidenziare con questo contributo è che il soggetto umano, nel suo interfacciarsi con l'IA, entra a sua volta a far parte del sistema e del processo trasformativo che ne coinvolge tutte le componenti, contestualmente quale attore e oggetto del mutamento.

La riflessione su *Bold Glamour* permette così di evidenziare la dimensione oggettuale della soggettività che può, dunque, essere trattata come un codice linguistico: l'IA come la conosciamo oggi, infatti, è stata resa possibile dalla svolta statistica di Jelinek²⁶, innovazione che, laddove applicata alla fisicità umana, con tutti i suoi marcatori sociali, può dare adito a un meccanismo di *feedback* normativo e

²¹ L. FLORIDI, *op. cit.*, 26 e ss.

²² C. WEST, D. H. ZIMMERMAN, *Doing gender*, in *Gender & Society*, 2, 1987.

²³ C. WEST, D. H. ZIMMERMAN, *op.cit.*, 135.

²⁴ H. R. MATURANA, F. J. VARELA, *Autopoiesi e cognizione. La realizzazione del vivente*, Venezia, 2022.

²⁵ B. LATOUR, *Reassembling the social: An introduction to actor-network-theory*, Oxford, 2007.

²⁶ J. MCMAHON, F.J. SMITH, *A Review of Statistical Language Processing Techniques*, in *Artificial Intelligence Review*, 12, 1998, 347–391.



dettare i criteri dell'inclusione e dell'esclusione. In questo emerge la natura di sistema socio-tecnico²⁷ dell'IA, che non solo assume un ruolo attoriale, ma potrebbe persino giungere a declassare l'utente nel suo fare e partecipare. Gli effetti di questo tipo di pratiche attengono quindi alla sfera dell'inclusione sociale: resta da determinare il ruolo dei sistemi di IA in relazione allo sviluppo della persona.

Appare però evidente come, tornando a Baudrillard, il simulacro - che qui possiamo intendere come un *avatar* - e i suoi *standard* di leggibilità abbiano un impatto ampio e profondo sulla realtà sociale e sui processi di soggettivazione che vi intervengono, in particolare, in questo contesto, rispetto al sistema di genere, che trova nelle applicazioni di IA un nuovo e potente dispositivo disciplinare.

3. Implicazioni etiche e tutela dei diritti fondamentali nell'utilizzo del filtro Bold Glamour

Le tecnologie di IA, dunque, potrebbero certamente paragonarsi a prodotti *prêt-à-porter*, essendo riproducibili, utilizzabili e trasportabili da chiunque. Tali caratteristiche, sintomatiche di un avanzamento tecnologico inarrestabile²⁸ e imprevedibile²⁹, rendono necessario analizzare i risvolti etici e i risultati che questi sono idonei a conseguire. Non è un caso, come già anticipato *supra*, che diversi attori siano coinvolti affinché i sistemi di IA siano puntualmente disciplinati, si faccia riferimento alla rilevante attività dell'Unione europea, intenzionata a ripetere quanto avvenuto con il GDPR³⁰, auspicandosi l'effetto Bruxelles³¹. In tal senso, si può richiamare l'*AI Act*³², il cui articolato, rispetto al contenuto proposto dalla Commissione europea, risulta maggiormente incentrato sulla tutela dei diritti fondamentali³³, apparendo candidarsi, ad ogni buon conto, alla concezione di un IA *human-centric* e *trustworthy*³⁴, seppure sembri piuttosto avventato pronunciarsi sull'efficacia delle disposizioni³⁵. In estrema sintesi, l'assetto dell'articolato è basato sul rischio (RBR), che risulta il fattore fondante dello sviluppo della normativa³⁶, che differisce in larga misura in base al livello di rischio riscontrabile nell'impiego dell'IA³⁷: a seconda della pericolosità e della predittività dei risultati dell'utilizzo di tali sistemi aumenta o diminuisce la forbice di discrezionalità dell'ideatore, nonché la libertà nel contenuto della co-regolamentazione³⁸. Il Regolamento, seppur trovi la propria base giuridica negli articoli

²⁷ Si v. *AI Act*, nota 8.

²⁸ V. P. CONTUCCI, *Rivoluzione intelligenza artificiale. Sfide, rischi e opportunità*, Bari, 2023.

²⁹ G. MOBILIO, *op. cit.*

³⁰ Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione dei dati e che abroga la direttiva 95/46/CE, del 27 aprile 2016, in GUUE L 119, del 4 maggio 2016, 1.

³¹ Da intendere l'idoneità dell'Unione europea di candidarsi come legislatore globale. Si v. A. BRADFORD, *The Bruxelles Effect: How the European Union Rules the World*, Oxford, 2020; E. CIRONE, *op. cit.*

³² Regolamento UE 2024/1689 del Parlamento Europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale.

³³ Considerando 7 e 8 del Regolamento, cfr. E. CIRONE, *op. cit.*; G. MOBILIO, *op. cit.*

³⁴ A. D'ALOIA, *op. cit.*

³⁵ E. CIRONE, *op. cit.*

³⁶ Cfr. G. DE GREGORIO, P. DUNN, *The European risk-based approaches: Connecting constitutional dots in the digital age*, in *Common Market Law Review*, 2, 2022, p. 475.

³⁷ Tanto emerge chiaramente dalla suddivisione in Capi del Regolamento.

³⁸ Cfr. G. MOBILIO, *op. cit.*

16 e 114 TFUE – vale la pena sottolineare che i sistemi di IA sono considerati dei servizi³⁹ – risponde, perlomeno astrattamente, a quella concezione per cui l'IA debba essere portatrice di un *ethical purpose*⁴⁰. Concezione che emerge parimenti da quanto evidenziato negli Orientamenti etici per un'IA affidabile⁴¹, documento in cui la Commissione europea fornisce le linee guida, nonché la definizione di un sistema IA etico ed affidabile, il quale deve rispondere a determinate componenti sintetizzabili in *legalità, eticità e robustezza*, operanti armonicamente e capaci di sovrapporsi⁴². La guida può essere applicata a qualsiasi macchina, tra cui *Bold Glamour*: l'accessibilità del filtro è pressoché universale ed è proprio l'universalità a rendere fortemente indispensabile che questo sistema rispecchi e risponda a *standard* di eticità e affidabilità elevati, considerando che l'utenza è composta in larga misura da minori che potrebbero mal beneficiare di soluzioni volte alla modifica del proprio viso per adattarlo a *standard* artificiali e irraggiungibili di bellezza. Posti i primi ostacoli verso un'effettiva e strutturale analisi della corrispondenza dei requisiti della *check-list* del documento della Commissione⁴³, l'indagine proverà a comprendere se *Bold Glamour* costituisca o meno un esempio di IA affidabile ed etica. È oltremodo chiaro, come già sottolineato, che l'applicazione del filtro al volto dell'utenza comporti delle implicazioni etiche specialmente dal punto visuale dei diritti fondamentali, della affidabilità, della trasparenza⁴⁴, nonché della diversità, non discriminazione ed equità, valori che, stando al contenuto della linee guida di Tiktok, si ascrivono tra i principi della *community*. Si pone, anzitutto, il dubbio circa l'effettiva analisi dell'impatto negativo che l'effetto può sortire: quali sono le tutele poste allo scopo di evitare l'eccessiva sessualizzazione⁴⁵ degli utenti – soprattutto minori – derivante dalla modifica del loro aspetto in soggetti più maturi e piacenti? Esistono strumenti idonei a limitare la sottoposizione passiva alla visione delle immagini standardizzate e *perfezionate* dal filtro? Ad oggi, non risulta essere presente alcuna opzione che consenta all'utente di richiedere la disattivazione di tali contenuti, non esiste, in altre parole, un «pulsante d'arresto»⁴⁶. Invero, tra le forme di tutela previste per i minori, nella sezione loro dedicata, «sicurezza e benessere dei più giovani», il rimando a filtri costruiti su sistemi di IA idonei a modificare i caratteri fisionomici della persona è assente, mancando nella lista di contenuti ascrivibili come pericolosi per la salute mentale e il benessere dei giovani, tra cui si annoverano, tra altri, i contenuti sui DCA, salvo siano di promozione verso un percorso riabilitativo, e quelli di incentivazione di pratiche di chirurgia estetica⁴⁷. Pertanto, il filtro non sembrerebbe soddisfare i requisiti relativi ai diritti fondamentali richiesti dalla *check-list* della Commissione⁴⁸. Ciò, tuttavia, può dirsi anche in riferimento alla sua affidabilità e alla sua riproducibilità, non appare possibile selezionare e segnalare eventuali errori commessi dall'IA in sede di applicazione dell'effetto,

³⁹ Cfr. M. INGLESE, *Il regolamento sull'intelligenza artificiale come atto per il completamento e il buon funzionamento del mercato interno*, in *Quaderni ASIDUE*, 2, 2024.

⁴⁰ Cfr. A.D'ALOIA, *op. cit.*

⁴¹ Commissione europea, *Orientamenti etici per un'IA affidabile*, Bruxelles, 2019.

⁴² *Ibidem*, 5 e ss.

⁴³ Si segnala, infatti, che nei Termini e politiche di Tiktok non sono specificate le modalità di raccolta dei dati per ciò che concerne l'impiego di filtri che si applicano al volto.

⁴⁴ Sulla spiegabilità T.E. FROSINI, *op. cit.*

⁴⁵ Cfr. W. PENDERGRASS, *op. cit.*

⁴⁶ Commissione europea, *op. cit.*, 34.

⁴⁷ Sezione «Salute mentale e comportamentale» in «Salute e benessere dei più giovani».

⁴⁸ Commissione europea, *op. cit.*, 34.



es. la diversificazione del filtro tra donne e uomini, oltre ad apparire estremamente binaria e discriminatoria, è risultata fallace anche nel non discernere chi si definisce donna e chi uomo a causa dei propri tratti più *femminili* o *maschili*. Il sistema è, dunque, poco preciso e trasparente – requisiti per l’ottenimento di un’IA affidabile –: nel valutare l’opportunità di applicare o meno il *make-up*, il sistema non è affatto spiegabile – appare ascrivibile tra i sistemi noti come *black box*⁴⁹ – non vengono ricostruite le ragioni poste alla base della decisione. La carenza di spiegabilità è naturale portatrice di problematiche, essa non solo risulta fondamentale per rispondere ai requisiti di trasparenza⁵⁰, bensì è posta alla base dell’eventuale sindacabilità del sistema di IA⁵¹, nonché degli strumenti di educazione dell’algoritmo⁵².

Quanto finora premesso non risulta affatto incoraggiante se si considera che non sono stati posti, ancor’oggi, dei correttivi ai diversi *bias* ivi presenti. *Bold Glamour*, invero, fallisce appieno il test sulla diversità, non discriminazione ed equità richiesti dalla *check-list*. È chiaro – o perlomeno, sperabile, laddove non si voglia accettare l’intenzionalità della proposizione di *standard* di bellezza occidentali – che alcuna strategia sia stata prevista per evitare di creare, rafforzare e/o perpetrare i c.d. *unfair bias*, basti far riferimento alle più palesi discriminazioni di cui si è reso protagonista il filtro: persone nere a cui la pelle e gli occhi vengono schiariti, donne a cui viene automaticamente applicato del *make-up*, uomini a cui, al contrario, viene automaticamente eliminato, con un costante assente: le persone *queer*. *Bold Glamour*, pertanto, sembra rispondere ad una società ormai obsoleta, attanagliata dal binarismo e dalla concezione escludente e discriminatoria della bellezza. Il disegno restituito è in grado di rafforzare i c.d. *confirmation bias*, pregiudizi di conferma, sulla psicologia umana: la bellezza è e dev’essere solo occidentale⁵³.

Peraltro, stando alla *check-list* in esame, al fine di succedere negli *standard* di trasparenza rileva finanche lo scopo dell’IA, la precisazione dello stesso e il vantaggio che l’utente può trarre dall’utilizzo, è forse d’uopo domandarsi se un effetto che propone *standard* di bellezza inarrivabili sia etico e restituisca un valore aggiunto alla società? Inoltre, di quanta conoscenza godono effettivamente gli utenti dell’utilizzo del filtro da parte di altri *users*? Seppure la sua applicazione sia rintracciabile qualora il realizzatore del video lo posti sulla piattaforma originale, il video sarà salvabile e ripubblicabile e il *disclaimer* non sarà presente, rendendo quasi impossibile a chi osserva discernere se quello è l’aspetto effettivo della persona: gli utenti non avranno la consapevolezza di interagire «con un sistema di IA»⁵⁴. Tanto può essere contenitore di molteplici conseguenze quali attaccamento, influenza o svilimento dell’essere umano⁵⁵. Lo scenario risulta più scoraggiante considerando che nella *policy* di Tik-

⁴⁹ Sul punto cfr. A. D’ALOIA, *op. cit.*

⁵⁰ La spiegabilità dei sistemi di IA è dedicato il Capo IV dell’*AI Act*.

⁵¹ Cfr. T.E. FROSINI, *op. cit.*

⁵² Cfr. A. D’ALOIA, *op. cit.*

⁵³ Commissione europea, *op. cit.* punto 44, 15. Sulla polarizzazione dei sistemi di intelligenza artificiale si rimanda a Commissione europea, *Draft Ethics Guidelines for Trustworthy AI*, 2018; Cfr. A. D’ALOIA, *op. cit.*

⁵⁴ *Ibidem*, punto 45, 131.

⁵⁵ M. MADARY, T. K. METZINGER, *Real Virtuality: A Code of Ethical Conduct. Recommendations for Good Scientific Practice and the Consumers of VR-Technology*, in *Frontiers in Robotics and IA*, 3, 2016.

Tok relativa alla segnalazione di contenuti creati con l'IA⁵⁶, l'invito all'apposizione dell'etichetta non è richiesto se le «modifiche non alterano il significato principale del contenuto»⁵⁷.

Quanto detto è idoneo ad esacerbare le complesse vulnerabilità sofferte dagli utenti, un recente studio ha rilevato che il 94% delle donne e delle persone non binarie partecipanti hanno dichiarato di sentirsi sotto pressione e di dover assumere un determinato aspetto fisico a causa di contenuti simili⁵⁸. Sulla base di questa analisi, risulta inevitabile escludere il *Bold Glamour* tra i sistemi di IA affidabili ed etici. Sarebbe, forse, opportuno domandarsi se esperimenti del genere siano necessari in una società già sofferente ed altamente performativa e costituiscano un valore aggiunto alla società e perseguano l'*ethical purpose* di cui appare doversi connotare qualsiasi sistema di IA?⁵⁹

4. Alcuni cenni conclusivi

Alla luce dell'analisi sin qui condotta, preme da ultimo ragionare sulla mitigazione dei potenziali danni derivanti dall'impiego dei filtri: le prospettive di intervento sono molteplici – tecnologiche, legislative e sociali – e richiedono tanto il coinvolgimento di attori esterni quanto una consapevolezza interna da parte degli utenti.

Dal punto di vista normativo, è emerso come l'intervento legislativo spesso si dimostri inefficace se non controproducente: si veda ad esempio l'infruttuoso tentativo di alcuni paesi di contrastare l'insoddisfazione dell'immagine corporea mediante leggi che vietano la pubblicità di immagini potenzialmente fuorvianti⁶⁰ o ancora l'introduzione di *disclaimer*, che tuttavia sembrerebbero non solo non ridurre l'insoddisfazione, ma addirittura promuovere confronti sociali nocivi⁶¹. Dunque, il rischio è che l'azione legislativa, sebbene appaia un approccio diretto, finisca per risultare troppo ampia e lenta per affrontare il problema in modo strutturale.

Da un punto di vista tecnologico, l'eliminazione dei filtri incriminati, da parte del *provider* sembrerebbe una soluzione immediata, ma in concreto risulta di difficile attuazione, sollevando anche questioni legate al ruolo degli Internet Service Providers nella moderazione dei contenuti. Data la crescente diffusione della tecnologia IA, inclusa quella basata su GAN, la sua regolamentazione risulta poi complessa: anche se esistono meccanismi per identificare e segnalare tali filtri, come l'auto-identificazione o l'installazione di marcatori digitali, questi potrebbero essere facilmente aggirati⁶². Attualmente,

⁵⁶ Cfr. le linee guida della *community* di Tiktok, Integrità e autenticità, Contenuti multimediali modificati e contenuti generati dall'intelligenza artificiale (AIGC).

⁵⁷ *Ibidem*.

⁵⁸ R. GILL, *Changing the perfect picture: Smartphones, social media and appearance pressures*, 2021, https://www.city.ac.uk/_data/assets/pdf_file/0005/597209/Parliament-Report-web.pdf (ultima consultazione 10/07/2024).

⁵⁹ Cfr. D'ALOIA, *op. cit.*

⁶⁰ G. SHARP, Y. GERRARD, *The body image "problem" on social media: Novel directions for the field*, in *Body Image*, 41, 2022.

⁶¹ E.S. DANTHINNE, F.E. GIORGIANNI, R.F. RODGERS, *Labels to prevent the detrimental effects of media on body image: A systematic review and meta-analysis*, in *International Journal on Eating Disorder*, 53, 2024.

⁶² F. SHAN, A.R. CILLO, C. CARDELLO, D.Y. YUAN, S.R. KUNNING, J. CUI, C. LAMPENFELD, A.M. WILLIAMS, A.P. McDONOUGH, A. PENNATHUR, J.D. LUKETICH, J.M. KIRKWOOD, R.L. FERRIS, T.C. BRUNO, C.J. WORKMAN, P.V. BENOS, D.A.A. VIGNALI, *Integrated BATF transcriptional network regulates suppressive intratumoral regulatory T cells*, in *Science Immunology*, 7, 76, 2022.



la limitazione del tempo di condivisione dei video su TikTok contribuisce a mitigare l'impatto dei filtri, ma resta dubbia la sua utilità a lungo termine. Considerazioni analoghe riguardano la percezione delle immagini manipolate dalla tecnologia, che solleva interrogativi sull'equivalenza tra IA e l'uomo. Dunque, una soluzione tecnologica, sebbene possa sembrare praticabile, potrebbe anche in questo caso rivelarsi inefficace nel lungo periodo.

Da un punto di vista sociale esterno, l'attivismo per la consapevolezza rappresenta un approccio ampio per affrontare il problema, in questo senso vanno letti movimenti come il *body positivity* che sfidano gli ideali estetici convenzionali e promuovono l'accettazione del corpo in tutte le sue forme; del resto, l'impatto dei filtri dannosi è stato riconosciuto dalle piattaforme stesse allorché hanno provveduto alla loro rimozione. Come visto, problematico è l'uso dei filtri da parte degli *influencer* e la promozione di ideali estetici irreali, rispetto ai quali le sfide rimangono significative. Ciò che risulta cruciale è la consapevolezza individuale del contesto tecnologico e sociale circostante, ossia la capacità di riconoscere le distorsioni generate dalle tecnologie digitali e comprendere la propria relazione con le stesse. L'educazione riguardo alle alterazioni nelle rappresentazioni digitali e il sostegno reciproco tra pari possono, pertanto, contribuire a mitigare le conseguenze dannose delle immagini manipolate⁶³.

Risulta ancora una volta chiaro che l'approccio in tal senso debba essere sfaccettato su diversi livelli: mentre l'azione normativa e tecnologica possono fornire alcuni strumenti per mitigare il problema, la consapevolezza sociale e individuale resta fondamentale per contrastare le potenziali lesioni dei diritti fondamentali⁶⁴. In tal senso è auspicabile un ulteriore sviluppo dell'analisi circa l'impatto dell'uso dei filtri, in specie con riferimento a quei gruppi di cui mancano ad oggi dati chiari, come persone trans* e queer.

Da tanto premesso, è evidente come sia ancora lontana una soluzione a lungo termine circa l'utilizzo di filtri bellezza come *Bold Glamour*, che tuttavia non può non partire dalla responsabilità del *provider* di questi strumenti, che hanno un ruolo centrale nell'immissione e gestione dei contenuti realizzati con tale filtro: la piattaforma, detenendo il controllo sui filtri disponibili e sulle politiche di utilizzo, può essere decisiva nella mitigazione dei danni⁶⁵. Questo implica non solo la necessità di implementare sistemi di monitoraggio efficaci e di reagire prontamente alle segnalazioni di abusi, ma anche l'adozione di un approccio proattivo nella valutazione, nella limitazione e anche nell'opportunità dell'uso di filtri potenzialmente dannosi⁶⁶, implementando le *policy* della *community*. In tal senso, TikTok potrebbe collaborare con esperti di salute mentale e *stakeholders* per sviluppare linee guida più rigorose e, contestualmente, sensibilizzare gli utenti sugli effetti psicologici derivanti dall'uso eccessivo di tali filtri.

⁶³ M. TIGGEMANN, V.G. VELISSARIS, *The effect of viewing challenging "reality check" Instagram comments on women's body image*, in *Body Image*, 33, 2022.

⁶⁴ W. PENDERGRASS, *op. cit.*

⁶⁵ J.M. TRAMMEL, *Artificial Intelligence for Social Evil: Exploring How AI and Beauty Filters Perpetuate Colorism—Lessons Learned from a Colorism Giant*, in K. LANGMIA (a cura di), *Black Communication in the Age of Disinformation*, Londra, 2023.

⁶⁶ *Ibidem.*