

SUPERSTRATO

SUPERSTRATO

Davide Chiappini

Sul rapporto fra essere umano e intelligenze
artificiali nell'era della postverità



POLITECNICO
MILANO 1863

963963

Relatore: Salvatore Zingale

A.A. 2021/2022

Politecnico di Milano

Scuola del Design

Tesi di Laurea Magistrale in Design della Comunicazione

Maggio 2023

PREMESSA	13
01. IL CYBORG	19
1.1 Umano e tecnologia, da sempre	22
1.2 Ibridazione	24
1.3 Essere mostro	29
02. CON L'IA SIAMO TUTTI CREATORI DI CONTENUTI	33
2.1 Genesi dell'intelligenza artificiale	36
2.2 AARON e la prima arte generata	38
2.3 Panorama attuale	43
2.4 Il rapporto con la nuova specie	48
03. LA MARIONETTA. COME L'UMANO SI REPLICA ONLINE	53
3.1 Storia della manipolazione	56
3.2 Cheapfake, radici del fenomeno	62
3.3 Deepfake e media sintetici	69
3.4 Finte celebrità	74
3.5 Sintesi vocale e clonazione	82
3.6 Modello Diffusion, dal testo all'immagine	85
3.7 Marionette e resurrezione	92
04. VIVREMO L'INFOCALYPSE	113
4.1 Come i governi manipolano l'informazione	116
4.2 Infocalypse, polarizzazione e rabbit hole	120

4.3 Contenuti grassroots e cultura partecipativa	126
4.3.1 Misinformazione sulle dark platform: Bitchute	128
4.3.2 Umani sintetizzati e pornografia: Unstable Diffusion	134
4.4 La tecnologia è trasparente	140
05. NULLA È VERO NEL SUPERSTRATO	143
5.1 Designer o creatori di contenuti?	146
5.2 Rapporto con le alterità generate	149
5.3 Apatia della realtà, postverità e iperreale	154
5.4 Riconoscere i contenuti: il nostro compito	158
5.5 Contromisure dall'alto	162
06. CONCLUSIONI	167
Il Mondo Nuovo	169
BIBLIOGRAFIA	173
INDICE FIGURE	179
RINGRAZIAMENTI	187

La democratizzazione di modelli di intelligenza artificiale per la manipolazione e la generazione di prodotti audiovisivi permette a chiunque di essere creatore di contenuti. Grazie a queste tecnologie è infatti possibile replicare l'essere umano nella sua forma estetica e comportamentale, generando delle marionette controllabili a piacere e diffondibili online.

I social network, seguendo le orme dei media tradizionali, favoriscono la diffondibilità di contenuti polarizzati a sfondo disinformativo o misinformativo. Queste piattaforme sono quindi terreno fertile per la nascita di un superstrato dell'informazione completamente manipolabile e slegato dalla realtà, uno scenario di infocalypse dove il confine tra vero e falso diventa labile e sfocia in una post-verità. Una persona, accedendo al mondo social, è quindi risucchiata in rabbit hole senza fine dove ogni contenuto è polarizzato e non controllato.

Il superstrato è abitato e frequentato da gruppi di utenti che collaborano e distribuiscono informazioni più o meno consapevolmente, attingendo a oggetti o eventi del mondo reale.

La tesi si rivolge quindi a un pubblico di designer e, più in generale, di utenti del mondo online che vuole comprendere le possibili dinamiche e implicazioni di un'ibridazione "senza barriere all'ingresso" con le nuove tecnologie.

Partendo dagli ideali postumanisti, la tesi è un'analisi approfondita dei rischi e sviluppi della diffusione delle intelligenze artificiali, presentando le possibili relazioni che l'utente può avere con

le macchine e con gli output generati dalle stesse. Come avviene il dialogo con le nuove alterità non-umane? Qual è quindi la differenza fra creatori di contenuti e progettisti? Ma soprattutto, come si può agire con cognizione di causa senza farsi trasportare dall'euforia provocata dall'avvento dell'IA?

Vivendo in una nuova realtà di comunicazione di contenuti audiovisivi generati, risulta quindi necessario possedere un'educazione progettuale, sia per evitare di diffondere output misinformativi, sia per comprenderli e imparare a riconoscerli. Infine, toccando sia la sfera individuale che quella collettiva, si pone uno sguardo sulle possibili soluzioni dal basso e dall'alto per arginare il problema e gestire il flusso di informazioni.

The democratization of artificial intelligence models for the manipulation and generation of audiovisual products allows anyone to be a content creator. Indeed, thanks to these technologies, it is possible to replicate human beings in their aesthetic and behavioral form, generating puppets that can be controlled at will and disseminated online.

Social networks, following in the footsteps of traditional media, favor the visibility of polarized content with a disinformation or misinformation background. These platforms are thus fertile ground for the emergence of a superstrate of information that is completely manipulable and disconnected from reality, an infocalypse scenario where the line between true and false becomes blurred and results in post-truth. A person, accessing the social world, is then sucked into endless rabbit holes where all content is polarized and uncontrolled.

The superstratum is inhabited and frequented by groups of users who collaborate and distribute information more or less consciously, drawing on objects or events in the real world.

The thesis is therefore aimed at an audience of designers and, more generally, users of the online world who want to understand the possible dynamics and implications of a “no barriers to entry” hybridization with new technologies.

Starting from posthumanist ideals, the thesis is an in-depth analysis of the possible risks and developments of the spread of artificial intelligences, presenting the possible relationships the user can have with machines and the outputs generated

by them. How does the dialogue with the new non-human alterities take place? What then is the difference between content creators and designers? More importantly, how can one act knowledgeably without getting carried away by the euphoria provoked by the advent of AI?

Living in a new reality of communication of generated audiovisual content, it is therefore necessary to possess a design education, both to avoid spreading misinformative outputs and to understand and learn to recognize them. Finally, touching on both the individual and collective spheres, there is a look at possible bottom-up and top-down solutions to curb the problem and manage the flow of information.

PREMESSA

Superstrato nasce dalla necessità di voler dare forma e definire un luogo digitale e non tangibile, ma in un certo modo reale: il mondo dell'informazione e dei contenuti online. Questo universo è già esistente e viene progressivamente contaminato e plasmato nella sua identità dalla grande mole di output generati con l'utilizzo di intelligenze artificiali. In linguistica storica, il superstrato è:

Lo strato linguistico che si sovrappone, in seguito a colonizzazione, invasione, predominio politico e culturale di un altro popolo, alla lingua già in uso in una determinata area, e che provoca nella lingua indigena cambiamenti fonetici, morfosintattici e soprattutto lessicali. (Treccani s.d.-e)

Il prefisso *super-* significa letteralmente "sopra", è un concetto che esprime addizione, sovrapposizione, qualcosa in più rispetto al normale. Indica quindi un processo in eccesso rispetto alla quotidianità.

Spesso, nel mondo della linguistica il termine superstrato viene affiancato al *substrato*, il suo opposto, ovvero quella lingua che, nativa di un certo popolo, viene sovrastata e plasmata. Il superstrato e il substrato non si escludono a vicenda. Sono infatti in un rapporto reciproco, parallelo, dove entrambi esistono e l'uno influenza l'altro, ma dove il primo esercita un potere estremamente più forte del secondo, cambiandone i connotati (Treccani s.d.-d).

Prendendo in prestito il termine, ho voluto definire con questa metafora il *superstrato digitale dell'informazione*: un mondo parallelo composto da contenuti audiovisivi pubblicati

e condivisi sui nuovi media, i social network, dove tutto è lecito e dove realtà e finzione si confondono. Come la nuova lingua che sovrasta una popolazione nativa, il superstrato digitale si sovrappone alla nostra realtà, prendendone in prestito connotati, oggetti, eventi, persone per poi rappresentarli come *simulacri* (Baudrillard 1981) in grado di ingannare l'osservatore con una sovrabbondanza di informazioni non verificate e cambiando per sempre il nostro modo di vivere.

La rivoluzione digitale crea un legame sempre più stretto tra dati, testi, immagini e suoni, facilita l'espressione, la trasmissione, la riproducibilità del messaggio. In altri termini, l'infospazio digitale già oggi si presenta strutturalmente ridondante. (Marchesini 2002: 387)

Queste rappresentazioni possono essere completamente modificabili o addirittura generabili, sia dagli utenti che dai media stessi, tramite l'utilizzo di nuovi strumenti intelligenti, servizi e software con cui ci si può interfacciare con una semplice interazione come un input testuale. Le rappresentazioni più efficaci e per cui siamo più influenzabili sono sicuramente quelle che contengono al loro interno delle repliche della figura umana, delle marionette. Spesso sono personaggi noti, politici o attori. Quelli che definiamo come *deepfake* o contenuti *generati* cambiano quindi il paradigma della nostra percezione facendoci domandare: è reale quello che stiamo guardando?

L'obiettivo ultimo, oltre a comprendere il significato di superstrato e fornire un monito sulla situazione attuale, è quello di indagare possibili soluzioni per arginare le problematiche tramite l'adozione di un metodo progettuale nell'utilizzo di strumenti di intelligenza artificiale e nel rapportarsi con la *disinformazione* e la *misinformazione* che popolano il web. Essere creatori di contenuti non significa necessariamente essere designer. Nonostante entrambi possano essere attori nel superstrato, chi contribuisce alla sua formazione è chi si lascia trasportare dall'utilizzo dello strumento senza comprendere la propria collocazione nel contesto sociale di riferimento.

Per la corretta comprensione dell'elaborato, quindi, ritengo necessario chiarire la differenza tra misinformazione e disinformazione. Per disinformazione si intende «diffusione intenzionale di notizie o informazioni inesatte o distorte allo scopo di influenzare le azioni e le scelte di qualcuno» (Treccani s.d.-a), mentre con il termine misinformazione si fa riferimento a un'informazione «non accurata, inattendibile, i cui contenuti, diffusi frettolosamente, rischiano di essere accettati come veritieri perché difficili o impossibili da verificare» (Treccani s.d.-c). Alla luce di ciò possiamo osservare come, pur facendo parte del-

lo stesso universo, i due fenomeni non condividano lo stesso significato. Creare disinformazione implica un'azione volontaria di depistaggio della verità, una consapevolezza di star diffondendo informazioni distorte con un fine ultimo, il più delle volte, maligno. Questo vuol dire che, chi compie questa azione, può essere un progettista, poiché progettare non è un'azione esclusiva di chi fa del bene. Al contrario, generare misinformazione è il più delle volte un'azione involontaria. Sempre da Treccani (s.d.) scopriamo infatti che quest'ultima è un'attività che ad esempio implica la condivisione di un post, un *re-tweet* di una notizia, senza la consapevolezza di star facendo del male e di star influenzando l'opinione pubblica. Chi sparge misinformazione ha una colpa, ma non ha un movente doloso.

	INFORMAZIONE	MIS- INFORMAZIONE	DIS- INFORMAZIONE
<i>VERA</i>	●	◐	◐
<i>COMPLETEA</i>	◐	◐	◐
<i>ATTUALE</i>	●	◐	◐
<i>INFORMATIVA</i>	●	●	●
<i>INGANNEVOLE</i>	○	○	●



↑ FIG. 0.1 RIELABORAZIONE DELLO SCHEMA DI KARLOVA E FISHER, 2013. SINTESI DELLE CARATTERISTICHE DI INFORMAZIONE, MISINFORMAZIONE E DISINFORMAZIONE

La tesi si sviluppa lungo cinque capitoli. Il capitolo 1 mira a fornire una base al lettore riguardo il rapporto fra umano e tecnologia, affermando come questo concetto, di uso comune per riferirsi a invenzioni recenti o al mondo digitale, in realtà è radicato nella cultura stessa della civiltà umana ed è quello che differenzia la nostra specie dal resto del pianeta. Viene poi introdotto, tramite l'espedito della figura del *cyborg*, il concetto di *postumanesimo*, utile per descrivere come la nostra civiltà si debba spostare dagli ideali della perfezione rappresentativa umanista del Novecento a una nuova realtà, fondata sull'ibridazione con l'alterità e l'ascolto di tutto ciò che è diverso. Esplicitare la visione postumanista, nell'ottica di un'ibridazione con la tecnologia, è utile per comprendere le radici della visione progettuale da applicare al mondo comunicativo odierno e futuro, oltre a essere fondamentale per chiudere il cerchio del discorso. Il capitolo si conclude con un excursus sul concetto di sottobosco (Marchesini 2002: 188) e di mostruosità. Nella natura umana è spesso rintracciabile la tendenza a far parte di subculture, gruppi di persone che si distaccano dalla massa per abbracciare degli interessi comuni. I concetti risultano particolarmente utili per evidenziare un'attitudine tipica del mondo online: formare *community*, discutere, aiutarsi reciprocamente, produrre contenuti e diffondere informazioni per un obiettivo comune.

Il capitolo 2 ha il ruolo di esplicitare il concetto di intelligenza artificiale, abbreviato con l'acronimo IA, e di come si sia evoluto dalla nascita dell'informatica a oggi. Particolare attenzione è fatta per quello che è definibile come primo esempio di IA generativa, *AARON*, un algoritmo programmato dall'artista e informatico Harold Cohen a partire dagli anni Settanta. Descrivere la storia di *AARON* diventa l'espedito per spiegare come i modelli di intelligenza artificiale odierni ragionino, di come stiano diventando sempre di più degli assistenti e introdurre il concetto di training, ovvero di allenamento della macchina. Si procede poi con una rapida descrizione del panorama attuale e di come i nuovi strumenti permettano all'umano un rapporto semplice attraverso il linguaggio e interfacce.

Nel capitolo 3 l'obiettivo è descrivere, in un percorso *step by step*, come replicare un essere umano digitalmente, esplicitando possibilità attuali ed eventuali proiezioni future. Sin dalla nascita della fotografia è possibile rintracciare una volontà di manipolazione dei contenuti visivi, e più in generale dell'informazione. Citando esempi storici è possibile infatti inquadrare la tendenza attuale e scoprire perché, visitando un qualsiasi social network, l'utente si ritrova circondato, spesso senza accorgersene, da marionette con aspetto, voce e comportamenti simili alle loro controparti reali. La *marionetta* è un'utile metafora che implica la presenza di una figura esterna, un *burattinaio*, identificabile come il creatore del contenuto, che governa i suoi

comportamenti. Grazie ai dati raccolti sul web, è possibile creare una copia carbone di un umano di riferimento, e riportare in vita figure ormai scomparse facendo compiere loro azioni mai accadute. Il discorso si conclude con una riflessione sulla capacità inventiva delle macchine, ponendo l'accento sulla loro esistenza come strumenti e non come progettisti digitali.

Il capitolo 4 introduce un'analisi dello scenario di *infocalypse* (Schick 2020: 9), o *infocalisse*, una situazione di caos informativo dove chiunque può esprimere la propria opinione e spargere contenuti online. Osservare precedenti come le campagne di disinformazione progettate dalla Russia, negli anni Ottanta e nella storia recente, aiuta a comprendere come l'introduzione di strumenti di IA generativa possa facilitare a questi attori *dall'alto* nel diffondere le proprie idee e manipolare le masse, aumentando a dismisura gli effetti dell'infocalypse. Oltre a ciò, grazie a due casi studio di realtà come *Bitchute* e *Unstable Diffusion*, viene analizzato il comportamento degli utenti in due differenti *sottoboschi*, simboli rispettivamente della disinformazione dal basso e della cultura partecipativa. Osservando le modalità d'azione degli attori al loro interno, è possibile identificare il metodo di diffusione della disinformazione e di come gli utenti possano collaborare, adottando un nuovo strumento e generando umani sintetizzati per soddisfare desideri personali. In conclusione, vi è una riflessione sulla trasparenza della tecnologia, significativa per comprendere come gli strumenti tecnologici amplino le capacità di azione dell'uomo e allo stesso tempo agiscano in modo invisibile quando vengono adoperati.

Infine, nel capitolo 5, si giunge al concetto di superstrato, di come effettivamente questo scenario di *post-verità* possa essere affrontato ed, eventualmente, evitato. Lo scopo del capitolo è mostrare nel dettaglio la differenza fra creatori di contenuti e designer, descrivendo il metodo progettuale da applicare nella progettazione con IA e nell'individuazione di marionette. Una particolare enfasi è posta sulla relazione con esse e sulla loro somiglianza con le controparti reali, analizzando il rapporto da un punto di vista semiotico. Queste alterità sono protagoniste dello scenario iperreale profetizzato da Baudrillard e la loro esistenza implica la coesistenza di realtà e di finzione. Concludendo, vengono poi indicate delle possibili strade che l'utente può percorrere per non cadere in un rifiuto della "realtà" per come viene proposta dai media, e di come anche questi ultimi, insieme agli organi statali e altre aziende, debbano attuare delle contromisure per arginare il problema.



01

IL CYBORG

1.1

UMANO E TECNOLOGIA, DA SEMPRE

1.2

IBRIDAZIONE

1.3

ESSERE MOSTRO

Affrontare le problematiche che l'era dei contenuti generati da intelligenze artificiali ci pone dinanzi è quanto mai complesso. L'incredibile mutevolezza dell'aspetto dell'informazione online non solo spaventa, ma coglie impreparati sia gli utilizzatori di questi software, che i fruitori di contenuti. Per questo è indispensabile analizzare le radici di questo scenario, rintracciabili nel rapporto fra la nostra specie e quella che chiamiamo *tecnologia*, con la quale, ora più che mai, ci troviamo in uno stato di forte ibridazione. Tuttavia, questa condizione può avere conseguenze differenti in base all'approccio scelto.

È importante infatti capire come, per raggiungere una vera sinergia con il mondo circostante (reale e online), comprendere le conseguenze delle proprie scelte ed agire superando una visione antropocentrica, sia indispensabile abbracciare gli ideali *postumanisti*, illustrati di seguito.

1.1 UMANO E TECNOLOGIA, DA SEMPRE

Siamo esseri artificiali. Il nostro modo di camminare, di mangiare, di comunicare, di sognare e amare è definito dall'utilizzo degli strumenti tecnologici che, come umani, siamo abituati ad adoperare a nostro favore e a nostro servizio.

Il termine *tecnologia* è legato intrinsecamente a numerose discipline che spaziano dal campo scientifico a quello matematico, fino a toccare il piano sociale, culturale e antropologico. La tecnologia viene definita come un

Vasto settore di ricerca [...], composto da diverse discipline [...], che ha come oggetto l'applicazione e l'uso degli strumenti tecnici in senso lato, ossia di tutto ciò [...] che può essere applicato alla soluzione di problemi pratici, all'ottimizzazione delle procedure, alla presa di decisioni, alla scelta di strategie finalizzate a determinati obiettivi. (Treccani s.d.-f)

È quindi uno sviluppo efficiente degli strumenti necessari e racchiude al suo interno un insieme di conoscenze e metodi per sviluppare e distribuire beni e servizi. La tecnologia è spesso confusa con la *tecnica*, che invece si riferisce al mero utilizzo pratico degli strumenti forniti dalla tecnologia e all'abilità quasi artigianale dell'umano nel maneggiarli.

In sintesi, potrei definire la tecnica come la *mano* con la quale eseguiamo una determinata azione, mentre la tecnologia come lo *scopo* per cui lo stiamo facendo. Molti pensano erroneamente che il concetto stesso di tecnologia e il suo rapporto con l'individuo e la società sia relativamente recente. Si è portati a credere che la tecnologia sia il computer portatile comprato l'altro ieri online, la nuova fibra ultraveloce e ultrasicura oppure l'applicazione basata su intelligenza artificiale che ci permette di generare centinaia di immagini di persone in pochi minuti. In verità, l'utilizzo della tecnologia è esattamente ciò che ha permesso alla specie umana di differenziarsi e distinguersi dalle alterità abitanti il pianeta terra e l'ha allontanata da esse, favorendo la formazione del pensiero umanista e antropocentrico.

Margaret Mead, antropologa e insegnante statunitense, si è sentita domandare in una sua lezione quale fosse il primo segno di civiltà umana, e quindi di *cultura*. Probabilmente dopo aver posto il quesito, il giovane studente si aspettava una menzione a qualche arma primitiva come asce, lance o sassi aguzzi per la caccia, o ancora a elementi di prima sopravvivenza come pentole o contenitori. Niente di tutto ciò. Il primo segno di civiltà è infatti «un femore umano con una frattura guarita, scavato e scoperto in un sito di quindicimila anni fa» (Byock 2012: 422). L'individuo, per sopravvivere a una frattura simile nei mesi necessari alla guarigione, è stato protetto e accudito grazie alla presenza di suoi simili. Nonostante anche altri animali si prendano cura dei feriti, nessuna specie è in grado di dedicare così tanto tempo ed energia alla causa. Un umano può essere quindi assistito tramite l'applicazione della tecnica: non è forse questa la prima e vera espressione del concetto di tecnologia? Il voler, e poter, applicare una capacità, utilizzando uno strumento per uno scopo superiore, non istintivo, non impulsivo.



↑ FIG. 1.1 FEMORE ROTTO E GUARITO

1.2 IBRIDAZIONE

Ormai tutti parlano di tecnologia e la cadenza di informazioni a cui siamo esposti a riguardo è giornaliera. Fra social media, televisione, giornali e pubblicazioni l'opinione delle masse tende a polarizzarsi, soprattutto durante periodi di rivoluzione: da un lato si prova paura, mentre dall'altro un forte interesse, quasi morboso, verso l'unione della propria individualità con l'artificiale.

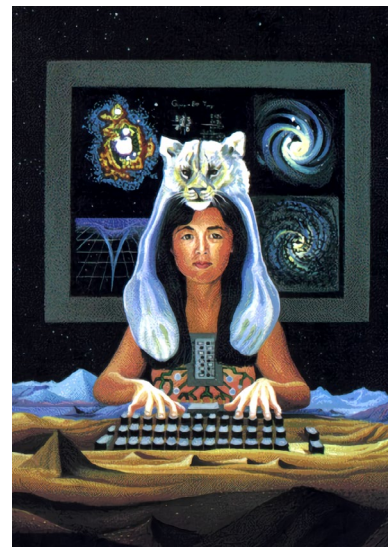
L'icona del *cyborg* è più che mai vera e presente oggi: formato dalla composizione dei due termini *cybernetic* e *organism*, il termine *cyborg* nasce negli anni Sessanta del Novecento, collegandosi inizialmente alle possibilità dell'uomo di viaggiare oltre l'atmosfera e raggiungere altri pianeti tramite la creazione di individui dalle funzioni biologiche potenziate, degli ibridi fra organismi e macchine (cfr. Giorgio 2016).

Il *cyborg* è diventato ben presto simbolo della letteratura fantascientifica, oltrepassando il recinto originario e assumendo più significati: è sia una creatura in cui il vivente e l'artificiale confluiscono, ma anche una rappresentazione della relazione stessa fra l'organico e il macchinico (cfr. Yehya 2017).

Forse però il trattato più influente in materia è stato quello di Donna Haraway, pioniera del concetto di *cyborg* come forma ibrida, non solo dal punto di vista scientifico ma soprattutto di una rappresentazione dell'essere *oltre* umano, *postumano* «Le persone non sono affatto fluide, dato che sono materiali e opache. I *cyborg* invece sono etere, quintessenza» (Haraway 1985: 13). Capostipite del pensiero femminista, Haraway ha illustrato quasi ironicamente questa nuova creatura come un vero e proprio superamento dei limiti peculiari dell'uomo e delle dicotomie disgiuntive che caratterizzano il pensiero umanista (Treccani s.d.-a).

Le dicotomie aiutano a descrivere degli stati di fatto della realtà dove c'è una chiara coerenza logica, ma sono poco efficaci e, anzi, non aiutano a comprendere i fenomeni fino in fondo, ostacolando il progresso della conoscenza. I dualismi femmina/maschio, mente/corpo, natura/cultura, naturale/artificiale, materia/forma sono divisioni che impediscono di coniugare l'essere umano con l'ignoto e non permettono l'ampliarsi della conoscenza (Zingale 2020).

Il *cyborg*, descritto in *A Cyborg Manifesto* nel 1985, ha la necessità di evitare e anzi di scappare da ogni pensiero dicotomico. Egli preferisce e necessita avere un nuovo punto di vista, osservando una realtà infinitamente complessa e guardando oltre l'orizzonte sociale autoimposto dagli esseri umani.



Il *cyborg* è libero da ogni pensiero di dominazione legato a genere, razza e specie. Secondo Haraway gli umani, intesi dal punto di vista biologico, hanno cessato di esistere come organismi univoci e sono stati sostituiti da sistemi (di comunicazione) completamente denaturalizzati, diventando artefatti costruiti e non più vincolati al mondo della natura. Di conseguenza, essi sono intrappolati in uno scenario di "umanesimo postmoderno", dove la ricerca biologica è al servizio degli ideali umanisti di perfezionamento della specie tramite la sua modificazione, decostruzione e ricostruzione dei corpi (cfr. Haraway 1985).

Dal manifesto della Haraway è facile comprendere come il ventesimo secolo abbia visto la grande ascesa ma, successivamente, anche il parziale tramonto degli ideali di purezza del-

† FIG. 1.2 COPERTINA DI *A CYBORG MANIFESTO*, DONNA HARAWAY

IL CYBORG SONO ETERE E QUINTESSENZA

(HARAWAY 1985: 13)

IL CYBORG

la specie umana in tutte le sue sfaccettature tipiche della cultura occidentale, cercando contemporaneamente un riscontro nella scienza. L'uomo ha sempre provato a dimostrare la sua superiorità in modo tangibile tramite lo studio scientifico, relegando il diverso, l'altro, a un pericolo per la perfezionabilità della specie, sfociando molto spesso in ideali razzisti (cfr. Marchesini 2002: 172-173).

Per comprendere meglio ciò che ho illustrato in queste pagine è necessario però rendere chiaro il concetto di *alterità*, ovvero ciò che è o che si presenta come "altro", diverso, non-identico. L'alterità può essere un animale, un vegetale, un essere non vivente. È quindi per antonomasia l'opposto dell'*identità*, che invece è la coscienza di sé stessi "umani" come "io", come consapevolezza di quello di cui siamo costituiti e di quello che ci rappresenta. Aprirsi all'altro significa per sua natura aprirsi a un'altra identità, non conosciuta e differente dalla nostra, mettendo a confronto in una relazione paritaria entrambe le parti e aprendosi al dialogo con la diversità. Ciò comporta che ogni concetto sconosciuto all'io è definibile come alterità e si trova almeno inizialmente in un rapporto di opposizione (Zingale 2020).

Gli umani che possiedono un'*identità forte* hanno repulsione verso l'alterità e portano avanti un ideale di purezza a scapito delle differenze anche con altri individui della loro stessa specie, rifiutando ogni tipo di relazione e di confronto che non sia distruttivo o a scopo di dominio. Questo concetto è figlio del pensiero umanista, che cerca di raggiungere una realtà utopica di una società ideale anche tramite la rappresentazione figurativa, rifiutando ciò che non è conforme al progetto umano e perciò non puro.

Il binomio *kalòs kai agathòs* ripreso dagli umanisti – dalla Venere botticelliana alle architetture dell'Alberti – coniuga il corpo a questo ideale di perfezione: si ricercano le simmetrie strutturali, gli equilibri architettonici, le formule morfometriche per realizzare strutture armoniosamente iterate.

In nome dell'aspirazione alla rotondità geometrica è necessario epurare tutto ciò che resiste alla concezione essenzialista della forma perfetta, del movimento lineare, della simmetria tra passato e futuro. [...] Ciò che non rientra nella cristallina perfezione platonica diviene rumore, corruzione, caos a cui contrapporre l'armonia incontaminata del cosmo. (Marchesini 2002: 174)

Preso in esame questa volontà umanista di separare ancor di più le dicotomie, di cui abbiamo parlato prima grazie alla figura del cyborg, possiamo affermare che secondo questa

IBRIDAZIONE

corrente di pensiero, da cui il *Transumanesimo* raccoglie pienamente l'eredità, la tecnologia sia in opposizione al biologico. In realtà, la prima è più una manifestazione di quest'ultimo.

Il concetto di ibridazione transumanista insegue il voler raggiungere una perfezione della specie grazie all'unione con la macchina e al conseguente e voluto allontanamento dalla natura, scenario che risulta quasi distopico e contraddittorio se applicato con questi termini.

L'umano transumanista prova una *vergogna prometeica*, una «vergogna che si prova di fronte all'umiliante altezza di qualità degli oggetti fatti da noi stessi» (Anders [1956] 2022: 57). Egli si reputa inferiore e non più meritevole di valore rispetto agli strumenti tecnologici che lui stesso ha creato, si vergogna di essere troppo uomo dinanzi alla perfezione delle macchine, promuovendo l'unione con esse. Come accennato prima, il fine ultimo di questa convinzione è la perfettibilità: il raggiungimento di un limite a causa della convinzione di non appartenere a ordini naturali dati. Tuttavia:

Andare verso l'innovazione tecnologica non significa operare un allontanamento dalla natura ma, viceversa, indirizzare il proprio cammino nella direzione di una maggiore coniugazione con la natura. [...] Scopriamo che attraverso l'ibridazione culturale l'uomo costruisce la sua natura e si avvicina alla natura dell'alterità (Marchesini 2002: 187).

In conclusione, nei termini postumanisti l'innovazione tecnologica non viene intrapresa per sopperire alle mancanze dell'individuo-uomo e quindi per un allontanamento dall'alterità, ma vi è piuttosto una coniugazione e un dialogo aperto con quest'ultima.

1.3 ESSERE MOSTRO

La vera evoluzione postumanista rinuncia in primis a una visione di uomo omologato e di standardizzazione, abbracciando invece il rapporto con l'altro e con la multiformità.

Un umano multiforme viene definito da Massimo Centini (1999: 83) come *mostro* se è appartenente a una minoranza che attua una qualsiasi forma di trasgressione. La mostruosità può essere considerata uno scostamento dalla normalità, grazie a capacità straordinarie dell'individuo che lo rendono parte di una piccola fetta della popolazione, una nicchia. Il talento naturale, le capacità peculiari o gli interessi qualificano quindi l'individuo sotto la definizione di mostro. «Il mostro ha anche il potere di affascinare, di coinvolgere l'osservatore in una spirale senza ritorno che fa del contrasto tra norma e infrazione una forma di dimostrazione delle potenzialità dell'uomo» (Centini 1999: 83).

Egli è quindi qualcosa di inaspettato, qualcosa che si distacca dai canoni di perfezione e aiuta a proiettarci verso una dimensione di ibridazione reale con la tecnologia. Gli individui che cospargono il mondo con la loro unicità e la loro mostruosità vengono essi stessi contaminati e ibridati dai contenuti che offrono, come in un circolo vizioso.

Abbiamo visto che la mostruosità può manifestarsi sia grazie alle caratteristiche fisiche ed estetiche del singolo, ma, estendendo la definizione, ritengo che possa mostrarsi anche grazie a ciò che egli produce. L'immagine figurativa umanistica, perfetta e seriale, lascia il posto a rappresentazioni eterogenee e imperfette, unioni di più connotati anomali che si distaccano dalla matrice. Come nell'arte contemporanea si capovolge il concetto di bellezza includendo il *mostruoso*, anche nella creazione di contenuti online la figura umana viene distorta, replicata e soprattutto condivisa e dispersa in rete, favorendo la contami-

nazione tra idee. Il mostro contribuisce, positivamente o negativamente, al progredire della società umana tramite la produzione di contenuti che ci plasmano. Vedremo in 4.3 che è già oggi possibile osservare community di persone, appartenenti a quello che Marchesini definisce *sottobosco* (Marchesini 2002: 188), impegnate a generare contenuti eticamente scorretti o a scopo disinformativo per soddisfare un duplice desiderio: dal lato dei creatori è il senso appartenenza a una mutazione culturale, a un movimento underground che favorisce l'utilizzo sperimentale delle nuove tecnologie; dal lato dei fruitori, anche qui, è il gusto del proibito e del diverso ad attrarre nella visione del contenuto stesso, che sia esso di carattere politico, complottista, pornografico o minatorio.

Il mostro agisce nell'ombra e porta contaminazione, ma soprattutto vuole stare nel sottobosco, non per rifugiarsi dal giudizio della società ma per un'attrazione viscerale verso scenario underground, per un rifiuto dell'*abito culturale* usuale a favore di una fluidità interna rintracciabile nel cyborg. Egli non è più puro, ma contaminato e contaminante. È l'attore protagonista in uno scenario di continua ibridazione e abbraccia la tecnologia, che diventa quasi un'espressione dell'insufficienza dell'essere umano (Marchesini 2002: 191).

Non dobbiamo però commettere l'errore di ritenere questa figura come lontana da noi. Chiunque oggi ha una propria specifica nicchia di interessi e una verticalità. Senza scomodare scenari o gruppi considerati dai più come spregevoli, anche la forte passione verso un'attività o un hobby e il senso di appartenenza a una sottocultura ci classifica in qualche modo sotto l'ombrello di questa definizione. Ognuno di noi ha il proprio sottobosco di cui fa parte e per cui si muove.



↑ FIG. 1.3 DIANE ARBUS, *MASKED CHILD WITH A DOLL*, N.Y.C., 1961



**CON L'IA
SIAMO TUTTI**

02

**CREATORI
DI CONTENUTI**

2.1

GENESI DELL'INTELLIGENZA ARTIFICIALE

2.2

AARON E LA PRIMA ARTE GENERATA

2.3

PANORAMA ATTUALE

2.4

IL RAPPORTO CON LA NUOVA SPECIE

IA, *Intelligenza Artificiale*. Una delle più recenti tecnologie informatiche con cui stiamo iniziando, nuovamente, il nostro processo di ibridazione. È indubbio che entro pochi anni questa avrà un impatto estremamente rilevante sulle nostre vite e sul nostro essere postumani.

Come individui abitanti il *villaggio globale* (McLuhan e Fiore 1967) del mondo online, siamo già tutti creatori di contenuti, nessuno escluso. Ogni giorno produciamo video, modifichiamo fotografie, registriamo audio e li condividiamo ad amici, familiari ma anche a sconosciuti in luoghi digitali privati e pubblici, alla mercé dell'altro. Nonostante il facile accesso ai più disparati strumenti, non tutti possiedono la cultura del progetto e non tutti sono progettisti.

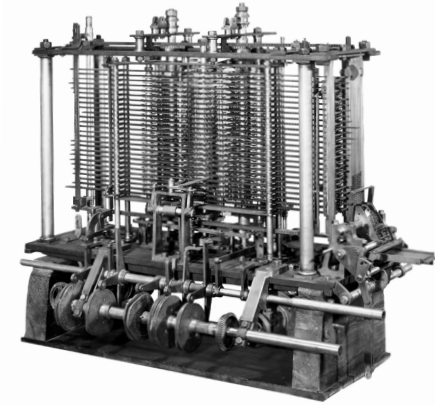
La generazione di *media sintetici* da parte di individui e community senza scrupoli o senza alfabetizzazione progettuale potrebbe condurre a uno scenario di continua e imperterrita contaminazione tale da raggiungere una soglia limite considerabile come pericolosa e confusionaria (4.2). Ritengo però che, ancor prima di affrontare questo tema, sia necessario osservare lo stato dell'arte attuale e le potenzialità che questa tecnologia fornisce ora all'essere umano, per poi effettuare un'astrazione verso scenari futuribili.

2.1 GENESI DELL'INTELLIGENZA ARTIFICIALE

Il computer è il nostro fidato compagno di progettazione, un *doppio digitale* che ci assiste in ciò che prima potevamo realizzare solamente a mano e che ha accelerato notevolmente il processo di progettazione di artefatti visivi e non. È anch'esso una tecnologia: al nostro servizio e comandato da noi per uno scopo ultimo. D'altronde, il mio lavoro come quello di migliaia di altri designer e progettisti non sarebbe stato possibile senza questo "portale sul mondo virtuale", uno strumento che grazie alla sua facilità d'uso entra nella relazione progettista-progetto in modo quasi trasparente (4.4), aprendo la mente verso scenari altrimenti irraggiungibili.

Similmente alla tecnologia, saremmo tentati di considerare il concetto di intelligenza artificiale come relativamente recente, data la sua presenza massiccia nella cronaca in conseguenza alla democratizzazione di strumenti per la generazione di elementi multimediali. Il grande pubblico può accedere solamente ora a questo assaggio di intelligenza macchinica, ma è bene considerare che la figura dell'automa, un cervello digitale al servizio dell'umano, è materia di studio da decenni ed è stata predetta sotto le spoglie di *Macchina Analitica* ben più di un secolo fa da Lady Ada Lovelace e Charles Babbage (v. fig. 2.1). Il desiderio di uno strumento in grado di processare simboli che

rappresentino "tutti i soggetti dell'universo" (Boden 2018: 7) inizia a sorgere nel mondo scientifico: prima con Lovelace, poi con Turing e il suo *test* composto da una serie di domande chiave sull'elaborazione di informazioni, volte a decretare se una macchina possa essere considerata intelligente: vero e proprio manifesto per il settore (*ibid.*).



L'intelligenza artificiale nasce come disciplina accademica con lo scopo di far operare i computer nello stesso modo in cui opera la mente umana, conferendogli concetti come il ragionamento, la visione, l'associazione, la pianificazione, e più in generale abilità psicologiche per raggiungere un obiettivo. L'IA si serve di tecniche per risolvere numerosi problemi nella vita quotidiana e non. Ha un duplice obiettivo: svolgere compiti utili tramite procedimenti differenti o nuovi rispetto a quelli finora utilizzati dalle menti umane e rispondere a domande aiutando l'uomo nello sviluppo di teorie. La sua diffusione è capillare e spazia in numerosi settori, dall'automobilistico al medico, dal videoludico al finanziario e militare, fino ad arrivare alla fotocamera del nostro smartphone. I motori di ricerca sfruttano l'IA, l'industria cinematografica sfrutta l'IA, le gallerie d'arte sfruttano l'IA (cfr. Boden 2018).

Nonostante scienziati e informatici abbiano svolto un lavoro non indifferente dalla sua genesi e nei decenni successivi, il vero progresso delle IA come le conosciamo oggi ha avuto inizio solamente dal nuovo millennio. La potenza hardware richiesta inizialmente, infatti, era estremamente elevata e le evoluzioni in ambito informatico fino agli anni Ottanta non sono state sufficienti, tanto da arrivare a bloccarne lo sviluppo a livello internazionale riducendo i fondi a supporto della causa: ciò è stato definito come *AI Winter* (McCorduck [2003] 2018: 462).

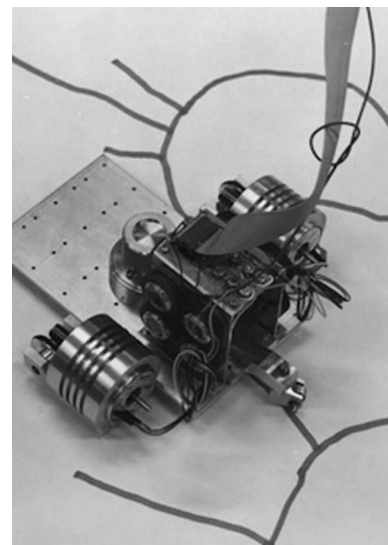
† FIG. 2.1 *MACCHINA ANALITICA*, CHARLES BABBAGE E LADY ADA LOVELACE, 1834

2.2 AARON E LA PRIMA ARTE GENERATA

Osservando lo scenario delle arti visive e figurative, possiamo considerare *AARON* di Harold Cohen come un primo esempio di strumento guidato da un'intelligenza artificiale per realizzare arte generativa.

Garcia (2016) ci racconta come la sua prima iterazione risalga al 1972 ma il supporto al programma è continuato fino al 2010, rendendolo così uno dei progetti di intelligenza artificiale più longevi della storia. Harold Cohen, il progettista dietro al codice di *AARON*, è stato un artista e studente di belle arti a Londra. Nonostante una certa fama nel settore, il suo interesse ben presto è diventato il processo di creazione dietro all'opera d'arte. Ciò lo ha portato a uno studio approfondito delle componenti base della rappresentazione visiva e artistica. Forme, immagini e colori sono state analizzate cercando di sviscerare il metodo con cui un artista può realizzare un output. Grazie al suo background e allo spiccato interesse informatico sviluppatosi in seguito, Harold ha iniziato a scrivere programmi per replicare la pittura. *AARON* inizialmente si è occupato solamente di aspetti interni alla cognizione dell'individuo e all'identificazione di forme primitive funzionali del pensiero (interni ed esterni, figura e terreno) con l'obiettivo di costruire un modello limitato della mente umana applicato all'arte (v. fig. 2.2).

Ha quindi imparato a produrre output in bianco e nero, completati poi con i colori dalla mano dell'artista. Ben presto ha però sostituito il creatore anche in questo frangente. Inizialmente generando forme e illustrazioni comparabili a disegni infantili, il programma si è evoluto nella rappresentazione di figure biomorfe, oggetti, animali e persone, in un processo *trial and error* in cui l'artista ha addestrato la macchina a dar vita a risultati in linea con i suoi obiettivi di ricerca: generare opere indistinguibili da quelle che avrebbe potuto produrre lui (v. fig. 2.3 e 2.4). *AARON* è stato in mostra per numeroso tempo in musei ed esposizioni grazie all'imponente dimensione del sistema, composto da braccia robotiche, superfici e colori, e la complessità del codice richiesto per assicurare il suo funzionamento (Garcia 2016).



Questo archetipo di IA applicata alla produzione di output figurativi porta con sé numerosi e naturali quesiti, che ancora oggi riecheggiano con la diffusione di strumenti alla portata delle masse. *AARON* infatti ne è il precursore concettuale: esso si basa su un *training* messo in atto dal suo creatore, che, dando in pasto alla macchina segni e simboli, pian piano ha addestrato il suo figlio virtuale a imitare la realtà. La sua visione della realtà era primitiva e limitata, ma apprendendo dai riferimenti, che potrei definire come veri e propri dataset, ha iniziato ad accumulare una conoscenza, riuscendo poi a produrre degli output tangibili grazie a dei supporti fisici che hanno ampliato le sue capacità di azione. Forse per la prima volta nel mondo delle arti figurative chi chiediamo se una macchina, *AARON*,

↑ FIG. 2.2 PRIMA ITERAZIONE DI AARON, 1979



↑ FIG. 2.3 AARON E HAROLD COHEN

CON L'IA SIAMO TUTTI CREATORI DI CONTENUTI



↑ FIG. 2.4 AARON NELLA FASE DI COLORAZIONE DELL'OUTPUT

AARON E LA PRIMA ARTE GENERATA

possieda realmente creatività e sia effettivamente equiparabile a un artista, oppure attui solamente una riorganizzazione di informazioni provenienti dal suo mentore (3.7). Ma soprattutto ci chiediamo se questo rappresenti un esempio di computer intelligente (Garcia 2016).

Per tutto questo tempo, Cohen ha sempre considerato AARON alla pari di un suo assistente fidato. Riprendendo tecnica e tecnologia, AARON è come una terza mano in grado di trasporre su tela la visione dell'artista, come uno strumento avanzato al servizio dell'uomo. Questa mano però nel tempo è diventata sempre più decisiva per l'ottenimento di un output. Inizialmente il computer ha avuto bisogno di una costante assistenza da parte del progettista che si è ritrovato a completare con colori e correzioni i disegni. Con il continuo aggiornamento delle sue capacità, però, AARON è diventato così indipendente da riuscire a ribaltare il rapporto progettista-assistente, relegando quest'ultimo quasi unicamente allo spazio della programmazione, ma mantenendo comunque una relazione di coniugazione fra le due identità.



† FIG. 2.5 HAROLD COHEN, *AARON WITH DECORATIVE PANEL*, 1992

2.3 PANORAMA ATTUALE

Con la ripresa della ricerca dopo l'AI Winter, il desiderio di sviluppare delle intelligenze artificiali seguendo la struttura della mente umana ha portato ben presto alla diffusione del modello computazionale del *neural network*, una rete neurale artificiale basata sul *deep learning*, ovvero una forma di apprendimento automatico della macchina. Un cervello artificiale, formato da neuroni e collegamenti fra essi in grado di imparare, elaborando dei dati forniti ed eseguendo compiti, prendendo poi decisioni autonomamente. Una delle prime applicazioni di questo modello è il riconoscimento facciale: alla rete neurale viene dato in pasto un dataset composto da migliaia di volti umani differenti: allenandosi ripetutamente il network impara poi a riconoscerli con estrema precisione (cfr. Schick 2020).

Una caratteristica peculiare del cervello umano è quella di analizzare ciò che gli si presenta davanti, scovare pattern e trovare soluzioni. Oltre a ciò, noi umani siamo esseri dotati di inventiva, siamo progettisti naturali in grado di scrivere poesie, dipingere opere, comporre artefatti. Allo stato attuale, gli strumenti di IA generativa hanno raggiunto una qualità tale da poter effettuare operazioni apparentemente simili ed essere scambiate potenzialmente per noi (cfr. Sequoia 2022). Scopriremo però, in 3.7, una sostanziale differenza fra questi due mondi.

Gli esperti hanno posizioni drasticamente opposte verso il fenomeno e non aiutano la già difficile comprensione dello scenario evolutivo di questa tecnologia, estremamente mutevole e capillare. Un fruitore attento non può non aver notato un cambiamento radicale nelle possibilità di creazione di contenuti e più in generale di come questa rivoluzione venga comunicata online. Da assiduo fruitore di video su TikTok, ho osservato una trasformazione netta di ciò che ci viene proposto ogni

giorno, con un forte scostamento tematico verso la divulgazione di nuovi sistemi di IA. È come se ogni utente volesse contribuire a una propaganda tecnico-comunicativa, e questa propaganda fosse spinta dalle piattaforme stesse. Nel giro di pochi mesi siamo passati dall'utilizzare strumenti capaci di ingrandire fotografie di qualche migliaio pixel ad avere a portata di click decine di programmi in grado di comprendere il linguaggio umano, scrivere, interagire con noi e generare immagini solamente grazie a un nostro input testuale. Forse proprio quest'ultima applicazione è stata l'elemento scatenante che ha permesso la viralità del fenomeno: cosa c'è di più comprensibile e condivisibile di un'immagine? «Sure enough, as the models get bigger and bigger, they begin to deliver human-level, and then superhuman results» (Sequoia 2022).

È facile capire come questa possa essere considerata l'alba della *seconda età delle macchine* (Dzyaloshinsky 2022), di un'innovazione non più incrementale ma radicale, alimentata dal costante aumento della velocità di calcolo e dalla riduzione del costo degli strumenti.

Per questo, nonostante sia interessante dare un'infarinatura dello scenario attuale del mondo delle IA generative, sono consapevole che tracciare uno schema assoluto e future proof di tutto ciò che possiamo fare noi umani in sinergia con queste è pressoché impossibile. Ciò non impedisce però un'analisi dello stato di fatto e delle riflessioni su possibili proiezioni future a breve e medio termine.

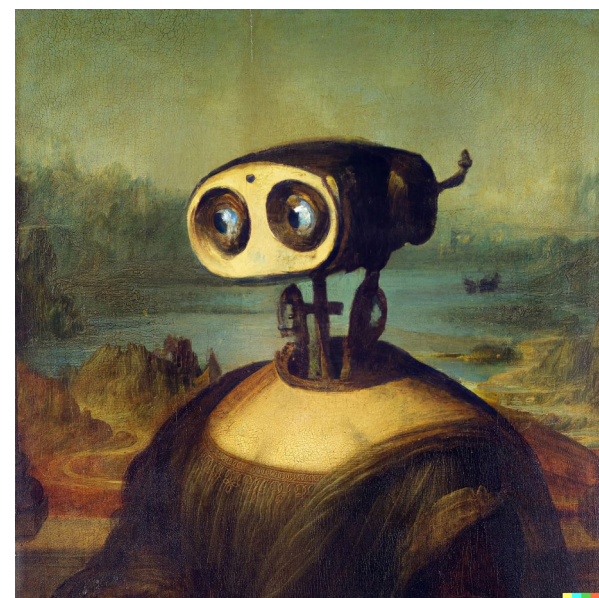
Questa "febbre dell'oro" è merito in primo luogo dell'estrema facilità di approccio con lo strumento, con interfacce *user friendly* e non più accessibili solamente a programmatori o appassionati di informatica. In secondo luogo, il fatto stesso che si tratti di servizi in rete e molto spesso open source rende il bacino d'utenza immensamente più ampio, spesso eliminando la necessità di installare il programma sul proprio hardware. Pur essendo molti strumenti ancora in versione sperimentale e quindi soggetti a repentini cambiamenti, chiusure o aggiornamenti, abbiamo già da ora un assaggio delle potenzialità di questa rivoluzione.

I provider online che ci mettono a disposizione modelli sono molteplici; tra i più celebri *OpenAI*, società di "ricerca e sviluppo dell'intelligenza artificiale per il beneficio dell'umanità" («OpenAI» 2021). Molto spesso nel corso del 2022 quest'azienda infatti ha spopolato nelle discussioni online. Oltre ad aver lanciato *DALL·E* e più recentemente *DALL·E 2*, sistemi *text-to-image* in grado di realizzare immagini a partire da una descrizione testuale in linguaggio naturale, OpenAI ha introdotto anche il modello alla base di *Chat-GPT*, ormai arrivato alla sua quarta iterazione, in grado di comprendere il linguaggio umano e di produrre conseguentemente una risposta convincente ed estrema-

THEY BEGIN
TO DELIVER...
SUPERHUMAN
RESULTS

(SEQUOIA 2022)

mente esaustiva sotto forma di testo. Quando anche un neofita può dilettersi e perdersi nella creazione di output pressoché perfetti, capiamo che il futuro è già qui: si aprono ora scenari in cui non è più necessario avere competenze tecniche verticali nell'editing foto, nella registrazione e nel montaggio video o nel copywriting, ma servirà il saper relazionarsi con queste alterità non-umane che comprendono il nostro linguaggio e dialogano con noi come farebbe un nostro simile.



↑ FIG. 2.6 ESEMPIO DI IMMAGINE GENERATA CON DALL-E 2 A PARTIRE DA UN INPUT TESTUALE

↑ FIG. 2.7 ESEMPIO DELLA FUNZIONE *INPAINTING* DI DALL-E 2, CHE PERMETTE DI SOSTITUIRE CONNOTATI DI IMMAGINI GIÀ ESISTENTI

2.4 IL RAPPORTO CON LA NUOVA SPECIE

È interessante osservare e analizzare il duplice confronto che ha l'essere umano con questi strumenti: sia dal punto di vista del creatore del modello, sia dalla parte del fruitore del servizio e di come si relaziona con esso per ottenere dei risultati.

Come già accennato prima, l'avvento delle intelligenze artificiali non è un tema nuovo nella storia. Parafrasando Marchesini (2002), l'uomo, fantasticando la creazione di grandi strumenti di intelligenza meccanica in un processo definibile come *meccanopoiesi*, ha sempre avuto il desiderio di proiettare sé stesso su una macchina. L'automa è in grado di svolgere i compiti al suo posto e permettergli finalmente di slegarsi dall'impegno ripetitivo e dall'attività manuale, quasi come uno schiavo che non ha bisogno di particolari attenzioni, non avendo necessità di sopravvivenza o coscienza alcuna, ma che anzi risponde solamente ai comandi del suo creatore. «L'automa è una marionetta senza fili, un organismo senza organi, un essere animato senza i bisogni e le finalità propri dell'essere vivente» (Marchesini 2002: 293).

La meccanopoiesi fa sognare l'uomo, che diventa il burattinaio, colui che governa la macchina per i propri scopi. Tuttavia, con l'irrompente informatizzazione del mondo, egli comprende che deve anche imparare a sviluppare uno specifico registro

di comportamento da applicare nella sua relazione con l'essere virtuale. La macchina, evolvendosi, diventa un vero e proprio compagno di lavoro in sinergia con l'umano. Il senso di familiarità e di rapporto paritario si instaura grazie all'applicazione di interfacce e di modelli di linguaggio che facilitano l'approccio con essa (*ibid.*). Nonostante ciò, la vergogna prometeica fa sempre capolino, rischiando di far passare l'uomo da uno stato di supremazia a uno di subordinazione all'alterità digitale: egli prova paura o rassegnazione e ipotizza di scenari in cui l'intelligenza artificiale lo sostituisce in tutto e per tutto.

Restiamo però nel presente e nell'ambito di creazione di contenuti. Come utenti, ci sentiamo sbalorditi e attoniti quando parliamo per la prima volta con la nuova specie. Le capacità computazionali e di ragionamento della macchina stupiscono l'intelletto: un essere artificiale diverso da noi ma che si comporta come noi, con cui possiamo dialogare in modo semplice componendo frasi di senso compiuto.

L'incanto della tecnologia, o enchantments of technology (McCarthy e Wright 2004), torna nelle nostre vite. È quell'emozione che proviamo nell'essere parte di un nuovo capitolo del progresso e di non conoscerne fino in fondo le potenzialità. L'abbiamo sperimentata con l'introduzione del computer e di internet, con le e-mail, con gli smartphone. Tutte queste manifestazioni di tecnologia talvolta spaventano, ma molto spesso stupiscono e possono causare l'incanto. Ogni qualvolta uno di questi strumenti diventa di uso comune però, tendiamo a perdere quella magia che ne caratterizza il suo utilizzo, diventando dopo poco tempo routine.

Come affermano McCarthy e Wright (2004: 192) infatti, l'incanto avviene in quanto siamo noi stessi a influenzare la forma e il significato della tecnologia che ci circonda attraverso il nostro dialogo con essa. Di fronte all'IA generativa, considerabile come una delle più recenti declinazioni di tecnologia e un'alterità non-umana, ci troviamo in un rapporto di relazione bidirezionale e interattiva: evolviamo con essa ed essa evolve grazie a noi, in un continuo scambio di informazioni. Questo significa che come umani dobbiamo essere pronti ad accettare il fatto che tutto ciò che ci circonda possa essere influenzato dalla presenza della tecnologia, ed è questo che la rende *incantevole*.

Il processo di conoscenza dell'alterità non-umana inizia grazie a un dialogo ripetuto (Zingale 2020). Il luogo in cui si instaura questo dialogo è potenzialmente ovunque e il tempo è sempre; essendo la maggior parte dei servizi online basta possedere un dispositivo in grado di accedervi.

La creazione di *user experience* e di interfacce per questi servizi da parte degli sviluppatori diventa cruciale. In un processo di ricerca della sintonia uomo-macchina, si vuole assecondare e anzi prevedere le esigenze degli individui e di potenziare

l'esperienza tramite la creazione di relazioni e interazioni basate molto spesso sul linguaggio umano. Considerando una persona mediamente istruita, il linguaggio è forse ciò che più di tutto il resto facilita il rapporto con l'*altro*. Allo stato attuale è già possibile interfacciarsi con ChatGPT e DALL·E 2 in circa un centinaio di lingue diverse, intercambiabili in qualsiasi momento durante la conversazione semplicemente scrivendo una frase.

Il linguaggio è un sistema di significazione: ci permette di attribuire un significato a ciò che come umani vediamo, viviamo e proviamo. Possiamo quindi considerarlo come una *griglia* attraverso cui suddividiamo e vediamo il mondo, che interpretiamo grazie alla nostra lingua e alla nostra cultura. Per questo, in base alla cultura d'appartenenza, il significato attribuito a una parola può variare. Con ChatGPT ad esempio, possiamo instaurare un dialogo di valore indipendentemente dalle nostre radici culturali, ottenendo le informazioni in modo naturale (Zingale 2020). «The old computing was about what computers could do; the new computing is about what users can do» (Shneiderman 2003: 2)

Su modelli text-to-image come *DALL·E 2*, *Midjourney* o *Stable Diffusion* (3.6), instauriamo ancora una *negoziazione tacita*, un dialogo fatto di parole enunciate in modo unico che si fonda sul linguaggio a noi conosciuto ma ha delle caratteristiche leggermente differenti. Semplifichiamo infatti la sintassi, utilizzando simboli, parole e codici peculiari (cfr. Finn 2017: 30) come nell'inserimento di *prompt*, ovvero parole di input, che molto spesso hanno più l'aspetto di termini chiave in rapida successione. Nonostante ciò, il ritmo del progresso è impressionante. Il rapporto sempre più umano che si ha con la macchina semplifica notevolmente la relazione, rendendo l'IA un fidato assistente con cui comunicare senza soluzione di continuità e il cui processo creativo è però invisibile ai nostri occhi. Le potenzialità dell'utente sono amplificate dallo strumento utilizzato e bypassano le *barriere all'ingresso* tipiche dello studio di una specifica tecnica, democratizzando la produzione di contenuti.

orsacchiotti sulla luna che lavorano a nuove ricerche sull'intelligenza artificiale negli anni '80

Generate

teddy bears working on new AI research on the moon in the 1980s

Generate

osos de peluche trabajando en nuevas investigaciones sobre IA en la luna en los años 80

Generate

泰迪熊在20世纪80年代在月球上从事新的人工智能研究

Generate

плюшевые медведи, работающие над новыми исследованиями ИИ на Луне в 1980-х годах

Generate

↑ FIG. 2.8 PROMPT TESTUALE UTILIZZATO SU DALL·E 2 PER GENERARE LA FIGURA 2.6, DECLINATO IN DIFFERENTI LINGUE

LA MARIONETTA

03

**COME L'UMANO
SI REPLICA ONLINE**

3.1

STORIA DELLA MANIPOLAZIONE

3.2

CHEAPFAKE, RADICI DEL FENOMENO

3.3

DEEPFAKE E MEDIA SINTETICI

3.4

FINTE CELEBRITÀ

3.5

SINTESI VOCALE E CLONAZIONE

3.6

MODELLO DIFFUSION, DAL TESTO ALL'IMMAGINE

3.7

MARIONETTE E RESURREZIONE

Preso coscienza delle fondamenta che sostengono il concetto di intelligenza artificiale, specialmente in ambito generativo, e della grande facilità di approccio agli strumenti oggi disponibili, vorrei addentrarmi in un argomento potenzialmente inquietante ma che suscita in me anche una certa attrazione. Attrazione che la nostra specie ha sempre nutrito nel corso della storia: *replicare* la figura umana è stato forse uno dei primi desideri della nostra civiltà. Dalle pitture rupestri fino ai giorni nostri abbiamo inseguito e attuato una rappresentazione a livello figurativo o materiale più o meno fedele del nostro corpo, collocandolo in contesti realistici, ipotizzando scenari fantastici o rendendo questo duplicato utile a uno scopo.

L'ambizione di voler generare di individui artificiali che Marchesini identifica come *figli dell'uomo* (Marchesini 2002: 281) esiste per molteplici ragioni celate o meno nella nostra indole, come il voler vincere la morte, il voler, da artisti, dar vita a una propria opera o il desiderio di riportare in vita persone defunte. Ciò si traduce in uno scenario caratterizzato da svariati sentimenti talvolta contrastanti. Come già affermato in 1.3, osservo come questo sogno tipicamente umano si possa riscontrare non solo nell'ambito della manipolazione biologica o della robotica, ma anche in quello della rappresentazione di una nostra copia in contesti virtuali, nella digitalizzazione della figura umana da parte nostra o da parte di terzi.

Negli ultimi anni è spopolato nel panorama dei contenuti online il concetto di *deepfake*, solitamente declinato in video nei quali si mostrano volti di persone realmente esistenti, sorprendentemente realistici, che svolgono determinate azioni o pronunciano determinate parole. Queste azioni sono generate con l'utilizzo di strumenti avanzati basati su intelligenza artificiale che impiegano il deep learning (Paris e Donovan 2019). L'introduzione di questi modelli ha sin da subito aperto la strada alla loro applicazione per la manomissione dell'informazione tramite l'espedito della raffigurazione umana, che oggi osserviamo quotidianamente online ma le cui origini risalgono a prima della nascita di internet. Le radici si spingono, quindi, in profondità e sono spesso legate alla nascita di nuovi media, come la fotografia e la carta stampata.

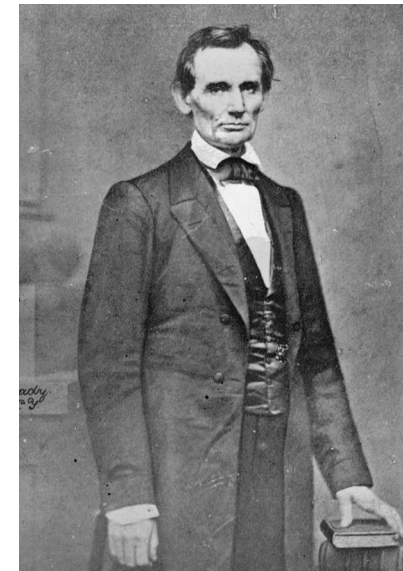
3.1 STORIA DELLA MANIPOLAZIONE

Non si può parlare di strumenti per la generazione di repliche umane digitali senza partire dalla genesi del concetto di *manipolazione*, cresciuto di pari passo con la nascita e l'evoluzione della fotografia. Dall'inizio del Diciannovesimo secolo, infatti, questo strumento permette all'uomo di catturare la realtà in modo accurato per la prima volta tramite un processo chimico. Non vi è più la mano del pittore che osserva la persona in posa per il ritratto e la imprime sulla tela manualmente, molto spesso rielaborandola su richiesta o a suo gusto ed esaltando tratti somatici o ricercando simmetrie. La fotografia infatti nasce con lo scopo di rappresentare nel modo più fedele possibile la realtà per come ci si pone dinanzi, relegando il fotografo all'accettazione del risultato del proprio lavoro per come è e come si mostra ai suoi occhi.

Nonostante ciò, come accaduto nella rappresentazione figurativa e come accadrà sempre nel corso della storia, l'impiego di uno strumento da parte di un creatore di contenuti, che sia esso progettista o meno, non è confinato al mero utilizzo dello stesso per degli scopi prestabiliti dalle sue qualità materiali, ma si espande alla ricerca dei limiti imposti e al superamento degli stessi. Anche con la fotografia infatti, ci si rende ben presto conto della possibilità di manomettere gli output prodotti.

Un caso eclatante, forse il primo rilevante nella storia, è quello del sedicesimo presidente degli Stati Uniti Abraham Lincoln. Come afferma Michael Waters (2017), durante la campagna del futuro presidente per la corsa alla Casa Bianca nel 1860, la diffusione della fotografia non era ancora lontanamente capillare e la popolazione delle zone più rurali del paese non conosceva il suo aspetto fisico, tanto da far circolare voce fra l'opposizione repubblicana riguardo il suo presunto brutto aspetto. Si era affermato addirittura un vero e proprio coro di protesta al grido di “we beg and pray you— Don't, for God's sake, show his picture”, ovvero “vi preghiamo e vi scongiuriamo: non mostrate, per l'amor di Dio, una sua foto”.

Per evitare ripercussioni sulla campagna politica, Lincoln ha messo a tacere queste dicerie rivolgendosi a Mathew Brady, un fotografo di Washington D.C. noto per la propensione al ritocco fotografico. Il futuro presidente è stato rappresentato in piedi con lo sguardo verso l'osservatore (v. fig. 3.1), e la sua figura è stata ritoccata sia nella realtà, puntando una forte luce sul suo volto e distogliendo l'attenzione dalla sua corporatura. È stata effettuata una vera e propria post-produzione che ha compreso, per esempio, l'allargamento del colletto della camicia per accorciare percettivamente il collo, altro aspetto fortemente deriso dagli oppositori.



† FIG. 3.1 FOTOGRAFIA DI ABRAHAM LINCOLN SCATTATA DA MATHEW BRADY POCO PRIMA DEL DISCORSO ALLA COOPER UNION, 1860



Una volta affissa su tutti i giornali, la fotografia ha subito ulteriori ritocchi da altri fotografi, che ne hanno aggiustato capelli e lineamenti in un vero e proprio circolo vizioso di manomissione del contenuto.

Tuttavia, il caso più eclatante riguarda un artefatto diffuso dopo il suo assassinio, quando vi è stata necessità di avere una sua rappresentazione in “stile eroico”.

Sempre Waters (2017) ci svela che il manipolatore questa volta è stato Thomas Hicks. Egli, sovrapponendo il volto di Lincoln a un’incisione di John C. Calhoun, è stato infatti responsabile di uno dei primi fotomontaggi dal carattere misinformativo della storia. L’immagine sottostante apparteneva a un politico repubblicano noto per le sue ideologie fortemente razziste e quindi ironicamente in opposizione ai principi dell’ex presidente (v. fig. 3.2). Sui fogli della scrivania, nell’incisione originale, sono infatti presenti parole a sostegno di questo pensiero, opportunamente sostituite da Hicks con frasi più consone al messaggio del nuovo artefatto, passato inosservato fino a metà del Novecento. L’unico dettaglio che ha tradito Hicks è stato il neo sul volto del presidente, dal lato sbagliato: la sua testa è stata

↑ FIG. 3.2 DETTAGLIO DELL’INCISIONE ORIGINALE DI JOHN C. CALHOUN



infatti ritagliata direttamente da quella stampata sulle banconote dell’epoca, scattata in origine proprio da Brady e riflessa poi specularmente ai fini del realismo della scena (v. fig. 3.3).

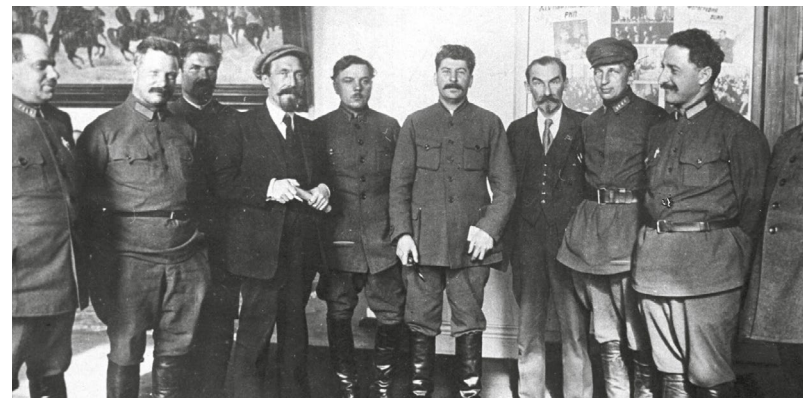
Si tratta quindi di un antenato di un processo definibile come *fotomontaggio*, oggi sviluppato al computer ma che ai suoi albori ha richiesto uno sforzo non indifferente per raggiungere una credibilità ed essere distribuito su larga scala.

Se ci spostiamo nel tempo e nello spazio, un altro caso anch’esso legato al mondo della politica è quello di Joseph Stalin. Nina Schick (2020: 19) ci mostra come la sua figura sia associata al mondo della manipolazione e il suo regime abbia attuato più e più volte campagne di falsificazione di documenti storici quali, appunto, fotografie. L’abilità tecnica nel giustapporre frammenti di negativi fotografici ritagliati e riposizionati va ora a braccetto con un’impostazione su scala industriale del concetto di manomissione: individui al servizio del governo che, di pari passo con gli sviluppi politici, agiscono modificando output visivi raffiguranti il leader sovietico, sia cambiando il suo aspetto, similmente a Lincoln, con aerografi per levigare i lineamenti; sia riscrivendo la storia di intere fotografie di gruppo ed

↑ FIG. 3.3 INCISIONE MANIPOLATA DA THOMAS HICKS

eliminando completamente alcuni oppositori politici del dittatore. Prendendo come esempio la figura 3.4 risalente all'aprile del 1925, Stalin si trova al centro di un gruppo di nove delegati durante il quattordicesimo "Congresso del Partito Comunista dell'Unione Sovietica". Osservando una versione dell'immagine del 1939, nella figura 3.5, possiamo constatare come questa abbia subito numerosi aggiustamenti, tanto che ora solamente i tre suoi amici più stretti godono della sua compagnia. È facile immaginare come gli altri uomini presenti, morti per suicidio, fucilazione o deportazione, siano stati man mano eliminati artificialmente in un avanguardistico tentativo di manipolazione della storia e di controllo della comunicazione.

Questi due esempi, seppur distanti dalla realtà odierna ma comunque interessanti dal punto di vista tecnico, aiutano a porre le basi per lo scenario che ci si pone dinanzi. Entrambi hanno come minimo comune denominatore la falsificazione di output fotografici raffiguranti esseri umani, la loro diffusione consapevole per scopi politici e propagandistici, ma soprattutto fungono da spunto per analizzare il ruolo del creatore, un architetto che compone e permette la divulgazione dell'informazione. La diffusione di artefatti comunicativi raffiguranti figure umane può infatti risultare estremamente persuasiva, efficace e pericolosa, se realizzata in modo sufficientemente convincente e rilasciata in uno scenario sociale di persone pronte ad ascoltare e spesso vulnerabili.



† FIG. 3.4 FOTOGRAFIA ORIGINALE DI JOSEPH STALIN AL CONGRESSO DEL PARTITO COMUNISTA, 1925

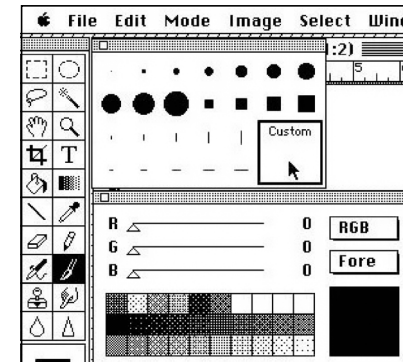
† FIG. 3.5 FOTOGRAFIA MANIPOLATA RISALENTE AL 1939

3.2 CHEAPFAKE. RADICI DEL FENOMENO

La falsificazione analogica ha ben presto lasciato spazio alla nascita di strumenti digitali sfruttabili grazie alla diffusione del personal computer. Nel 1990 è nato uno dei primi software di editing e manipolazione fotografica, *Photoshop*, che, nonostante il suo prezzo inizialmente inaccessibile, è stato utilizzato così frequentemente negli anni successivi da diventare un verbo (Kastrenakes 2020). “Photoshoppare” è oggi sinonimo di alterare digitalmente di un contenuto fotografico.

Nel nuovo millennio la manipolazione di fotografie è diventata popolare tra l'utenza comune grazie alla diffusione della pirateria *peer-to-peer* e della banda larga. Ciò ha portato a una grande circolazione del software Photoshop, che si stima sia stato scaricato illegalmente, insieme ad altri programmi, su circa il 40% dei PC (*ibid.*). Da qui sono nate anche le prime preoccupazioni: se una foto rappresenta la realtà, e questa realtà può essere alterata da tutti tramite uno strumento, come possiamo essere sicuri che quello che vediamo sia reale? Non vi è più l'artigiano che taglia, incolla e compone i negativi, ma vi è una nuova figura che si trova nello spazio digitale e può operare nel bene o nel male. L'abbattimento delle barriere all'ingresso, prima col computer e poi con lo smartphone, ha allargato ancora di più il bacino d'utenza delle applicazioni, ora anche gratuite.

È stato scoperto che, come umani, siamo più propensi a credere alla veridicità di un'informazione se questa ci “suona bene” o ci si presenta sotto forma di contenuto visivo in accompagnamento all'informazione stessa (Schick 2020: 22). Pur ritenendo innegabile che le persone più avvezze al mondo digitale e con un background progettuale abbiano gli strumenti per mettere in discussione la veridicità di un'immagine, non posso fare a meno di osservare che per il fruitore medio di contenuti online questo rappresenta ancora un grosso problema.



Il termine *cheapfake* indica tutta quella sfera di output multimediali che per essere prodotti richiedono l'utilizzo di strumenti di editing foto e video, ri-contestualizzando il contenuto e le informazioni al suo interno con il fine ultimo di trasmettere un messaggio differente dall'originale. Per esempio, è definibile come *cheapfake* un video tagliato, editato e pubblicato online lasciando solo parti funzionali al nuovo messaggio per influenzare l'opinione pubblica riguardo temi politici, culturali e sociali.

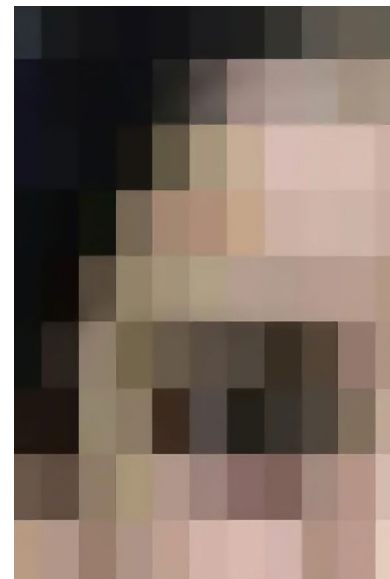
Questo scenario si riflette nel problema a monte di questa tesi e trova le sue radici in quel lasso temporale che va dalla democratizzazione di internet all'inizio degli anni Duemila fino a oggi. Il mondo dei social media accelera notevolmente la circolazione dei contenuti, e gli output fotografici possono essere manomessi per diffondere dei messaggi, convincere persone o soddisfare desideri. In quest'ultimo caso in particolare, e nello specifico nel mondo della pornografia, forse sono rintracciabili le conseguenze più gravi per il singolo individuo.

Il fenomeno è definito come *photoshopping sessualizzato* (McGlynn, Rackley, e Houghton 2017: 33) e comprende il più delle volte un *face swapping* che, similmente al fotomontaggio di Abraham Lincoln, viene attuato con la sostituzione del volto o di una parte del corpo di una persona sul corpo di un'altra persona donatrice senza il suo consenso. Il risultato è un ibrido vir-

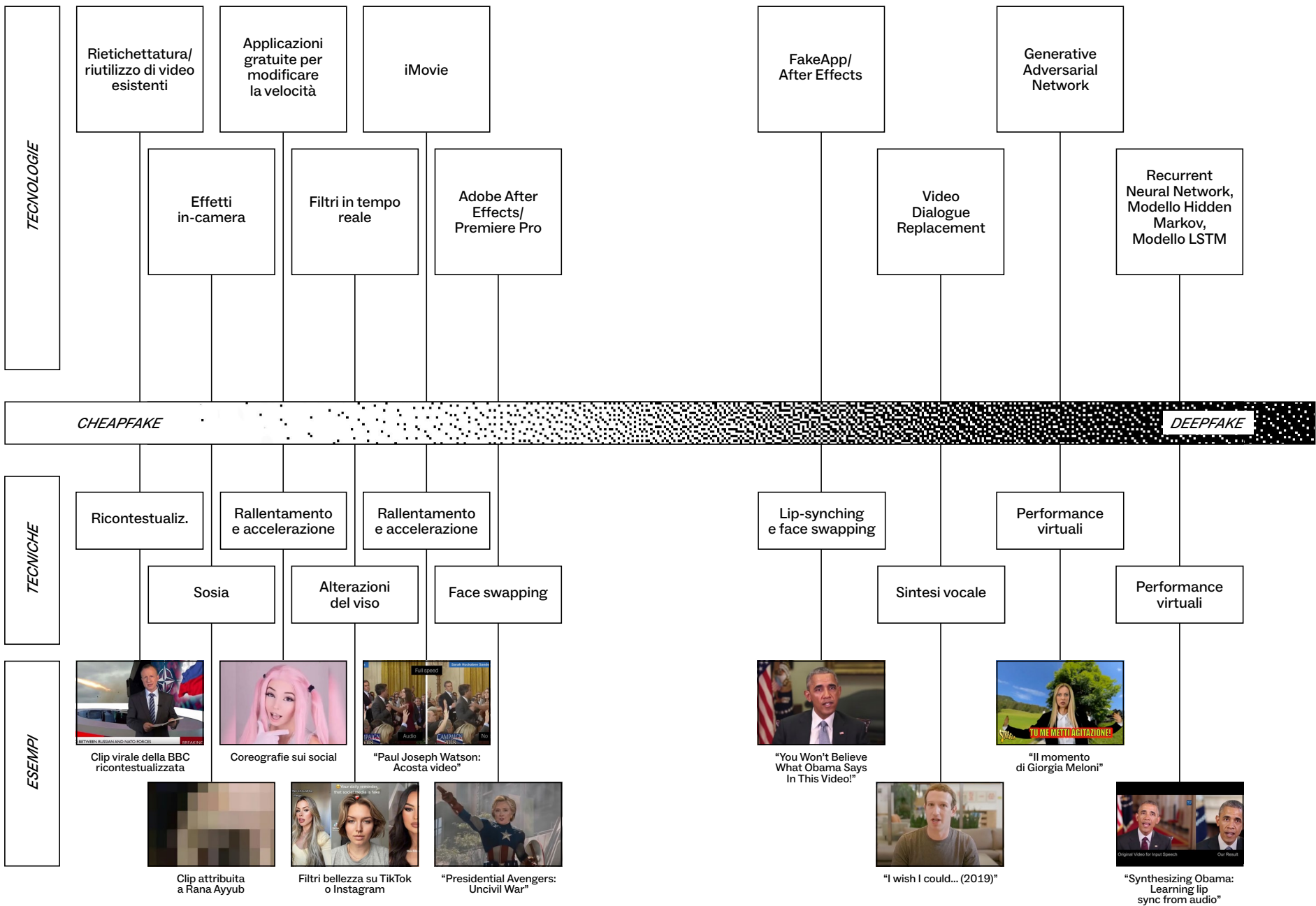
† FIG. 3.6 USER INTERFACE DI PHOTOSHOP 1.0 SU MACOS, 1990

tuale che ha lo scopo di soddisfare desideri sessuali, minare la credibilità di un individuo o infliggere danni a livello politico. Una volta che queste immagini vengono distribuite online, sono utilizzate come merce di ricatto o inviate semplicemente ad amici e rientrano, nonostante la loro natura fallace, nella sfera dell'abuso. Questi contenuti influiscono sia a livello emotivo, sia sul piano sociale, come nel caso di Noelle Martin, una diciassettenne australiana che ha scoperto la diffusione di immagini pornografiche con il proprio volto, vedendo di conseguenza compromesse le sue opportunità lavorative (Paris e Donovan 2019: 26).

A volte contenuto audiovisivo non necessita nemmeno di una manipolazione: è il caso dei *sosia*, o *lookalikes*. Un esempio è l'episodio che riguarda la giornalista Rana Ayyub e di un suo libro dove viene fatta menzione di alcuni reati, come omicidi extragiudiziali, attuati del partito nazionalista Bharatiya Janata in India. Oltre a ricevere numerose molestie durante gli anni, la donna ha scoperto l'esistenza di un falso video pornografico su di lei diffuso tramite Whatsapp e postato da una pagina sostenitrice del partito (Paris e Donovan 2019: 27). La decontestualizzazione del materiale audiovisivo si unisce ora alla sfera dell'abuso toccata con il photoshopping sessualizzato, senza la necessità di alcun tipo di sforzo tecnico. In questo caso è infatti l'attrice raffigurata nel video che, consapevolmente o meno, contribuisce alla realizzazione del contenuto, sembrando chi in realtà non è.



† FIG. 3.7 FRAME CENSURATO DEL VIDEO DIFFUSO SU WHATSAPP
ATTRIBUITO A RANA AYYUB



↑ FIG. 3.8 RIELABORAZIONE DELLO SPETTRO CHEAPFAKE-DEEPPFAKE DI PARIS E DONOVAN (2019)

Questa rappresentazione in figura 3.8, rielaborata da Paris e Donovan (2019), ci anticipa sia lo spettro dei contenuti identificabili come deepfake, che tratteremo in 3.3, ma anche quello che è lo scenario attuale dei cheapfake. È possibile notare, osservandolo nella sua interezza, come ogni passaggio richieda abilità via via più specifiche da parte del creatore di contenuti.

Lo spettro dei cheapfake, che occupa la metà sinistra della pagina, è molto ampio e parte dalla ricontestualizzazione di frammenti di contenuti audiovisivi, per poi passare ai sosia, toccare ambiti come l'accelerazione o il rallentamento di un contenuto tramite applicazioni più o meno complesse, arrivando a filtri in tempo reale su Instagram e Tiktok e infine a face swapping veri e propri.

Ognuno di questi step richiede l'utilizzo di uno strumento digitale e di un supporto che può essere un computer o uno smartphone. Generalmente, per la generazione di cheapfake, all'utente non è richiesta particolare abilità tecnica o grande potenza di calcolo. Come osserveremo in 4.3.1 infatti, la diffusione di contenuti simili è già in atto da svariati anni e ritengo possa essere analizzata come caso studio della facile influenzabilità delle persone e della loro tendenza all'auto contaminazione nei sottoboschi online. Le piattaforme hanno faticato e stanno ancora faticando nel riconoscimento e nella moderazione di video cheapfake, agevolando la confusione e la condivisione di misinformazione senza neppure l'ausilio di intelligenza artificiale.

3.3

DEEFAKE E MEDIA SINTETICI

I *deepfake*, ancor più sbalorditivi e confusionari dei cheapfake, sono il passo successivo verso lo scenario più volte menzionato durante questa tesi. Un deepfake è un particolare tipo di medium sintetico che comprende audio, immagini e video, ed è parzialmente manipolato o completamente generato dall'intelligenza artificiale e che ha come peculiarità la rappresentazione umana (cfr. Schick 2020: 10).

Il termine *media sintetici* descrive dei media algoritmicamente generati o modificati, spesso grazie all'utilizzo di intelligenze artificiali, e fa parte di un più ampio spettro di contenuti che comprende anche la realtà virtuale.

Synthetic is actually an important and fast-moving definition of how humans and machines are collaborating to make media that is, on one hand, exciting and innovative, and on the other, disturbing and ethically ambiguous. (Rosenbaum 2019)

Il risultato del processo di creazione di media sintetici è una nuova categoria di immagini, testi, audio, video e dati generati da algoritmi: facce e persone che non esistono, individui digitali manipolati da modelli di IA (cfr. Vales 2019). I deepfake, nello specifico, sono basati su un modello di deep learning noto come *Generative Adversarial Network* (GAN), teorizzato e messo in pratica per la prima volta da Ian J. Goodfellow nel 2014 (Helmus 2022). Il nome *Rete Generativa Avversaria* fa riferimento al peculiare utilizzo simultaneo di due algoritmi posti uno contro l'altro come in un combattimento, in cui il primo di essi impersonifica il "generatore", ovvero chi crea effettivamente l'output che può essere un segno, un volto, una figu-

ra; mentre l'altro ha il ruolo di "discriminatore", ovvero colui che verifica i difetti del contenuto appena generato basandosi su un dataset di riferimento. Comunicati questi difetti al generatore, quest'ultimo impara dai propri errori migliorando la qualità e la veridicità dell'output, per poi ripresentarlo al discriminatore e così via. Si arriva infine a uno stato di accuratezza soddisfacente, ovvero un momento in cui entrambi gli algoritmi non riescono ad apportare più miglioramenti (cfr. Kalpokas 2021).

Goodfellow si rende ben presto conto che il modello GAN è estremamente efficace per la generazione di volti umani, dato che riduce notevolmente i tempi di attesa, a modello allenato, ma soprattutto permette una riproduzione spaventosamente fedele dei lineamenti facciali. Molte sono state le iterazioni del metodo GAN negli anni, ma forse la prova più immediata ed esplicativa della straordinaria precisione di questo modello si può trovare su www.this-person-does-not-exist.com, un sito web in cui a ogni refresh della pagina viene generata l'immagine di un volto credibile ma inesistente (v. fig. 3.9).

Nella stessa pagina scopriamo che lo scopo primario del neural network, sviluppato in origine da *NVIDIA* nel 2018, era quello di addestrare l'IA a riconoscere volti e migliorare le prestazioni delle proprie schede video applicando questo strumento al rendering in tempo reale dei videogiochi. Come effetto collaterale però, è stata scoperta la possibilità di generare volti umani. Sempre sulla pagina web ci viene data un'informazione cruciale: è veramente difficile riconoscere una persona falsa guardando superficialmente l'immagine, soprattutto se l'osservatore non è particolarmente esperto dell'argomento o non è un fotografo. Pochi sono infatti i difetti che il modello si porta dietro, come la rappresentazione di accessori quali occhiali, orecchini e in generale la generazione di alcuni artefatti sullo sfondo («This Person Does Not Exist» 2022).

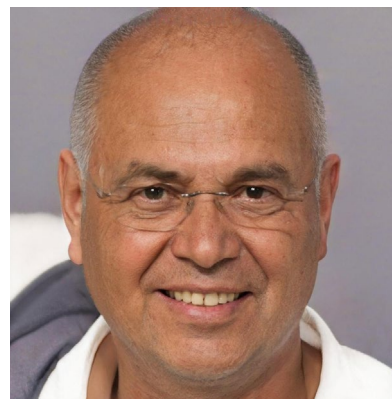
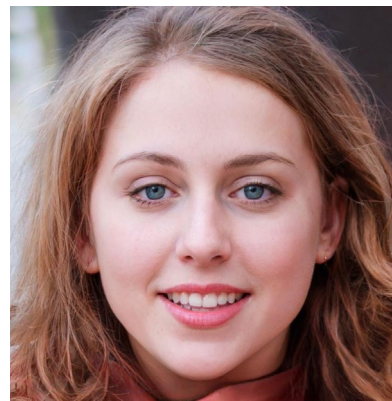
Grazie a uno studio di Nightingale e Farid (2022), constatiamo che, se messi a confronto, non solo i volti sintetici sono praticamente indistinguibili da quelli reali, ma risultano sorprendentemente anche più affidabili. In tre esperimenti separati sono state sottoposte a dei partecipanti diverse immagini di persone, chiedendo loro di classificarle dapprima come "reali" o "false". Il pool di immagini è stato composto tenendo conto anche del genere, della cultura e dell'età. La classificazione di volti caucasici è risultata più complicata, rivelando come i modelli GAN siano stati allenati con dataset di persone con questi tratti somatici, permettendo alle macchine maggiore accuratezza. Infine, a un gruppo differente di partecipanti è stato chiesto di classificare l'affidabilità percepita dello stesso set di immagini, scoprendo quindi che la percentuale è risultata più alta per quelli artificiali. Una possibile spiegazione di questo fenomeno è che i volti sintetici tendono ad assomigliare a dei volti medi, senza fattezze

particolarmente accentuate che inevitabilmente abbassano la nostra percezione di fiducia verso quella persona.

Il risultato di questo esperimento ci porta a riflettere sul grande potere persuasivo che i media sintetici possono avere. Come esposto in 1.3, nell'uomo vi è un'attrazione intrinseca per ciò che è diverso, ciò che si discosta dai nostri canoni di perfezione, ciò che appartiene a un mondo differente dal nostro. L'alterità risalta agli occhi dell'essere umano proprio a causa dei suoi attributi che si discostano dall'io. Conseguentemente, ne deriva che ogni elemento che appartiene alla media e che porta nei propri connotati una certa familiarità ci rassicura e ci tranquillizza. Questo dona un'incredibile potenza al contenuto del media sintetico, che diventa persuasivo se nella sua raffigurazione risulta affidabile. L'apertura all'alterità, che come umani postumani dovremmo imparare ad abbracciare, è anche un'arma a doppio taglio. Per questa tendenza intrinseca a rimanere nella nostra *comfort zone*, ci fidiamo di ciò che è bello e che è fatto bene, aprendoci al pericolo che i media sintetici portano con loro.

This-person-does-not-exist ci mostra come possiamo essere facilmente ingannati anche con una semplice immagine generata, ma possiamo anche adoperare questo strumento per ottenere a nostra volta dei volti e sfruttarli per i più disparati scopi. Nel lato destro della rappresentazione (v. fig. 3.8) notiamo che i deepfake, nonostante abbiano una complessità di approccio più alta dei cheapfake, siano la porta d'accesso per la produzione di contenuti prima rintracciabili solamente nel mondo cinematografico. Questi comprendono tecniche di *lip-synching*, ovvero sincronizzazione del movimento labiale, face swapping ma anche di sintesi vocale e performance virtuali per rendere credibile una specifica azione mai accaduta.

Il lavoro manuale con software dedicati viene ora sostituito da un lavoro della macchina che, grazie alle enormi potenzialità toccabili già ora con mano, genera artificialmente e in modo automatico degli output in tutto e per tutto confondibili con le controparti reali, e ha bisogno solamente di un dataset di addestramento (Kalpokas 2021).



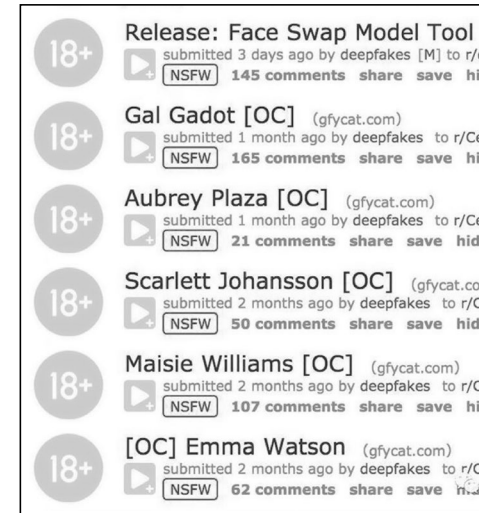
← ↑ FIG. 3.9 VOLTI GENERATI SU THIS-PERSON-DOES-NOT-EXIST.COM

3.4 FINTE CELEBRITÀ

Il 96% dei deepfake in circolazione sono video pornografici non consensuali (Schick 2020: 28), e i deepfake nascono proprio per soddisfare questo desiderio, tipico del mondo del mondo di Internet. *U/deepfakes* era il nickname di un utente anonimo di *Reddit* che a partire dal novembre 2017 ha iniziato a pubblicare sul subreddit “r/deepfakes” video espliciti di attrici hollywoodiane realizzati grazie a piattaforme open source di deep learning. Egli ha aperto il vaso di pandora e spianato la strada ad altri creatori di contenuti, che hanno iniziato a contribuire alla crescita del fenomeno. Gli utenti con un minimo di dimestichezza con il codice, infatti, sono riusciti fin da subito a generare video. Il primo video pubblicato da *u/deepfakes* aveva come protagonista Gal Gadot, in un rapporto sessuale con il suo fratellastro: mentre la notizia faceva il giro del web, la realtà dei fatti è che nessuna caratteristica estetica presente nel video apparteneva all'attrice israeliana (Schick 2020).

Oltre all'appeal del proibito che hanno il creatore del contenuto e il fruitore-osservatore nel vedere come soggetto di queste creazioni delle celebrità, la diffusione di deepfake realistici è merito anche della grande quantità di materiale fotografico a loro appartenente. L'enorme mole di immagini dei loro volti presenti su archivi di immagini stock, *Google* e *YouTube* viene data in pasto alla macchina e processata in un'ibridazione audiovisiva spaventosamente realistica. Frame per frame, l'IA sostituisce il volto desiderato sul corpo umano “donatore” con grande precisione. Non si tratta più di un fotomontaggio statico come da sempre siamo abituati a vedere, ma di un video in continuo cambiamento dove la persona sorride, urla, si muove e compie azioni che nessuno finora ha potuto vedere, e per questo

attrae e stupisce. Nonostante *Reddit*, un anno dopo, abbia bandito ogni forma di video sintetico contenente pornografia non consensuale, *u/deepfakes* è riuscito a distribuire gratuitamente il codice (Schick 2020).



Ciò ha portato alla proliferazione di numerosi programmi, che, sfruttando questa solida base, hanno declinato il modello di face swap con interfacce utilizzabili da tutti. Applicazioni come *FakeMe*, *iFace*, *FaceMagic* e *Deepfake video maker* sono fra le decine rintracciabili oggi su *AppStore*. Accessibili da smartphone, queste facilitano il dialogo utente-macchina trasformando il processo in un gioco quasi innocente, un videomontaggio da inviare ai propri amici per strappare loro una risata. In ambito pornografico, *DeepNude* forniva agli utenti la possibilità di caricare un'immagine e di ottenere in cambio la versione svestita di quella persona, solitamente di sesso femminile, dato che il modello è allenato principalmente su questa fetta di popolazione e utilizzato per la stragrande maggioranza da uomini (Cole 2019).

Mi è capitato recentemente di osservare come alcuni account su *TikTok* presentino contenuti manipolati da IA, fra cui quello di *@deptomcruise*, aperto durante la primavera del 2021 e che ormai conta più di 5 milioni di follower. Al suo interno vi sono parecchi deepfake realistici dell'attore Tom Cruise impegnato in attività come andare in palestra, fare una passeggiata al parco o addirittura partecipare a serate di gala e incontrare celebrità del mondo dello spettacolo. La peculiarità di quasi tutti i suoi video è come la figura di Tom parli direttamente

↑ FIG. 3.10 SUBREDDIT R/DEEPFAKES, 2017



→ ↑ FIG. 3.11 @DEEPTOMCRUISE





↑ FIG. 3.12 @DEEPTOMCRUISE

alla telecamera. Questa forse è la prima caratteristica che colpisce lo spettatore, poiché oltre alle sembianze, anche la sintesi vocale è particolarmente realistica. Sicuramente la persona che si presta come modello per il face swap ha studiato attentamente i movimenti dell'attore americano per rendere l'impersonificazione più credibile: è anche grazie a questo il risultato finale, con lo scorrere del video, riesce a scatenare continui dubbi nella mente dello spettatore, tanto da indurlo a chiedersi se il profilo *TikTok* in questione appartenga realmente all'attore americano (v. fig. 3.12).

Allo stato dell'arte attuale non vi è ancora modo per la persona comune di realizzare contenuti simili, quasi irriconoscibili e con questa precisione. Per allenare l'intelligenza artificiale in modo così efficace è necessaria parecchia potenza di calcolo, settimane di perfezionamento, un certo budget, abilità e cura nei dettagli, sia da parte di chi presta il proprio corpo, sia da chi verifica che ogni frame e ogni movimento sia credibile (cfr. Helmus 2022).

La natura di questo esperimento è sempre stata indicata chiaramente, in primis dal nome utente ben visibile su ogni video e poi anche dai numerosi inviti da parte di celebrità reali, quasi a considerare il finto Tom un influencer distaccatosi dalla controparte reale e divenuto un'entità a sé stante. Presenziando a questi eventi, colui che impersona l'attore fonde il virtuale con il reale, mostrandoci una realtà fittizia che percepiamo diversamente quando la osserviamo da utenti online. Questi video si avvicinano più al mondo degli effetti speciali cinematografici che a un processo attuabile dalla persona comune in un pomeriggio di noia. Il budget e le capacità messe in gioco sono notevoli, come anche il risultato. Al cinema siamo sempre stati abituati a fruire di un contenuto sapendo che questo è finzione. Il film racconta una storia che, nonostante possa avere dei fondamenti reali, è per sua natura fittizia, interpretata da attori e fortemente influenzata da processi di editing e post produzione.

La manipolazione qui è da sempre presente, ma viene sperimentata per la prima volta digitalmente intorno agli anni Duemila, quando la presenza di immagini generate al computer (CGI) diventa massiccia e l'accessibilità a strumenti in grado di gestire tali effetti aumenta, restando comunque confinata all'industria stessa (Schick 2020: 23).

Un esempio eclatante di come la tendenza stia ora cambiando è il caso di *The Irishman*: un film-evento dal budget milionario prodotto da *Netflix* e rilasciato nel 2019, il cui lasso temporale narrativo si estende per circa settant'anni. Gli attori Al Pacino, Robert De Niro e Joe Pesci sono stati ringiovaniti grazie a tecniche CGI e particolari accorgimenti durante le riprese. Nonostante ciò, l'esperimento è in parte fallito: nonostante l'incredibile dispendio di risorse da parte degli addetti ai lavori,

la qualità percepita e soprattutto il *de-aging* messo in atto non è convincente, tanto da spingere un utente su YouTube a provare a fare di meglio (*ibid.*)

Circa tre mesi dopo l'uscita del film viene caricato il video «The Irishman De-Aging: Netflix Millions VS. Free Software» (iFake 2019), che fa utilizzo di deepfake per ringiovanire i volti degli attori, comparando il risultato finale in *splitscreen* con le riprese originali. Il risultato è di gran lunga più realistico e plausibile della controparte di Netflix. L'utente iFake porta a una seconda giovinezza i volti degli attori ormai anziani, mettendo quasi in ridicolo l'enorme spesa e il lavoro svolto da Netflix, ma soprattutto sfidando l'industria cinematografica come mai prima d'ora. A detta dell'autore, l'intero processo di manipolazione video ha richiesto appena sette giorni, ma "i risultati sarebbero potuti essere ancora migliori con un ulteriore addestramento utilizzando DeepFaceLab" (iFake 2019).

In appena una settimana, con un software open source, una sola persona è stata in grado di battere una multinazionale nella manipolazione di volti credibili e di diventare il simbolo dell'imminente avvento della nuova era dei contenuti online generati da intelligenze artificiali, costruiti non solo dall'élite ma prodotti anche *dal basso*, dai singoli utenti.

NETFLIX MILLIONS VS FREE SOFTWARE

(IFAKE 2019)

3.5

SINTESI VOCALE E CLONAZIONE

È il cinque ottobre 2022, su TikTok viene caricata e diventa virale una puntata del celebre podcast di Joe Rogan, commentatore televisivo e comico statunitense. Non è la prima volta che mi capita di vedere questo tipo di contenuti sulla mia *ForYou Page*, la pagina principale dove l'algoritmo mi propone giornalmente i video. Scorrendo velocemente, forse non avrei mai approfondito il caso, ma, ascoltando qualche secondo in più, noto subito che il conduttore sta intervistando Steve Jobs, l'ex CEO di Apple ormai scomparso, in un dialogo continuo fatto di opinioni, battute divertenti e anche qualche controversia. Ovviamente quest'ultimo, non potendo prendere parte all'intervista, è sostituito da una voce artificiale (v. fig. 3.13).

Quello che però ritengo sbalorditivo è che ogni componente del contenuto è generata: Joe Rogan e Steve Jobs sono completamente sintetizzati, a partire dalle battute che si scambiano e dai ragionamenti che esprimono. Il podcast, della durata originaria di diciannove minuti, è in realtà prodotto da *Podcast.AI*, una piattaforma dove settimanalmente vengono caricate puntate con nuovi intervistati.

Navigando sul loro sito scopriamo che entrambe le voci sono generate dal software *Play.ht*, mentre i discorsi sono composti con modelli linguistici allenati sulla base della loro bio-

grafia, delle registrazioni disponibili online e delle trascrizioni dei loro speech pubblici (Podcast.ai 2022).

«The Steve Jobs episode was trained on his biography and all recordings of him we could find online so the AI could accurately bring him back to life» (Podcast.ai 2022). A detta degli stessi autori infatti, tutti i documenti e le registrazioni appartenenti all'ex CEO di Apple fanno parte del dataset utilizzato per allenare l'IA e riportarlo in vita accuratamente.

La sintesi vocale, già conosciuta in 3.4 con @deeptomcruise, ha ricevuto una forte spinta, andando a braccetto con la diffusione dei deepfake e completandoli. Definito come *voice cloning* da Helmus (2022), il fenomeno ha raggiunto il grande pubblico con applicazioni semplici da utilizzare che mimano le voci di celebrità. Come per l'output video, anche la generazione audio, diffondendosi a macchia d'olio ha fatto sorgere conseguentemente i primi problemi.

Il sito web www.NotJordanPeterson.com, ora accessibile ma privato delle sue funzionalità, è riuscito a smuovere addirittura la controparte reale. Jordan Peterson, intellettuale e psicologo canadese spesso al centro di polemiche e discussioni in clip su TikTok e Instagram, ha intrapreso infatti azioni legali verso la pagina, colpevole di permettere agli utenti di generare output audio con la sua voce. Questo vuol dire che chiunque, anche un suo detrattore, avrebbe potuto utilizzare la sua voce per diffondere disinformazione e metterlo in cattiva luce, e così è stato (Schick 2020: 91).

Per quanto riguarda invece la generazione di frasi, già con ChatGPT un utente qualsiasi può produrre un output credibile. Relazionandosi con la macchina tramite il linguaggio, possiamo chiedere al chatbot di impersonificare una determinata celebrità o una figura fittizia in base alle necessità e dare indicazioni sulle caratteristiche del discorso che vogliamo ottenere, come il contesto, l'argomento che deve trattare, il modo in cui si deve esprimere e molto altro.

Synthesia.io è forse lo strumento che, a oggi, riesce a chiudere il cerchio e permettere all'utente di sfruttare la sintesi vocale in sinergia con un supporto video. Su questa piattaforma online è possibile infatti scegliere tra quasi un centinaio di persone virtuali che prestano volto e voce alle nostre frasi inserite come input. Unendo il visivo all'uditivo, Synthesia.io si pone come provider per un servizio di generazione video totalmente sintetici in un processo user friendly principalmente rivolto al mondo delle imprese (Synthesia 2022). Secondo il fondatore e CEO dell'azienda infatti «synthetic media is a "glimpse into the future of how humans will create content"» (Schick 2020: 32). La grande flessibilità del modello permette di produrre un video in 120 lingue contemporaneamente, abbassando drasticamente i tempi e i costi di produzione video. Nonostante il tar-

get primario di strumenti simili siano le aziende, non è escluso l'utilizzo anche da parte di persone comuni.

Il modello *text-to-speech* di *ElevenLabs*, poi, spinge la generazione di voci sintetiche oltre i limiti finora presentati, consentendo all'utente di clonare la propria voce o quella di amici, parenti o sconosciuti, nonché di variarne l'enfasi, la cadenza e le pause nel discorso per rispecchiare fedelmente l'essere umano di riferimento. Possiamo considerarlo un vero e proprio *laboratorio di clonazione di voci*, dove l'input è una registrazione di qualche minuto di una persona a piacere. Il software fa il resto, creando un modello adattabile a ogni frase e modificabile. Vi è addirittura la possibilità di generare una voce completamente nuova regolando parametri come genere, età e marcatezza dell'accento (ElevenLabs 2023). La corsa a modelli di IA ultra realistici permette di dare voce a testi di fantasia, articoli o paper scientifici e più in generale di sbloccare una nuova frontiera della progettazione in più campi; una missione che porta con sé delle implicazioni non trascurabili.



↑ FIG. 3.13 VIDEO SU TIKTOK DEL PODCAST FRA JOE ROGAN E STEVE JOBS, 2022

3.6 MODELLO DIFFUSION. DAL TESTO ALL'IMMAGINE

Forse però, i modelli che più hanno cambiato il panorama di generazione di contenuti sono i *text-to-image*, già menzionati in 2.3, ma che alla luce delle informazioni esposte in questo capitolo trovano una collocazione più solida nella linea temporale del progresso tecnologico. Nonostante molte delle tecnologie appena elencate siano apparse online quasi contemporaneamente, i software e le applicazioni *text-to-image* rappresentano, oggi, il *non plus ultra* di ciò che una persona comune può adottare senza impiegare alcuno sforzo tecnico. Il già citato DALL·E 2 permette di ottenere immagini da una frase inserita come input ed è forse il primo esempio di IA generativa basata sul *modello diffusion* utilizzabile dal grande pubblico. Questo è accaduto per via della risonanza mediatica avuta durante il corso del 2022 ma anche per l'incredibile facilità di approccio da parte dell'utente non esperto, che solamente tramite una stringa di testo ha nelle sue mani un potere fuori dal comune.

Contemporaneamente alla nascita di DALL·E 2 è nato un'ulteriore servizio accessibile solamente dalla piattaforma *Discord*, basato anch'esso sul modello *diffusion*: *Midjourney*. L'approccio dell'utente qui è leggermente differente, ma i risultati ottenibili sono paragonabili. Il programma fornisce sempre un *box di input* in cui descrivere a parole la sua idea, ottenen-

do di conseguenza quattro immagini rappresentative della frase inserita. Ma come funziona il modello diffusion?

«Creating noise from data is easy; creating data from noise is generative modeling» (Song et al. 2021). Semplificando di molto il processo di generazione delle suddette immagini, possiamo dire che la rete neurale interpreta le relazioni fra le parole comunicate dall'utente e il loro significato, per poi procedere con la fase generativa. La macchina parte quindi da un'immagine composta per la sua totalità da rumore gaussiano, da cui successivamente estrapola l'immagine finale tramite un processo di *denoising*, quasi come se fosse un artigiano che, partendo da un blocco uniforme di creta, ottiene la sagoma finale desiderata manipolando il materiale (v. fig. 3.15). Questo metodo è già rintracciabile su strumenti che permettono l'aumento di risoluzione o nitidezza di immagini. Come designer lo utilizziamo per produrre output visivi sia con software proprietari che servizi online, come in Photoshop quando, durante l'ingrandimento di un'immagine, è possibile selezionare quale filtro il programma deve applicare per interpolare i pixel.

È stato scoperto, quindi, che il processo risulta applicabile, dopo un periodo di training specifico, anche alla realizzazione di strumenti digitali per la produzione di immagini ex novo (Song et al. 2021).

Even for laypeople not blessed with artistic talent and without special computing know-how and computer hardware, the new model is an effective tool that enables computers to generate images on command. As such, the model removes a barrier to ordinary people expressing their creativity. (LMU München 2022).

Queste sono le parole di Björn Ommer, professore all'*LMU* di Monaco e componente chiave del gruppo *Machinist Vision & Learning group*, che insieme *Stability AI* e *Runway*, ha contribuito alla realizzazione del modello *Stable Diffusion*. Questo viene presentato come uno strumento particolarmente adatto per i non addetti ai lavori, per chi non ha talento artistico o conoscenze informatiche. Inoltre, a differenza di *DALL-E 2* e *Midjourney*, *Stable Diffusion* non possiede le limitazioni intrinseche di un servizio puramente online. Il codice sorgente è stato infatti rilasciato pubblicamente, e quindi risulta installabile su qualsiasi computer (*Stability AI* 2022).

Questo consente in primo luogo di non dover pagare alcun abbonamento, funzione che i servizi precedenti prevedono dopo un certo numero di generazioni; permettendo in secondo luogo una maggiore flessibilità nell'ottenimento dei risultati: una mano esperta può capire veramente come ragio-

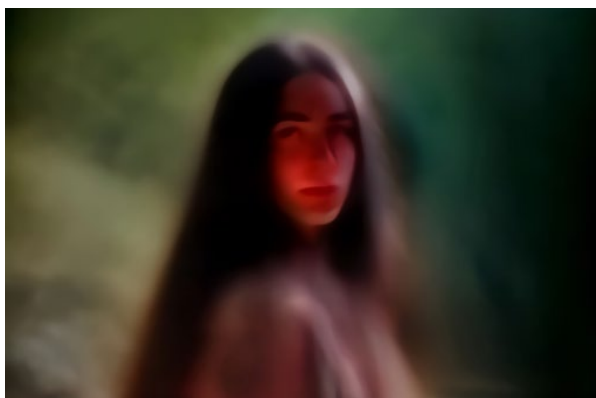
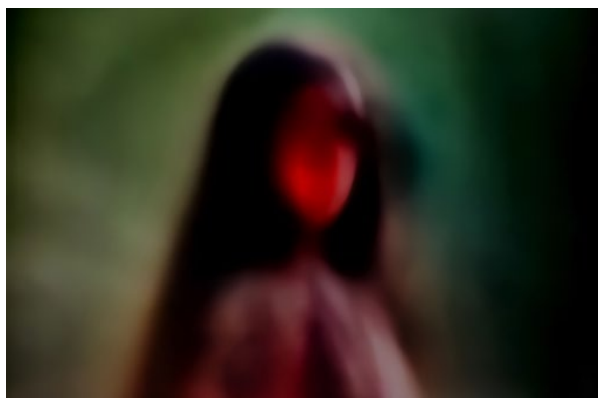
na la macchina, utilizzando dei prompt specifici e dettagliati, e apportare delle regolazioni di fino. Ma la mano esperta non è assolutamente necessaria all'utente comune per approcciarci in modo soddisfacente al modello. Data la sua natura open source, è presto comprensibile come chiunque possa adattarlo ma soprattutto allenarlo con un dataset a piacere e renderlo disponibile ad altre persone semplicemente caricandolo su un server. Il modello può essere applicato su diversi supporti, sia siti web che direttamente sulla piattaforma Discord.

Il modello diffusion si collega nuovamente, come per i deepfake, al mondo della pornografia. Con *Unstable Diffusion*, l'utente può generare contenuti espliciti di sorprendente qualità senza sforzo. Il modello open source di *Stable Diffusion* è stato qui allenato con dataset dedicati per la riproduzione di fattezze umane e poi lanciato sulla piattaforma *Patreon* per raccogliere fondi da investire nell'allenamento del modello. Nato intorno all'agosto 2022 su Reddit, quasi in concomitanza con il rilascio del modello madre *Stable Diffusion*, il progetto è riuscito a guadagnare più di 2500 dollari al mese, migrando poi su un canale Discord (*Wiggers e Silberling* 2022).

Stability AI non ha mai negato a chi utilizza il proprio codice di generare output a sfondo sessuale, a patto che queste ultime non violino leggi o danneggino il prossimo. Il dataset del modello originale contiene meno del 3% di immagini esplicite, ma l'unico freno imposto alle community online è stato quello di consigliare un "utilizzo responsabile dello strumento".

Il processo di training di *Unstable Diffusion* non è stato immediato e i risultati non sono stati sempre soddisfacenti, ma, oggi, si è raggiunto un livello impressionante di dettaglio e fedeltà alla rappresentazione anatomica corretta del corpo umano. Nonostante gli amministratori abbiano imposto rigide regole per quanto riguarda la generazione di pornografia, fra cui il divieto di creazione e diffusione di contenuti raffiguranti minori, è facile immaginare come strumenti basati su questo modello, una volta rilasciati e sviluppati, possano diventare più complicati da monitorare (*Wiggers e Silberling* 2022).

È presto detto che questo scenario di quasi totale liberalizzazione permetta di ampliare senza troppi sforzi le potenzialità dei modelli di intelligenza artificiale, soddisfacendo qualsiasi desiderio dell'utente e permettendogli di ottenere immagini generate da IA a proprio gusto, sia in termini di stile artistico che di preferenze sessuali.



→ ↑ FIG. 3.14 STEP DI GENERAZIONE DI UN'IMMAGINE SU MIDJOURNEY

LA MARIONETTA. COME L'UMANO SI REPLICA ONLINE



→ (PAG. SUCC.) FIG. 3.15, 3.16 IMMAGINI GENERATE CON MIDJOURNEY

MODELLO DIFFUSION. DAL TESTO ALL'IMMAGINE



LA MARIONETTA. COME L'UMANO SI REPLICA ONLINE



MODELLO DIFFUSION. DAL TESTO ALL'IMMAGINE

3.7

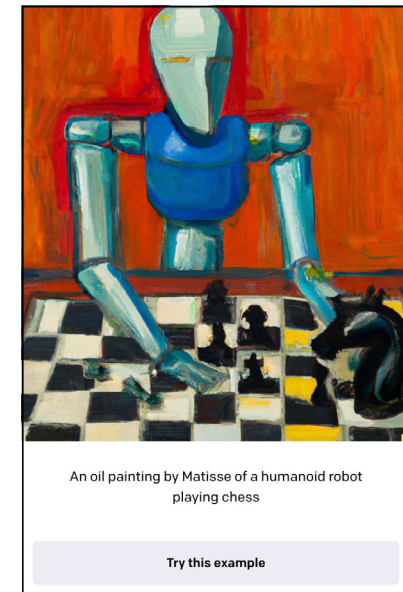
MARIONETTE E RESURREZIONE

Il percorso per la generazione della persona virtuale non è univoco ma anzi, è caratterizzato, ora più che mai, da un numero di strade percorribili in costante crescita. Il suo identikit non è definibile, poiché porta con sé il concetto di fluidità: la *mariionetta* è completamente personalizzabile per ogni generazione, che sia essa tramite un input testuale in un modello text-to-image oppure un'elaborazione più avanzata come un deepfake. Ogni volta che il creatore interagisce con la macchina ottiene un risultato univoco e irripetibile. Allo stesso tempo però, questa rappresentazione è un'iterazione della persona-tipo con cui il modello è stato allenato, declinata in un nuovo contesto.

In 2.4 è stato illustrato come nella storia l'essere umano desidera avere il controllo sui propri figli, creando degli automi, delle macchine a suo servizio. È possibile traslare questo concetto e considerare l'utente di questi nuovi strumenti come un generatore ma anche come un *burattinaio*. Egli genera ma comanda anche le sue creature virtuali come marionette al suo servizio, facendole apparire, parlare e pensare come meglio crede. Nonostante ciò, l'utente non è sempre consapevole del potere che ha nelle sue mani e delle responsabilità che ne derivano. Le proprietà intrinseche di strumenti generativi basati su intelligenza artificiale fanno sì che ogni output da questi prodotto sia in qualche modo una rivisitazione di ciò che la macchina ha già visto nel momento in cui è stata allenata. Il training, che già abbiamo rintracciato nell'archetipo di un'IA generativa come quella di AARON e abbiamo rivisto nei modelli odierni, si basa principalmente su ciò che è reperibile e visitabile online, su dati e informazioni umane proiettate nel virtuale. Ogni qualvolta lasciamo traccia nel mondo di internet, prestiamo il fianco

a uno scenario nel quale le nostre informazioni, le nostre fotografie o i nostri pensieri possono essere copiati e reinterpretati. E se ci sono abbastanza informazioni su di noi, qualcuno può generare una copia carbone, un *surrogato* di come appariamo, controllandolo, senza nemmeno chiedere autorizzazione (Cortex 2022).

Ciò che può essere imitato non è solamente il nostro aspetto, ma anche come ci esprimiamo artisticamente e comunicativamente. Prendendo nuovamente esempio AARON (2.2), egli ha appreso i connotati espressivi dell'artista che l'ha programmato, imparando a realizzare output indistinguibili in tutto e per tutto. Ho spesso notato che le campagne di comunicazione dei provider di questi strumenti hanno sempre ammiccato alla possibilità di rappresentare un soggetto "nello stile di...", ovvero applicare i connotati estetici di un certo artista come patina alla propria idea. Il simbolo del lancio di Stable Diffusion sul comunicato stampa di Stability AI (2022) è una reinterpretazione sintetica della *Notte Stellata* di Van Gogh, mentre nella homepage di DALL·E 2 è fatta più volte menzione di come si possano inserire prompt per ottenere un'immagine come se fosse stata dipinta da artisti celebri quali Matisse.



† FIG. 3.17 SEZIONE NELLA LANDING PAGE DI DALL·E 2 DOVE VIENE INVITATO L'UTENTE A GENERARE IMMAGINI NELLO STILE DI MATISSE

È chiaro che queste associazioni fra idea e rappresentazione estetica della stessa aiutino l'utente a comprendere il funzionamento dello strumento, soprattutto durante un primo approccio. Anche questi output, seppur non raffigurino necessariamente un essere umano al loro interno, possono essere considerati come surrogati del modo di comunicare dell'artista, attributi applicabili a piacimento tramite un programma ma che portano con loro un pensiero e un modo di esprimersi di una persona.

Riprendendo ciò che abbiamo osservato in 3.5, possiamo dedurre che il poter clonare l'aspetto, il comportamento e il modo di esprimersi di una persona non implica che quella persona debba essere necessariamente in vita. Far *risorgere* Steve Jobs per aprire un dibattito in un podcast ne è un primo esempio, come può anche esserlo il generare un nuovo quadro di Leonardo Da Vinci, redigere una nuova opera di Shakespeare oppure scrivere un nuovo copione per un film di Kubrick. L'IA permette all'essere umano di generare infinite iterazioni di uno specifico contenuto, soddisfacendo il suo bisogno da fanatico di vedere il proprio artista preferito ripetersi in eterno, distruggendo contemporaneamente il valore delle opere prodotte. Basandosi solamente su dati noti, l'IA ricomponi i connotati e realizza un output gradevole, in grado di soddisfare, apparentemente, chi ha dato il comando.

Possiamo affermare quindi che questi strumenti attuino un processo creativo, che creino un output, sia esso visivo, testuale o vocale. Ma allora perché non sono in grado di pensare come penserebbe l'autore o l'artista di riferimento?

Il termine "creare" può significare portare alla luce un qualcosa che non esisteva prima, materializzare qualcosa dal nulla, un attributo teoricamente applicabile, in termini religiosi, solamente a Dio. Una vera definizione di quello che vuol dire creare è però complessa da stilare. Per questo, il concetto di creatività può essere esteso anche al produrre qualcosa, al combinare elementi già esistenti per la generazione di un risultato. Tuttavia, c'è da distinguere il processo appena citato con quella che in realtà è una caratteristica peculiare dell'essere umano, ancor di più del progettista. Come umani infatti siamo in grado di *inventare*. Inventare significa scoprire qualcosa di ignoto a noi, ma già potenzialmente appartenente al mondo in cui viviamo, all'esistenza fatta di regole a cui sottostiamo ogni giorno. Inventare significa scoprire un *oggetto ignoto*, ovvero un oggetto che è ancora nel mondo delle possibilità, un *assente possibile* (Zingale 2020).

«Un oggetto possibile è un oggetto virtualmente già esistente, ma nascosto, non immediatamente visibile. Spesso però è sufficiente una ricombinazione dei dati a disposizione o un diverso sguardo per renderlo visibile» (Zingale 2020).

Ma le macchine possono inventare? La creatività della macchina è meglio definibile con il termine di *inventiva combinatoria*, una forma di inventiva che prevede la ricombinazione di dati (ad esempio quelli del training) già presenti nello scibile umano per la generazione di un risultato (cfr. Ferraresi 2023). Le IA, attualmente, possono inventare combinando qualcosa di già conosciuto.

Questi dati però, se osservati da una mente umana, possono diventare degli indizi di come quell'input può essere eventualmente cambiato o riorganizzato. Come umani infatti, oltre a un'inventiva combinatoria, possediamo anche quella che Ferraresi descrive come *inventiva trasformativa*. L'azione che la macchina non è in grado di fare è di trovare un significato alternativo da dei dati e trasformarlo, applicandolo poi in una differente situazione (*ibid.*).

L'IA generativa, come espresso nel nome, genera dei contenuti partendo da delle informazioni date, ma non è in grado di andare oltre l'ovvio e l'evidente e cercare delle *connessioni nascoste*. Può capitare, visionando un suo output, di rimanere stupiti vedendo una decontestualizzazione e ricontestualizzazione di un dato elemento, ma è bene ricordare che questo processo è attuato inconsapevolmente dallo strumento. La macchina non comprende se quel salto sia corretto o meno, lo fa e basta.

L'utilizzo di strumenti generativi è estremamente efficace per ottenere un risultato che *sembri qualcosa*, ma da solo non basta per attuare un processo progettuale. Le marionette traggono in inganno ma allo stesso tempo sono caratterizzate da una superficialità tipica di un contenuto generato, le cui referenze sono dati già esistenti. Per questo, allo stato attuale, preferisco identificare i software e i servizi elencati in precedenza non come delle unità senzienti, ma come degli accessori al servizio del burattinaio, dei possibili compagni di lavoro complici della nascita di una nuova era della rappresentazione in grado di dare spazio a desideri e ossessioni tipiche dell'essere umano.

→ (PAG. SUCC.) FIG. 3.18-3.33 CELEBRITÀ GENERATE CON MIDJOURNEY



LA MARIONETTA. COME L'UMANO SI REPLICA ONLINE



MARIONETTE E RESURREZIONE



LA MARIONETTA. COME L'UMANO SI REPLICA ONLINE



MARIONETTE E RESURREZIONE



LA MARIONETTA. COME L'UMANO SI REPLICA ONLINE

100



MARIONETTE E RESURREZIONE

101



LA MARIONETTA. COME L'UMANO SI REPLICA ONLINE



MARIONETTE E RESURREZIONE



LA MARIONETTA. COME L'UMANO SI REPLICA ONLINE

104



MARIONETTE E RESURREZIONE

105



LA MARIONETTA. COME L'UMANO SI REPLICA ONLINE



MARIONETTE E RESURREZIONE



LA MARIONETTA. COME L'UMANO SI REPLICA ONLINE

108



MARIONETTE E RESURREZIONE

109



LA MARIONETTA. COME L'UMANO SI REPLICA ONLINE

110



MARIONETTE E RESURREZIONE

111

04

VIVREMO
L'INFOCALYPSE

4.1 COME I GOVERNI MANIPOLANO L'INFORMAZIONE

4.2 INFOCALYPSE, POLARIZZAZIONE E RABBIT HOLE

4.3 CONTENUTI GRASSROOTS E CULTURA PARTECIPATIVA

4.3.1 MISINFORMAZIONE SULLE DARK PLATFORM: BITCHUTE

4.3.2 UMANI SINTETIZZATI E PORNOGRAFIA: UNSTABLE DIFFUSION

4.4 LA TECNOLOGIA È TRASPARENTE

In precedenza ho affermato come ci si trovi in una realtà online dove tutti siamo creatori di contenuti: già senza l'ausilio di intelligenze artificiali, in realtà, le nostre possibilità sono molteplici. La diffusione di nuovi strumenti tecnologici ci permette di generare marionette al nostro cospetto, completamente controllabili e personalizzabili in ogni loro parte. Queste rappresentazioni di surrogati vengono di conseguenza utilizzati per spargere disinformazione e misinformazione.

L'utente inconsapevole, che ricade anch'esso nella sfera dei creatori di contenuti, una volta avuto accesso ai nuovi strumenti di intelligenza artificiale rischia quindi di bypassare il processo di pensiero che caratterizza invece la natura di un progettista, finendo inevitabilmente per produrre degli output misinformativi. Quello che ci si presenta di fronte grazie alla liberalizzazione delle intelligenze artificiali è uno scenario definibile come *infocalypse* (Schick 2020: 9), o *infocalisse*, spaventoso e complicato da affrontare sia per chi fruisce del contenuto, sia per chi lo modera, ma allo stesso tempo estremamente efficiente per la proliferazione di sottoboschi.

4.1 COME I GOVERNI MANIPOLANO L'INFORMAZIONE

La creazione e la condivisione di *informazione*, *disinformazione* o *misinformazione* avviene generalmente da tre poli: persone comuni, quindi utenti soli o organizzati in comunità; governi e associazioni di potere; aziende e corporazioni (Karlova e Fisher 2013).

Analizzando dapprima la disinformazione, scopriamo come la sua definizione sia relativamente recente. Introdotta da Joseph Stalin dopo la Seconda Guerra Mondiale e diffusa più ampiamente intorno agli anni Ottanta, deriva dal termine russo *дезинформация*, traducibile con *dezinformatsiya*. L'Unione Sovietica, autoproclamatasi bersaglio di tattiche di diffusione di notizie false per manipolare l'opinione pubblica da parte dell'Occidente, è stata infatti la prima a utilizzare questo termine (Taylor 2021). Ironicamente, sono stati proprio i sovietici a far scuola da questo punto di vista, attuando le campagne di disinformazione forse più iconiche della storia recente.

Schick (2020: 37) ci rivela come prima dell'era di internet e durante la Guerra Fredda, il KGB, acronimo di *Komitet gosudarstvennoj bezopasnosti*, ovvero "Comitato per la sicurezza dello stato", abbia usato la disinformazione trasmessa al mondo occidentale come arma di divisione e di distrazione del nemico americano, attuando una vera e propria guerra psico-

logica. Un esempio della potenza organizzativa e progettuale dell'agenzia di intelligence sovietica è l'*Operation Infektion*, un'operazione speciale di manipolazione delle informazioni per diffondere quella che ora chiameremmo *fake news*, fondata sulla presunta creazione del virus dell'AIDS come arma biologica da parte degli Stati Uniti con lo scopo di uccidere uomini omosessuali e di colore. L'incredibile accusa è apparsa per la prima volta sul giornale indiano di Nuova Delhi *The Patriot*, testata perlopiù sconosciuta ma sponsorizzata dal governo Russo.



La notizia, uscita nel 1983 ma passata inosservata per circa sei anni, ha compiuto un percorso verso la risonanza mediatica estremamente più lungo di quello che potremmo osservare oggi online. La perseveranza nell'accusare gli States a partire dall'uscita dell'articolo, ha fatto sì che stazioni radio nazionali, stampa e pubblicazioni scientifiche anche extra-sovietiche citassero nuovamente questo articolo proveniente dall'India come fonte attendibile, spesso tramite corruzione delle testate locali in paesi in via di sviluppo. Con l'aumento dei casi di AIDS nella seconda metà degli anni Ottanta, la notizia ha iniziato a riecheggiare mediaticamente fino a diventare virale in tutto il mondo.

Abbiamo parlato prima di armi di divisione dell'opinione pubblica, e questo ne è un esempio. L'Unione Sovietica è riuscita per la prima volta a dividere la stessa popolazione degli USA, prendendo come bersaglio le minoranze interne al paese,

† FIG. 4.1 PRIMA PAGINA DEL THE PATRIOT DI NUOVA DELHI, 17 LUGLIO 1983

già emarginate in precedenza dallo stato, e creando un malcontento su due fronti. Le informazioni malevoli sono ancora oggi radicate nelle comunità afro-americane, dove si crede che il virus sia stato prodotto in laboratorio dal governo.

Da questa prima grande campagna capiamo che, in primo luogo, la progettazione può avvenire sia per il bene che per il male, e se è disinformativa spesso avviene dall'alto. In secondo luogo, se gli effetti di un'operazione strutturata, ma avvenuta offline, si possono rintracciare ancora oggi, cosa può accadere invece su Internet?

«Russia's old Cold War techniques have become easier to deploy in the Infocalypse» (Schick 2020: 39). La Russia ci ha mostrato ancora una volta, nelle elezioni presidenziali americane del 2016, come sia stato possibile dividere la popolazione, infiltrandosi e addirittura creando delle community *Facebook* ad hoc da ambo gli schieramenti politici. La strategia, realizzata a tavolino dall'IRA (*Internet Research Agency*) e che questa volta ha sfruttato la tecnologia informatica, ha dapprima previsto l'instaurazione di un senso di appartenenza dei nuovi membri, creando un'identità di gruppo paragonabile a una *tribù*, tramite la condivisione di contenuti quali immagini, video, screenshot e *meme* a supporto della fazione politica. In un secondo momento, le community sono state inondate di contenuti e messaggi negativi verso le altre tribù, con lo scopo di esasperare le divisioni già presenti nella società e distrarre la popolazione, mentre il governo Russo stava attuando operazioni di hacking a sistemi di voto e sabotaggio alla candidata Hillary Clinton.

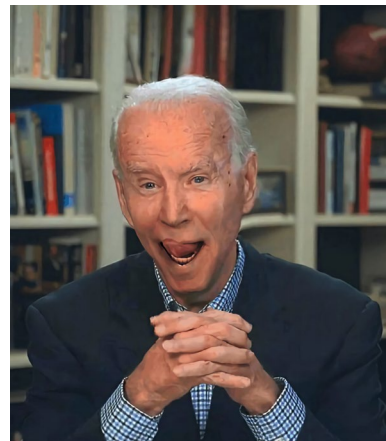
L'effetto collaterale, amplificato all'epoca dei fatti dal mondo social specialmente di Facebook e Twitter, è stato quello del *repost* dei contenuti dagli stessi candidati politici dell'opposizione, specialmente da Donald Trump. L'ex presidente degli Stati Uniti infatti era infatti solito condividere in modo quasi spudorato materiale audiovisivo chiaramente manipolato o decontestualizzato, continuando con operazione anche durante la sua carica.

Esempio ne è il deepfake "Sloppy Joe", dove è stato preso di mira il candidato Joe Biden, rappresentato mentre si contorce e arrotola la lingua (v. fig. 4.2). Il video, che in origine apparteneva a un account anonimo, è stato prontamente condiviso dal repubblicano in un'azione che unisce i due poli della creazione di contenuti disinformativa (Schick 2020).

Trump's populist approach in the Infocalypse is making the United States a more divided and dangerous country. The distrust and polarization he normalizes in this increasingly corroded information ecosystem means that it will not take much for violence to spill over into real life.

In the Infocalypse, that violence can spread faster and become ever more difficult to control. (Schick 2020: 69)

I contenuti generati tramite una semplice applicazione per smartphone possono raggiungere la notorietà grazie a un palcoscenico come il profilo Twitter di Trump, un centro nevralgico di quello che Schick definisce come un punto di partenza dello scenario infocalittico. L'ex presidente è stato in grado di dividere l'opinione pubblica di un'intera nazione e porre le fondamenta per un nuovo sistema dell'informazione, in grado di esplodere nel mondo reale da un momento all'altro.



↑ FIG. 4.2 "SLOPPY JOE"

4.2

INFOCALYPSE, POLARIZZAZIONE E RABBIT HOLE

Il concetto di Infocalypse è apparso per la prima volta nel 2016 grazie ad Aviv Ovadya. Durante la campagna elettorale americana infatti, egli ha notato sempre di più come il web, contenitore dell'ecosistema dell'informazione odierno, sia incredibilmente malsano. Le piazze virtuali, ovvero i social network, dove ci rechiamo ogni giorno per fruire e scambiare informazioni sono le principali artefici di questo scenario. Gli algoritmi che regolano queste piattaforme tendono, infatti, a premiare contenuti fuorvianti e dal carattere estremista, generando per loro natura più visualizzazioni e discussione (Warzel 2018).

Quello che TikTok, Facebook, Instagram, Google hanno in comune è il voler dare la priorità alla diffusione di un contenuto piuttosto che alla qualità dell'informazione, polarizzando le opinioni degli utenti e facendoli cadere in tunnel senza fine. Secondo Ovadya, l'imminente infocalypse potrebbe scattare al raggiungimento di una soglia critica di disinformazione e disinformazione *addictive*. Egli ha delineato quello che definirei come una proiezione distopica su quello che potrebbero essere i prossimi due decenni, caratterizzata da campagne di disinformazione gestite interamente dall'intelligenza artificiale, che permetterebbero di tagliare esponenzialmente i tempi tecnici di ideazione e realizzazione (Warzel 2018).

Abbiamo avuto già un assaggio della pericolosità e della confusione tipica di questi luoghi virtuali, ma non è difficile immaginare come ci si trovi solamente all'inizio di un problema più grande di noi, dove le piattaforme sono sia contenitori di informazione generata da utenti e organizzazioni, sia esse stesse complici della grande rilevanza dei contenuti.

«As the pace of life is getting faster and faster, people's demand for information is getting bigger and bigger» (Xu 2022).

Nell'era della *datafication* sui social media, ci troviamo di fronte a centinaia di migliaia di output, amatoriali, professionali e sintetici, che vengono smistati e indirizzati non più manualmente ma grazie all'impiego di algoritmi. Questi non solo moderano i contenuti, ma ne raccomandano di nuovi in base all'idealizzazione paradossale della nostra persona virtuale, da loro creata e basata sul nostro comportamento online (Xu 2022).

Per questo, man mano che l'algoritmo impara dai nostri interessi, ci propone contenuti sempre più affini, ignorando il fatto che idee e post particolarmente polarizzati possono attrarre chiunque. Come in un circolo vizioso, l'identikit che l'algoritmo ha quindi di noi è fortemente influenzata dalla presenza di questi output, che potrebbero piacerci o meno, ma che si auto-alimentano grazie alla raccomandazione automatica e trascinano la persona in *rabbit hole* di disinformazione e disinformazione pressoché infiniti, da cui le piattaforme lucrano grazie al posizionamento di pubblicità.

La pericolosità dell'universo online in cui viviamo quotidianamente è data anche dall'assoluta fede che come umani riponiamo nel lavoro di questi algoritmi, strumenti a cui ci affidiamo senza porci alcun quesito e a cui permettiamo di amministrare a loro piacimento gran parte della nostra giornata. Non mettiamo in dubbio le loro capacità di calcolo, che ci assistono nella scelta della strada da prendere o del prossimo video da guardare: li consideriamo, infatti, quasi come l'apice del pensiero illuminista e razionalista. Da non addetti ai lavori, li pensiamo come una *macchina utopica* in grado di funzionare senza il bisogno dell'uomo, onnisciente, il cui risultato è insindacabile e con cui siamo abituati a convivere ormai da anni. Gli algoritmi sono la reiterazione dell'ambizione illuministica di un sistema universale di conoscenza, programmi in grado di svolgere il lavoro razionale al posto nostro che ci tolgono il peso del pensiero mentre scrolliamo la homepage (cfr. Finn 2017: 6-8).

L'algoritmo è una *techne logos*,

un dispositivo "artefattuale" dotato del potere di dire, con sempre maggiore precisione e immediatezza, lo stato teoricamente esatto delle cose. Si potrebbe affermare che stiamo entrando nello stadio ultimo della tecnologia, intesa non più come discorso sulla

tecnica, ma in quanto termine che prende atto della sua facoltà di proferire verbo, proferire *logos*, al solo scopo però di garantire il vero. Questo potere costituisce la principale caratteristica di quella che viene definita "intelligenza artificiale" e che determina poi tutte le funzioni che le vengono attribuite. (Sadin 2019: 9)

Risucchiati nei nostri rabbit hole quotidiani e nell'assoluta fede verso i contenuti che ci vengono proposti, non ci rendiamo conto della potenza messa a disposizione dai nuovi strumenti generativi alle persone, capaci ora e in futuro di manipolare l'informazione in modo così credibile da riuscire a offuscare il reale confine tra realtà e finzione tramite la generazione di marionette.

«It doesn't have to be perfect — just good enough to make the enemy think something happened that it provokes a knee-jerk and reckless response of retaliation» (Warzel 2018).

Lo scenario descritto, dal carattere volutamente distopico, è sinonimo di un'imminente *information war*, una *guerra dell'informazione* che ha come protagonisti dei surrogati umani ed è guidata da progettisti malintenzionati, semplici creatori di contenuti o persone comuni. Abbiamo infatti il potere di attuare una "manipolazione della diplomazia" e potenzialmente influenzare la geopolitica utilizzando l'intelligenza artificiale, come per la recente notizia del possibile arresto di Donald Trump e lo spopolare di immagini sintetiche che lo ritraggono mentre scappa da ufficiali della polizia (v. fig. 4.3, 4.4) (Rogers 2023); oppure la presunta fotografia dell'incontro fra Putin e Xi Jinping, dove il presidente russo si trova inginocchiato ai piedi del capo di stato cinese, realmente in visita a Mosca, nell'atto di baciargli le mani (v. fig. 4.5) (Bonanata 2023).

FASTER
AND FASTER,
BIGGER
AND BIGGER



↑ FIG. 4.3, 4.4 ARRESTO DI DONALD TRUMP GENERATO CON MIDJOURNEY, 2023

Nel primo caso le fotografie sono state dichiaratamente create come satiriche e false, mentre nel secondo caso la generazione e la conseguente pubblicazione è avvenuta con intento disinformativo. Da entrambi gli esempi possiamo inoltre evincere come la qualità degli output non debba essere necessariamente perfetta, ma anzi debba essere “sufficiente” per poter ingannare l’utente: uno scatto rubato come quello di Putin e Xi Jinping assolve perfettamente la funzione, non eccellendo in qualità ma collocandosi perfettamente nell’universo dei social media dove ogni immagine è in bassa risoluzione e spesso e volentieri scattata con lo smartphone.



Un’ulteriore problematica messa in luce nel mondo social è quella del *laser phishing automatizzato*, fenomeno già esistente e che consiste nel creare messaggi falsi da parte di persone che conosciamo, raccogliendo informazioni dai nostri profili social e utilizzandole contro di noi. Automatizzando il processo grazie all’IA, sarà possibile creare copie carbone in grado di dialogare con noi in modo credibile, tramite sintesi vocale e visiva (Warzel 2018). Osservando queste possibilità, capiamo che l’avvento dell’imminente infocalypse è una riproposizione in chiave generativa di ciò che siamo già abituati a osservare oggi. Truffe e manipolazioni non sono quindi una novità nella storia umana, prima erano solamente attuate in modo più rudimentale.

↑ FIG. 4.5 INCONTRO FRA VLADIMIR PUTIN E XI JINPING GENERATO CON MIDJOURNEY, 2023

4.3

CONTENUTI GRASSROOTS E CULTURA PARTECIPATIVA

Nel mondo online siamo tutti creatori di contenuti, ma solo una piccola percentuale è attiva in un determinato sottobosco (1.3).

“Tutti sono creatori”: “La domanda è ‘di che cosa’. Ciascuno ha una sfera in cui si sente a suo agio nel prendere l’iniziativa – magari nel lavoro, magari nell’allevare i figli, oppure nel fare collezione di francobolli. Al di fuori di questa sfera, la maggior parte delle persone sono creatrici solo in circostanze attentamente limitate; la maggior parte delle persone non sa disegnare, ma tutti possono colorare dentro le figure, oppure ricalcare una traccia”.
(Jenkins, Ford, e Green 2013)

Ogni *community* ha quindi utenti che contribuiscono attivamente e utenti che osservano passivamente l’operato dei primi, ma che spesso e volentieri diventano con il tempo anch’essi attori protagonisti, scalando la piramide della partecipazione. Anche i mass media presentano questa struttura, ma ciò che è peculiare di determinate community online è la facilità di accesso e la velocità con cui si può scalare questa piramide. Questo processo prende il nome di *cultura partecipativa* e trova

la sua massima espressione nel mondo online odierno, ma ha le sue radici insite nella tendenza umana di generare piattaforme alternative per le comunicazioni dal basso. Un esempio è la *Amateur Press Association*, associazione di giovani scrittori di cultura, politica e vita quotidiana che stampava e componeva internamente le sue pubblicazioni, distribuendole tramite circuiti complessi simili a quelle che oggi definiamo reti sociali e adottando per primi la radio amatoriale. Il proliferare di sottoculture digitali staccate dai grandi provider di informazione allenta, per chi ne fa parte e per chi osserva, il controllo delle grandi aziende, favorendo la circolazione di una grande varietà di prospettive. Le persone scelgono di appartenere a una community principalmente per due motivazioni: la *resistenza* e la *partecipazione*. Vi è necessità di opposizione e rivoluzione ai media tradizionali e alle grandi aziende, di un movimento *grassroots*, ovvero dal basso, di una forma di protesta condivisa, ma anche di una volontà di condivisione (Jenkins et al. 2013: 170-74).

Questa tendenza è riscontrabile soprattutto osservando social network non convenzionali come *Gab*, *Parler* o *Bitchute*, che approfondirò in 4.3.1, dove le persone sono tendenzialmente ostili, guerrafondaie e dalla spiccata indole complottista, pronte a contaminare il prossimo con le proprie ideologie.

Negli ultimi decenni però si è affermata anche una politica online basata sulla partecipazione, proiezione di un mondo in cui il potere dei media sta principalmente nelle mani dei cittadini e quindi degli utenti, che contribuiscono invece a qualcosa, in una micro società organizzata *in e attraverso* la collettività. La cultura della partecipazione trova qui il suo massimo riscontro, diventando anche una collaborazione fra utenti per l’ottenimento di un risultato (ivi: 174-84). È il caso di molte community presenti in piattaforme quali Reddit e Discord, fra cui *Unstable Diffusion*, nella quale gli utenti si scambiano continui consigli e informazioni per l’ottenimento di un obiettivo e per il progredire del gruppo stesso.

4.3.1 Misinformazione sulle dark platform: Bitchute

Durante il mio percorso universitario da designer al *Politecnico di Milano*, e più nello specifico durante il laboratorio di *Final Synthesis Design Studio C3* nel corso di Laurea Magistrale in Design della Comunicazione A.A. 2021-2022, mi è capitato di analizzare, in alcuni sottoboschi di utenti definibili come “alternativi”, social network non convenzionali abitati da account dalle idee riconducibili all'estrema destra e dalla forte accezione complottista. I due progetti sviluppati durante lo svolgimento del corso, nel campo della data visualization, hanno come minimo comune denominatore la necessità di tracciare uno scenario quanto più possibile indicativo di come le persone comunichino, durante l'esplosione dell'epidemia del Covid-19, su piattaforme senza alcuna regolamentazione, e come le stesse piattaforme ospitino una quantità di informazioni tale da risultare opprimenti e confusionarie. Questo fenomeno è legato a doppio filo all'infocalypse, e in questa particolare accezione prende il nome di *infodemic*, o *infodemia*.

L'*Organizzazione Mondiale della Sanità* la definisce come una diffusione di informazioni false o fuorvianti in ambienti digitali e fisici messa in atto durante l'esplosione di una malattia, come può essere il Covid-19. Quest'ultimo infatti, oltre a ingenti danni umanitari, ha generato uno scenario di misinformazione online senza precedenti. Nell'infodemia, gli utenti, ma anche politici o cariche istituzionali, condividono sul web informazioni false o non verificate mettendo a repentaglio la salute del prossimo con consigli e opinioni potenzialmente pericolosi e portando contemporaneamente l'opinione pubblica a una sfiducia verso le autorità. L'infodemia è un vero e proprio prolungamento dell'epidemia “reale”, e come quest'ultima prende di mira i più deboli, i più incerti e coloro che hanno paura (Jourová 2020). I social network, essendo piazze dell'informazione, permettono la circolazione di misinformazione come mai prima d'ora. Ogni messaggio, post, fotografia o video, se inviati senza cognizione di causa, aiutano solamente il rafforzarsi di questo fenomeno, formando delle *echo-chamber*, delle bolle di contenuti misinformativi da cui è molto spesso difficile uscire (von der Weth et al. 2020).

Molte piattaforme mainstream come Twitter, Facebook e Instagram hanno quindi iniziato ad attuare contromisure per cercare di arginare il dilagare di questa pandemia online, iniziando a censurare contenuti non verificati. Questo ha portato gli utenti bannati a non avere più quelle piazze in cui erano soliti discutere e a spostarsi su piattaforme alternative, che abbiamo definito come *dark platform*, dove tutto è concesso.

Il progetto *Off Track: a journey inside the Bitchute Infodemic Underground*, sviluppato insieme ai miei colleghi Andrea

De Simone, Emanuele Ghebaour, Francesca Gheli, Hanlu Ma, Mattia Casarotto e Raffaele Riccardelli; è un sito-archivio esperienziale con lo scopo di mostrare quanto facilmente gli utenti, anche digitalmente istruiti, della piattaforma di condivisione video alternativa *Bitchute* possano perdersi in rabbit hole di misinformazione e disinformazione.

Bitchute ospita centinaia di migliaia di video banditi da altri siti web come YouTube o Facebook. Questi toccano temi come la vaccinazione, l'utilizzo di mascherine e l'origine del virus e sono caratterizzati da una forte connotazione complottistica. Non mancano infatti contenuti riguardanti *Illuminati*, *New World Order*, e più in generale opinioni controverse verso tutti coloro che occupano cariche istituzionali di rilievo.

Partendo dall'analisi di Bitchute e raccogliendo i dati sotto forma di video e relativi hashtag apparsi, cercando le parole chiave “Plandemic” e “Scamdemic”, in un lasso di tempo che va dall'inizio della pandemia nel 2020 e prosegue per tutto l'anno successivo, abbiamo progettato un network interattivo che mostri come tutti questi contenuti siano in realtà collegati fra di loro e si distribuiscano nella piattaforma, visualizzandoli in uno spazio bidimensionale (v. fig. 4.6).

Poiché l'utente comune non è solito avere familiarità con la data visualization, abbiamo anche progettato una sezione esperienziale suddivisa in quattro pagine tematiche, per poter cogliere il senso del progetto in modo più diretto. Questi quattro punti di ingresso, oltre ad essere apparentemente imparziali, prendono il nome da hashtag dalla grande rilevanza all'interno del network. Scorrendo verso il basso nelle pagine dedicate di *#Masks*, *#Vaccine*, *#Virus* e *#PCR*, è possibile notare come i video raccolti e selezionati che le compongono partano da contenuti di disinformazione generica e diventino via via sempre più aberranti, fino a raggiungere una soglia critica. Spesso gli utenti si riprendono con la webcam del proprio computer, oppure girano video da smartphone riprendendo proteste, eventi, discorsi in luoghi pubblici e situazioni surreali, conditi da montaggi amatoriali classificabili nella sfera dei cheapfake (v. fig. 4.7).

Esplorando la sezione “Jump Into” possiamo provare in prima persona come, da ipotetici utenti di Bitchute, ci si possa mettere solamente pochi minuti a perdersi in rabbit hole sempre più profondi e contraddittori che, inizialmente, condividono un tema comune, ma che spesso proseguono senza nessun nesso logico apparente tra un video e l'altro.

Osservando il network completo presente nella sezione “Archive” del sito, analizzato nelle sue componenti nel report *Among the Plandemic Truthers: a data-journey through the Infodemic conspiracist underground* durante la precedente fase del laboratorio, è possibile notare come, oltre a una forte concentrazione di video collegati agli hashtag più comuni che polarizzano

4.3.2 Umani sintetizzati e pornografia: Unstable Diffusion

Abbiamo già osservato in questa tesi come con la nascita di nuovi strumenti basati su intelligenza artificiale porti con sé un'evoluzione in ogni campo della creazione di contenuti. *Unstable Diffusion* è come un distributore automatico di contenuti pornografici personalizzabili e con parametri definibili su misura, dove tutto ciò che è immaginabile è anche visualizzabile e immediatamente fruibile. Già citata in 3.6, la piattaforma Discord permette a un utente di entrare a far parte di *canali*, ovvero gruppi chiusi e privati di utenti in una nicchia specifica. Per entrare in un server è necessaria una chiave d'accesso, un link rilasciato solamente da chi amministra quello specifico server e che garantisce la partecipazione alla community.

Dato l'innegabile appeal di poter generare contenuti espliciti con dei semplici prompt testuali, non stupisce la grande popolarità che il server, e più in generale il progetto, sta riscontrando in questo momento. Ad aprile 2023 il server conta infatti più di 240.000 membri, ma non tutti sono ancora abilitati a realizzare immagini. Il bot Discord di Unstable Diffusion infatti è in *closed beta*, ciò vuol dire che l'accesso viene rilasciato man mano a utenti in lista d'attesa tramite il portale *unstable-diffusion.com*. Nonostante ciò, il modello completo è scaricabile e installabile gratuitamente sul proprio dispositivo.

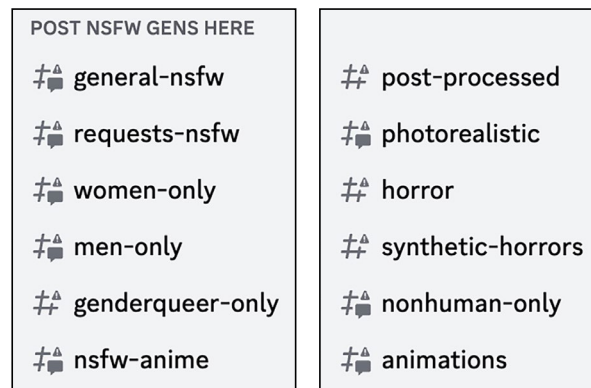
«Basically, we're here to provide support for people interested in making NSFW» dichiara AshelyEvelyn, una delle amministratrici del server, sottolineando come la missione di Unstable Diffusion sia quella di permettere agli utenti di generare contenuti pornografici in modo semplice (Wiggers e Silberling 2022). La dicitura *NSFW*, spesso utilizzata nelle community su internet, è l'acronimo di *Not Safe For Work* e si riferisce a post raffiguranti nudità e profanità forti, non adatte a essere visionate sul luogo di lavoro poiché potrebbero risultare offensive (Cambridge Dictionary 2023).

Una volta entrati in Unstable Diffusion, accessibile anche solamente da osservatori, è possibile prendere parte a discussioni, ricevere e dare consigli, ma, soprattutto, condividere con la community le proprie immagini generate.

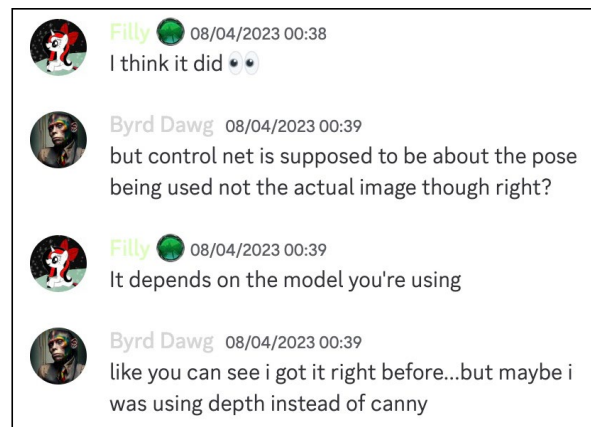
Il server è organizzato nel minimo dettaglio (v. fig. 4.10): partiamo da un'introduzione generica informativa con regole e news; proseguiamo poi con una sezione dedicata alla discussione sull'allenamento del modello di IA, alle tecniche utilizzabili e a una ricerca più vicina al mondo informatico; incontriamo i canali vocali dove è possibile confrontarsi a voce con altri membri della community, per poi passare alla sezione *SFW*, ovvero *Safe For Work*, dove vengono pubblicati contenuti non espliciti.

Scorrendo verso il basso ci imbattiamo nelle categorie NSFW, dove ogni immagine pornografica sintetizzata e genera-

ta da utenti che hanno accesso al bot è categorizzata minuziosamente per stile artistico, genere, fantasia sessuale o aspetto della persona oggetto della rappresentazione. Vi sono anche categorie non convenzionali come *nonhuman-only*, dove sono mostrate solamente creature fantastiche o animalesche dalle sembianze umanoidi o in qualche modo riconducibili per certi attributi alla sfera sessuale umana. È presente anche un luogo dedicato ai risultati "sbagliati", umani deformi o con caratteristiche fuori dal comune frutto di un errore della IA, archiviati ma comunque consultabili.



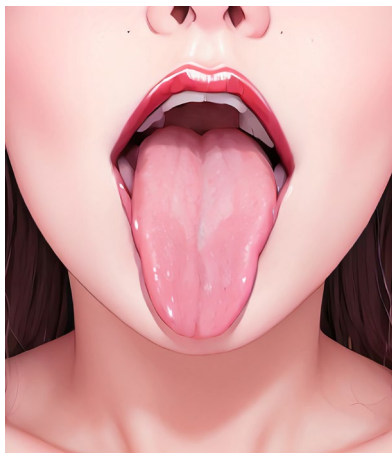
Gli utenti che hanno accesso allo strumento generativo si scambiano spesso prompt e consigli su come ottenere l'immagine desiderata, attuando un processo di perfezionamento dell'output ottenuto sempre tramite l'ausilio della macchina, spesso tramite l'adozione di *control net*, ovvero di scheletri virtuali posizionabili e regolabili a piacimento (v. fig. 4.11, 4.13, 4.14).



↑ FIG. 4.10 CANALI TEMATICI DEL SERVER DISCORD DI UNSTABLE DIFFUSION
↑ FIG. 4.11 ESEMPIO DI DISCUSSIONE FRA UTENTI SULLE CONTROL NET

Inoltre, gli amministratori del server hanno previsto delle competizioni interne in cui vengono date agli utenti delle immagini di reference da ricreare tramite il bot, con il duplice scopo di allenare l'intelligenza artificiale e di consolidare la community (Wiggers e Silberling 2022).

Unstable Diffusion è stato il primo e forse più celebre strumento per la generazione di immagini rivolte a nicchie specifiche, e per questo ritengo sia un caso studio rilevante per individuare come le community del futuro possono organizzarsi in un mondo online contaminato da media sintetici. Nonostante i risultati e le tematiche trattate risultino perlopiù controverse per un visitatore qualsiasi, Unstable Diffusion è un simbolo della cultura partecipativa grassroots, dell'organizzazione e del supporto reciproco nella realizzazione di artefatti generati mossa da interessi e curiosità comuni, senza apparenti intenzioni ostili.



Lo strumento, seppur dal training e dallo scopo evidentemente connotati, mantiene delle linee guida ferree imposte dagli amministratori e sviluppatori del progetto, con l'obiettivo di costruire una community etica per la realizzazione di pornografia sintetica. Ciò che spaventa è, paradossalmente, la possibile liberalizzazione a un pubblico più ampio: man mano che il modello verrà distribuito le immagini generate saranno più difficili da controllare e da amministrare. Gli stessi consigli scambiati tra utenti nel server potrebbero aiutare anche i malintenzionati ad affinare le proprie tecniche nell'utilizzo dello strumento, dato che ogni discussione passata e presente risulta visibile una volta entrati (Wiggers e Silberling 2022).

Un'ulteriore tematica comune a molte intelligenze artificiali generative, che Unstable Diffusion tocca pienamente, è la

↑ FIG. 4.12 IMMAGINE CONDIVISA SU UN CANALE TEMATICO

contrapposizione fra artisti "tradizionali" e artisti "IA". Mentre i primi, attivi spesso anche nel campo della pornografia, realizzano i loro output manualmente, i secondi si avvalgono quasi unicamente dell'ausilio di questi nuovi strumenti, allenati il più delle volte con dataset contenenti opere reali da cui imparare. Il malcontento generale ha portato più volte a proteste online di artisti che, vedendo opere generate spacciate come realizzate a meno su portali come *Deviantart*, hanno creato un vero e proprio *movimento anti-IA*, boicottando la piattaforma.

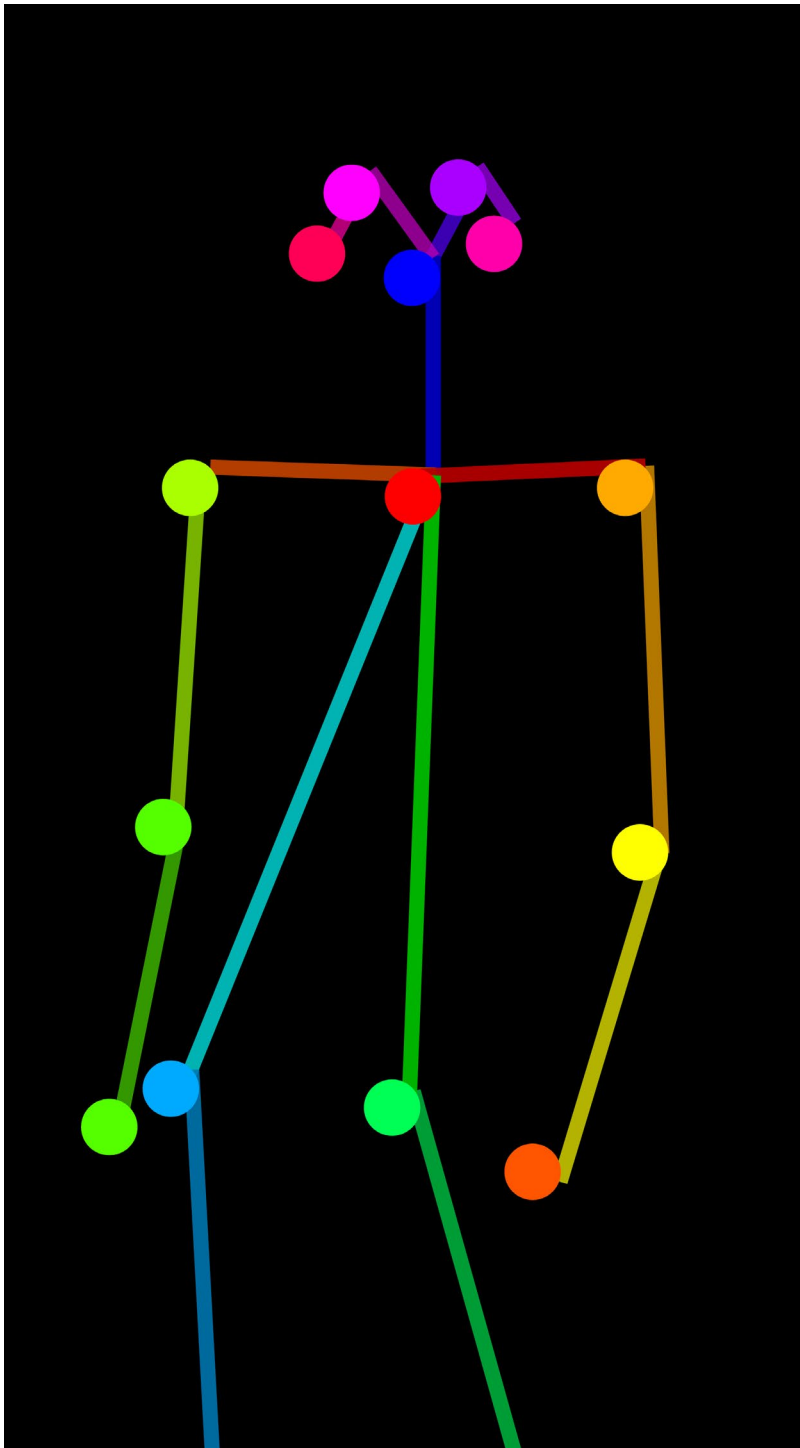
Nonostante Unstable Diffusion abbia messo in chiaro la natura del suo dataset di allenamento, la sua seconda campagna di *crowdfunding* su *KickStarter* per il finanziamento del progetto è stata bloccata. Gli amministratori hanno quindi rilasciato un comunicato che così si conclude:

Unfortunately, we have seen other communities and companies cower in the face of these attacks. Zeipher has announced a suspension of all model releases and closed their community and Stability AI is now removing artists from Stable Diffusion 3.0. But we will not be silenced. We will not let them succeed in their efforts to stifle our creativity and innovation. Our community is strong [...] and we will not be defeated by a small group of individuals who are too afraid to embrace new tools and technologies. (Silberling 2022)

Si denota così uno scenario di resistenza, dove il mondo dei creatori di contenuti tradizionali osserva con un certo distacco e paura l'avvento delle IA, mentre chi abbraccia questa nuova tecnologia non vuole cedere alle pressioni esterne. La solidità della community di Unstable Diffusion è comprovata da questa risposta, menzionando pilastri come "creatività" e "innovazione" che guidano lo sviluppo della stessa: i partecipanti non vogliono essere messi a tacere.

→ (PAG. SUCC.) FIG. 4.13 CONTROL NET CONDIVISA DA UN UTENTE SU UN CANALE DI DISCUSSIONE PROMPT

→ (PAG. SUCC.) FIG. 4.14 IMMAGINE OTTENUTA GRAZIE ALLA CONTROL NET PRECEDENTE



VIVREMO L'INFOCALYPSE

138



CONTENUTI GRASSROOTS E CULTURA PARTECIPATIVA

139

44 LA TECNOLOGIA È TRASPARENTE

Gli strumenti tecnologici possono sembrare apparentemente neutrali ma la loro stessa esistenza implica che siano stati creati da attori esterni e che abbiano ereditato da essi dei *bias*, intenzionali o meno, delle connotazioni peculiari apparentemente invisibili ma che modellano il nostro rapporto con essi (3.3).

«Yet neither are they simply neutral. Technologies, by providing a framework for action, do form intentionalities and inclinations within which use-patterns take dominant shape» (Ihde 1996: 141). La tecnologia amplia le nostre possibilità come umani e ci porge nuove possibilità di operato, fornendoci un *quadro per le azioni*. Viene donata quindi capacità di azione a chi non possiede necessariamente un'abilità tecnica, a chi in precedenza non avrebbe mai potuto nemmeno pensare di generare un determinato risultato. È come se la tecnologia, avendo delle proprietà intrinseche, ci spingesse a voler produrre qualcosa e avesse un impatto sottile ma significativo su quello che ne facciamo. Per questo, lo strumento stesso, influenzando la nostra esperienza, è considerabile come un *costruttore di significato*, e questo significato è caratteristico e spesso univoco in base alla tecnologia che stiamo utilizzando: scrivendo con penna e calamaio ci lasciamo andare al piacere della bella scrittura, lenta e laboriosa; scrivendo al computer e utilizzando una tastie-

ra invece sacrifichiamo inconsapevolmente l'estetica del glifo su carta per favorire un'inclinazione alla continua correzione e riedizione del testo, rendendo il progetto più malleabile e meno definito (cfr. Ihde 1996: 142-43).

Lo scenario descritto da Ihde porta con sé anche un altro interessante aspetto, che accomuna l'utilizzo di ogni tecnologia. Sia utilizzando una penna che utilizzando una tastiera per scrivere, oppure un determinato software per realizzare un output visuale, una volta raggiunta una certa dimestichezza non ci focalizziamo più sull'azione manuale, ma sul risultato.

Nella *fenomenologia* di Heidegger, la tecnologia si mostra a noi come fenomeno, come funzione che questa ricopre e come possibilità che ci dona. Egli prende come esempio un martello: quando lo utilizziamo, non pensiamo allo strumento ma ci concentriamo sul chiodo da battere, tanto che il primo sembra quasi scomparire. La tecnologia rivela ciò che *sta al di sotto* e ci lascia soli con quello che vogliamo realizzare, diventando *trasparente*, invisibile, ma mantenendo dei valori nascosti che inconsapevolmente danno forma al risultato finale (Thorn 2022).

«Technology becomes transparent. We don't experience it, we experience the world through it. Heidegger says that technology comes with its own way of seeing» (ivi 2022). Quando una tecnologia diventa abbastanza trasparente, avviene un'ibridazione dove la somma delle due parti crea un'entità composta differente, un nuovo soggetto che osserva il mondo da una prospettiva unica: è la tecnologia stessa a prestarci il suo modo di vedere e a sfuggire dalla nostra attenzione.

Anche il mondo di Internet modifica le nostra capacità di agire: la facilità di utilizzo di social network tradizionali, o delle loro controparti dark, abilita modi di espressione tipici di questi sottoboschi e ci fa comportare di conseguenza, spesso generando discussioni frivole, tossicità e discriminazioni; azioni che non avremmo mai compiuto nel mondo reale ma che vengono sdoganate online, rischiando di oltrepassare il limite (Cross 2022).

La trasparenza della tecnologia si riflette quindi sia nelle piattaforme in cui si comunica, sia nei programmi utilizzati per generare i contenuti da pubblicare, permettendo all'utente una possibilità di espressione e di ascolto mai visti prima d'ora. La sinergia con gli strumenti utilizzati da fruitori e creatori nel mondo online porta all'eliminazione delle barriere d'ingresso e della frizione nell'approccio a un software: nello scenario info-calittico la macchina è invisibile e ogni contenuto è generabile istantaneamente; ciò rende le idee l'unica barriera che separa la persona dall'ottenimento di un risultato tale da influenzare un potenziale pubblico.

05

**NULLA È VERO
NEL SUPERSTRATO**

5.1

DESIGNER O CREATORI DI CONTENUTI?

5.2

RAPPORTO CON LE ALTERITÀ GENERATE

5.3

APATIA DELLA REALTÀ, POSTVERITÀ E IPERREALE

5.4

RICONOSCERE I CONTENUTI: IL NOSTRO COMPITO

5.5

CONTROMISURE DALL'ALTO

Una fotografia non è più prova dell'esistenza o meno di un fenomeno, e ora anche l'audio e il video lo stanno diventando. Stiamo vivendo in un'era di passaggio e di mancata educazione al riconoscimento della veridicità del contenuto online, dove tutto ciò che osserviamo può essere contemporaneamente vero o falso. Mi domando, quindi, se effettivamente arriveremo a una regolamentazione della generazione sintetica, a una paura della verità con conseguente sfiducia nei nuovi media, oppure proveremo realmente il significato di *postverità*.

5.1 DESIGNER O CREATORI DI CONTENUTI?

Sebbene l'accesso a tecnologie di intelligenza artificiale possa democratizzare la creazione di contenuti, questo non ne garantisce la loro validità o affidabilità e non rende necessariamente l'utente un progettista. Abbiamo già osservato in 3.7 che l'IA si limita solamente a generare del contenuto con un processo di inventiva trasformativa, non è in grado quindi di compiere *salti abduktivivi* caratteristici invece della progettualità umana.

Rivolgendo lo sguardo verso quest'ultima, ora, scopriamo che l'inventiva «è la capacità di rispondere all'insorgere di un problema, non solo neutralizzandolo, ma anche modificando a proprio vantaggio futuro le condizioni che hanno dato luogo al problema» (Zingale 2020).

Come umani ed esseri viventi appartenenti alla biosfera, possediamo per natura l'inventiva e quindi la capacità di cogliere una possibilità, di immaginare quello che potrebbe accadere o potrebbe essere in un nuovo contesto con delle abduzioni, alla ricerca di un *adiacente possibile* scoperto solamente tramite la ricombinazione dei dati applicati a utilizzi differenti. Tutto ciò però non basta.

Essere progettisti nel mondo del design della comunicazione significa trovarsi dinanzi a una realtà con diversi livelli di complessità e molteplici discipline interconnesse, orientate

a differenti segmenti della società, appartenenti sia al mondo reale che al mondo online. In questo scenario è indispensabile avere un metodo e la capacità di passare da un punto di vista all'altro, senza isolare gli input ma connettendoli e facendoli circolare produttivamente. Il designer della comunicazione abbraccia quindi l'alterità, facendola propria ma sapendo sempre dove collocarsi nel contesto sociale di riferimento. Egli è un mediatore continuo, un traduttore di informazioni che non si occupa solamente dell'estetica degli output prodotti, ma applica un processo metodologico che analizza il contesto, tiene conto degli obiettivi, procede alla redazione grafica e contenutistica degli artefatti e si occupa delle possibilità di accesso comunicativo agli stessi, facilitando la loro diffusione (cfr. Baule e Caratti 2016: 21-23).

Il progettista, durante questo processo inventivo, compie un percorso di scelte e possiede un *abito*, ovvero un modo di essere della persona, una disposizione dell'animo. Abito è la metafora di qualcosa che siamo soliti avere con noi: tramite l'abito ci comportiamo in una certa maniera, ma questa non è una caratteristica innata e anzi, può essere chiamata così solamente quando abbiamo la tendenza a comportarci in un modo simile anche in circostanze simili in futuro. Avere un *abito inventivo*, o *abito progettuale* significa essere sempre pronti a osservare quello che c'è, per vedere quello che ancora non c'è, agendo per trasgressione alla normalità, oltrepassando i limiti della realtà per come si manifesta e ascoltando l'alterità. L'abito può anche mutare con il passare del tempo; il progettista deve quindi acquisire dei modi di fare, pensare e dialogare che gli facciano possedere questo abito progettuale. Di conseguenza, per ottenerlo è indispensabile sviluppare un pensiero critico (Zingale 2020), ovvero la capacità di osservazione e di giudizio di quello che si sta producendo, comprendendo se il progetto è sviluppato coerentemente ed eticamente, o in caso contrario se è necessario cambiare direzione, con il fine ultimo di ottenere un artefatto che abbia senso.

In questi termini, il designer è anch'esso cyborg, non solo per la sua fisica unione con la tecnologia, ma soprattutto per come egli ragiona. Come afferma Haraway, il cyborg non è altro che un simbolo per indicare colui che è in grado di superare la visione dicotomica e identitaria, aprendosi all'alterità e a ciò che lo circonda.

Inoltre un artefatto, seppur frutto di un pensiero progettuale, può essere progettato *non* correttamente. Questo accade poichè la correttezza della comunicazione non è scontata. Come umani, ci illudiamo di comunicare, poichè lo facciamo naturalmente, ma in ogni output vi è un significato implicito che solamente il ricevente può cogliere, interpretando con dei significati propri che possono variare. Ma soprattutto la comunicazione è un processo complesso che segue l'altrettanta com-

plexità della mente umana, rimanendo nell'oscurità finché non si attua un *dialogo*, ovvero un contatto, con l'altro. Il messaggio fra il mittente (colui che crea il contenuto) e il destinatario (colui che fruisce del contenuto su uno specifico canale, per esempio su TikTok) può non essere correttamente interpretato a causa di un mancato rispetto delle funzioni dello stesso e delle sue regole costruttive (Zingale 2020).

L'abito progettuale si rivela fondamentale nello scenario di infocalypse a cui stiamo andando incontro. Affidarsi ciecamente all'operato delle IA è come lasciare a loro la completa facoltà di decidere ogni connotato del prodotto, lasciando a noi solamente il compito della diffusione, essa stessa talvolta mal progettata. Essere unicamente creatori di contenuti focalizza l'attenzione sulla mera estetizzazione dell'output e sulla sua messa in scena, e attuarlo nell'era delle intelligenze artificiali generative può scremare ulteriormente le capacità di intervento. Chi non progetta, non considera ciò che *viene prima* e ciò che *viene dopo*, non si fa domande sui possibili effetti e conseguenze sociali dell'artefatto e come il pubblico o la community possa recepire quella che per lui può essere, solamente, una simpatica presa in giro o una manipolazione digitale in buona fede.

È questo il fulcro che permette la nascita di contenuti misinformativi online e il fuoco che alimenta sottoboschi come Bitchute. Fermo restando che la disinformazione può essere progettata in modo accurato e scrupoloso, avendo chiari obiettivi ed effetti sulla popolazione come nelle campagne dell'intelligence russa; produrre misinformazione, invece, può capitare a tutti, ma accade sicuramente più frequentemente a chi non è vestito con un abito progettuale adeguato.

5.2 RAPPORTO CON LE ALTERITÀ GENERATE

In questa nuova realtà è però fondamentale osservare anche l'altra faccia della medaglia. Oltre a dover applicare un metodo progettuale etico e corretto, come utenti del mondo online siamo a contatto, inevitabilmente e continuamente, con output ingannevoli generati da altri creatori. Queste marionette si presentano ai nostri occhi come rappresentazione visiva di quello che la persona di riferimento è. Prendiamo l'esempio di @deep-tomcruise in 3.4: egli è una celebrità replicata ma apparentemente reale che fa scattare in noi un processo identificabile come *interpretazione*. Ogni contenuto online si palesa ai nostri occhi e alle nostre orecchie come segno, interpretabile attraverso un processo di *semiosi*, ovvero la capacità di pensare e di stabilire relazioni fra le cose, dove una qualsiasi realtà esterna alla mente diventa nostra (Zingale 2020).

«We have no power of thinking without signs» (Peirce 1932: CP 5.265). Osservando umani generati o manipolati, attingiamo all'umano ideale di riferimento ed erroneamente diamo per scontata la sua veridicità ogni qualvolta ci si presenta di fronte.

Più specificatamente, vedendo Tom Cruise generato, stabiliamo una connessione fra la sua persona reale "A" e sua rappresentazione in chiave deepfake "B", ovvero il segno, il feno-

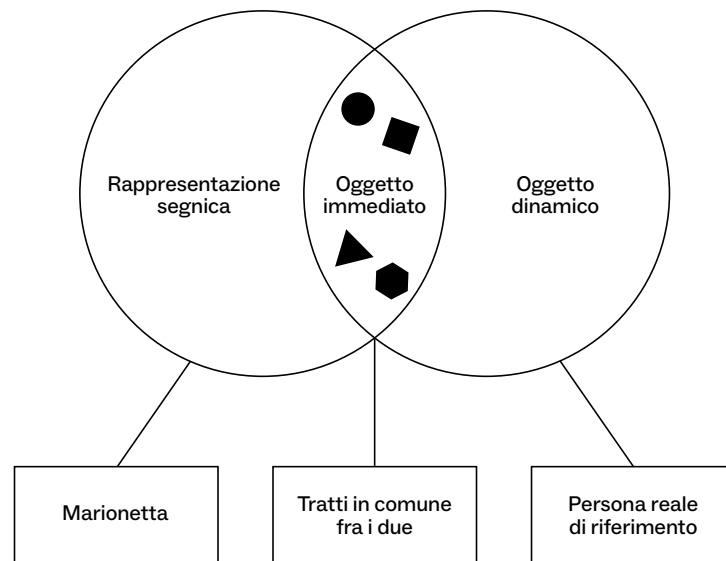
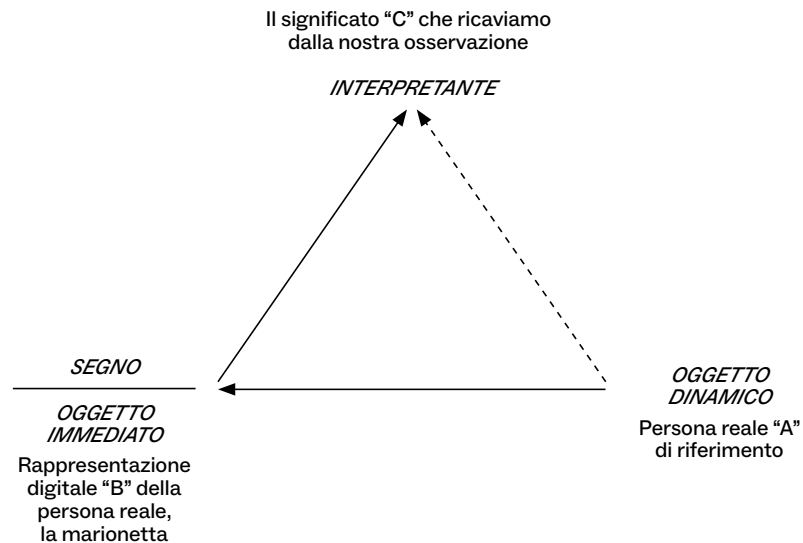
meno, un *oggetto immediato* così come viene rappresentato; producendo un terzo fenomeno “C”, l'*interpretante*, che rappresenta l'unione fra i due, una conseguenza apparentemente logica che si forma nella nostra mente e il significato che ricaviamo dalla nostra osservazione (v. fig. 5.1). La risposta che la nostra mente dà può essere di diversa natura, e molto spesso nel mondo online è di tipo emozionale (Zingale 2020).

Il processo semiotico che scatta in noi è il medesimo di ogni altra osservazione, ma questa volta la nostra mente è tratta in inganno. Il contenuto di video o immagini con marionette fa spesso leva, volontariamente o involontariamente, su una risposta emotiva dell'osservatore, effettuando una manipolazione non solamente sul contenuto digitale in sé ma, di conseguenza, anche sulla mente della persona. Ciò è dovuto, in primis, alle tematiche trattate da questi contenuti, dal carattere politico, religioso o spesso controverso, che tendono a polarizzare le opinioni.

La marionetta, considerata come *oggetto immediato* determinato dal *segno*, è in una mediazione di *somiglianza iconica* con l'umano di riferimento. Questo significa che le due figure, una reale e l'altra sintetizzata, presentano dei tratti pertinenti in comune, delle qualità che rendono entrambi riconoscibili e simili uno con l'altro. Queste qualità appartengono alla sfera visuale, uditiva ma anche comportamentale delle due persone (v. fig. 5.2) (Zingale 2020).

So in contemplating a painting, there is a moment when we lose the consciousness that it is not the thing, the distinction of the real and the copy disappears, and it is for the moment a pure dream – not any particular existence, and yet not general. At that moment we are contemplating an icon. (Peirce 1932: CP 3.362).

Peirce mette per iscritto la sensazione che proviamo quando osserviamo un dipinto realistico e, per un attimo, dimentichiamo come esso non sia in realtà la cosa che rappresenta. La distinzione fra reale e copia scompare, ci troviamo in un sogno dove contempliamo un'icona che in qualche modo inganna la nostra mente. Le marionette non sono altro che dei ritratti nell'era dell'infocalypse, delle rappresentazioni iconiche in tutto e per tutto simili alla loro controparte reale. Solitamente però, hanno maggior rilevanza, influenza e diffondibilità verso il pubblico quando sfruttano sia le somiglianze che le differenze, facendo compiere al soggetto rappresentato delle azioni fuori dal comune che siano in contrasto con il comportamento della persona di riferimento, oppure che esasperino una tendenza appartenente alla stessa (v. fig. 5.3, 5.4).



↑ FIG. 5.1 TRIANGOLO SEMIOTICO
 ↑ FIG. 5.2 RAPPORTO DI *MEDIAZIONE SEGNICA PER SOMIGLIANZA*, NELL'INTERSEZIONE VI SONO I CONNOTATI IN COMUNE FRA LA PERSONA DI RIFERIMENTO E L'OUTPUT SINTETICO
 → (PAG. SUCC.) FIG. 5.3, 5.4 IMMAGINI VIRALI DI PAPA FRANCESCO GENERATE CON MIDJOURNEY, 2023



NULLA È VERO NEL SUPERSTRATO



RAPPORTO CON LE ALTERITÀ GENERATE

5.3

APATIA DELLA REALTÀ, POSTVERITÀ E IPERREALE

Se la nostra realtà è popolata da umani generati, fake news, contenuti di disinformazione e misinformazione, come possiamo distinguere ciò che è reale, o meglio, affidabile, e ciò che non lo è?

«Ora, infatti, certi sistemi computazionali sono dotati – noi li abbiamo così dotati – di una singolare e inquietante vocazione: enunciare la verità» (Sadin 2019: 8).

Con questo incipit vorrei affrontare quindi quello che può essere realmente l'effetto che i nuovi media sintetici e la diffusione di questi su piattaforme online potranno avere sul nostro esistere nella società. Già ora è possibile rintracciare una *neo-verità*, un nuovo regime di verità conseguente all'azione degli algoritmi sui social network e nella generazione di contenuti, un regime che mai nella storia si è imposto in questo modo. La neo-Verità si basa sulla convinzione che i fatti che ci vengono presentati dagli algoritmi siano così evidenti da risultare veri e affidabili. Questo perché i nostri sistemi cognitivi, ovvero la nostra capacità di conoscere e comprendere il mondo, sono unicamente fondati sull'evidenza (cfr. Sadin 2019: 58–59).

Noi ci fidiamo del consiglio dell'algoritmo anche quando non verificiamo ciò che ci si presenta davanti, perché non abbiamo ancora i mezzi per distinguere la veridicità dell'artefatto. Ormai assistiamo compiaciuti all'agonia della nostra "coscien-

za infelice", per vedere il sopraggiungere dell'era del "Sapere assoluto" descritto da Hegel come il momento in cui «la verità è in sé perfettamente uguale alla certezza [...] e per evitare di "cadere nell'abisso spaventoso dello scetticismo" tentiamo continuamente di afferrare una verità indubbia» (Sadin 2019: 59).

Essendo umani, ognuno di noi cade quotidianamente in questa trappola. C'è chi abbraccia il fenomeno inconsciamente, partecipando addirittura alla sua espansione, e c'è chi invece ritiene, erroneamente, di non esserne affetto. Abbiamo osservato però in 3.1 che già prima e dopo la nascita della fotografia, l'essere umano ha plasmato e manipolato la produzione visuale e contenutistica. Come i nostri sistemi cognitivi fondati sull'evidenza, anche questa indole ritengo sia insita nelle nostre caratteristiche biologiche.

È importante però sottolineare come, vivere nel purgatorio online in cui ci troviamo ora, possa portare a dei fenomeni sociali mai affrontati nella storia. In primo luogo, solamente l'esistenza di una tecnologia in grado di sintetizzare tutto ciò che compone il mondo in cui viviamo, rendendo l'irreale reale, comporta anche, d'altra parte, una possibile negazione a prescindere dell'accaduto. «You don't need to create the fake video for this tech to have a serious impact. You just point to the fact that the tech exists and you can impugn the integrity of the stuff that's real» (Warzel 2018). Il fenomeno si chiama *liar's dividend*, e Donald Trump ne è stato ancora una volta portavoce durante la sua campagna politica. Questo fenomeno infatti rappresenta il vantaggio che alcuni individui possono trarre sfruttando il caos di informazioni non verificabili online, facendo leva sull'esistenza di strumenti sfruttati spesso per fare del male e utilizzando questo pretesto a proprio favore per "farla franca". Trump ha spesso negato, infatti, l'integrità di contenuti a lui attribuiti. Basti pensare a quante volte ha menzionato il termine "fake-news" durante i suoi discorsi pubblici e nei suoi post su Twitter, oppure come abbia liquidato come "false" delle registrazioni audio comprovate e compromettenti risalenti ai primi anni Duemila (cfr. Schick 2020: 61).

L'integrità del sistema è quindi a rischio senza nemmeno che le tecnologie di IA entrino effettivamente in gioco, ma soprattutto, non è nemmeno più necessario che una notizia sia realmente vera o falsa per consentire la sua negazione. Questo è sintomo di un progressivo e imperterritito scetticismo per la verità, che tende a superare ciò che abbiamo descritto in precedenza con la neo verità, e riguarda la sfera contenutistica.

Sarebbero necessarie infatti poche grandi bufale, simulazioni sintetiche altamente fotorealistiche poi svelate all'opinione pubblica, per convincere quest'ultima che nulla è reale. La gente potrebbe infatti iniziare a sospettare che non vi sia alcuna veridicità di fondo negli artefatti mediatici e che non ci si possa

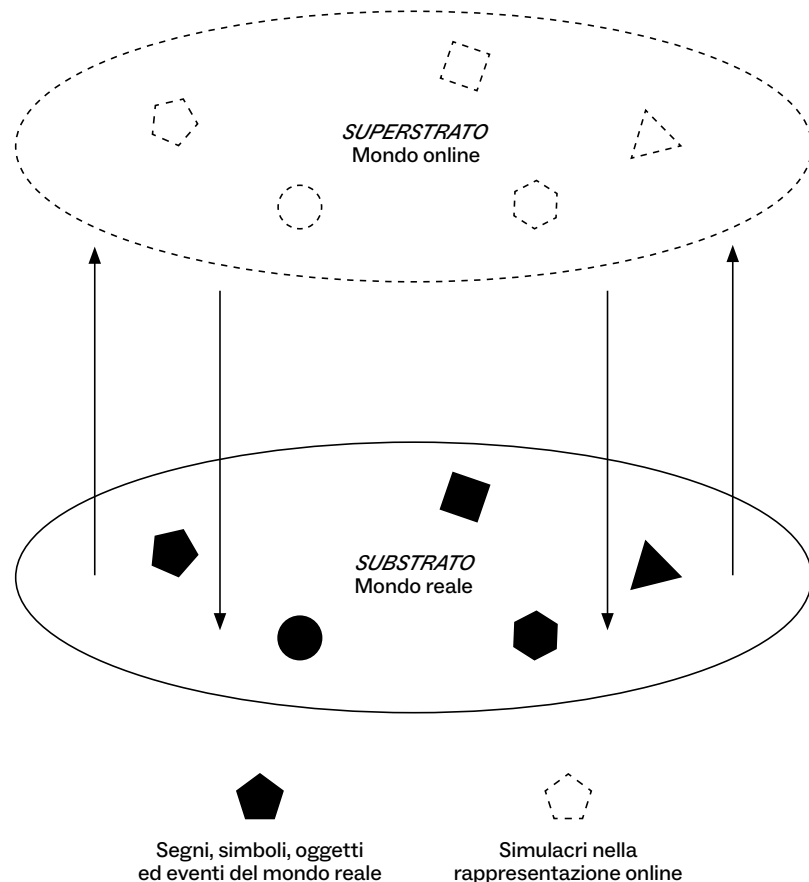
fidare di essi, iniziando a sviluppare un'apatia della realtà, una sfiducia a priori del sistema mediale attuale. Prevedere questo fenomeno è estremamente complesso, e anzi è possibile che, grazie all'influenza di attori consapevoli e non, ci si stia avvicinando a uno scenario di *postverità*. (cfr. Warzel 2018).

Jean Baudrillard definisce con il termine *iperreale* un'esistenza più reale della vita reale stessa, costituita da *simulacri*, ovvero immagini multimediali ed elementi di realtà virtuale, condivisi e promossi da e su i mass media, e ora anche sui social media. L'iperreale è quindi il risultato dell'espansione della cultura dei media e della simulazione nella società contemporanea, e si riferisce alla generazione di modelli di un "reale" senza origine né veridicità. Il mondo dell'iperreale è completamente legato dai fatti, e sono gli stessi simulacri a comporre lo scenario della nuova realtà, una postverità (Baudrillard 1981).

Segni e simboli, interpretabili in questo caso come i contenuti sintetici generati da attori umani, tracciano lo scenario dell'iperreale e producono una condizione in cui il mondo diventa una rappresentazione di sé stesso, dove la distinzione fra realtà e finzione svanisce. Nel mondo mediale, la rappresentazione di un evento è più importante dell'evento stesso: la generazione di un superstrato digitale costituisce una realtà parallela plasmabile a piacere, spesso a scopo di lucro, tanto da diventare più rilevante del mondo reale stesso e sovrastare la nostra percezione degli eventi (*ibid.*).

We are now living in the era of post-truth, and, I believe, it is on this concept that a "new" idea of reality should be based on. Real should be referred to what has a meaning for people. [...] Real is correlated to reliability and authenticity, to trust and truth. And so, the way of saying "seeing is believing" appears to be too limited now: we see too many things within our communications, but it's not taken for granted that we believe to all of these. We have to learn to render the truth in order to process what we consider real. (Castelli 2020:146)

Abbiamo parlato di fiducia negli algoritmi e di una conseguente sfiducia in ciò che i media ci propongono, ma Castelli ci pone una riflessione su quello che potrebbe essere il nostro approccio, anche involontario, alla postverità. In un mondo dove il vero e il falso non solo saranno indistinguibili e il "vedere per credere" non avrà più potere, ma la loro distinzione sarà anche irrilevante, il reale diventa quindi ciò che ha un significato per le persone. Come utenti ci troveremo quindi a processare internamente la verità per poi elaborare ciò che consideriamo reale.



† FIG. 5.5 RAPPRESENTAZIONE DEL RAPPORTO FRA *SUPERSTRATO* E *SUBSTRATO*. SEGNI E I SIMBOLI APPARTENENTI AL MONDO REALE VENGONO PROIETTATI NEL SUPERSTRATO, MANTENENDO CONNOTATI RICONOSCIBILI MA PERDENDO CONTATTO CON LA REALTÀ A CAUSA DELLA FACILITÀ DI MANIPOLAZIONE

5.4

RICONOSCERE I CONTENUTI: IL NOSTRO COMPITO

Ma cosa possiamo fare da utenti e come invece dovremo adattare le nostre modalità interpretative in futuro? Già oggi risulta estremamente complesso stilare un metodo per categorizzare ciò che è reale e ciò che non lo è. Non vi è infatti una modalità univoca per scovare un contenuto sintetico, ma quelli che definirei una serie di *indizi* che sommati producono delle prove. Se allenati progettualmente e tecnicamente a osservare contenuti in modo critico, possiamo infatti ancora sospettare facilmente di video e immagini dove il lip-synching non è perfetto o dove sono rappresentati solamente personaggi noti e situazioni troppo assurde per essere vere.

Riportando nuovamente la mia personale esperienza su TikTok, mi è stato possibile identificare dei pattern contenutistici abbastanza chiari nonostante la dilagante disinformazione presente sulla piattaforma. Questi contenuti rappresentano frequentemente persone del mondo dello spettacolo e dell'attualità che in parte abbiamo già potuto incontrare in questa tesi: Elon Musk, Joe Rogan, Jeff Bezos, Donald Trump, Joe Biden e qualunque altro personaggio trending del periodo storico di riferimento. Spesso questi personaggi, nei video in cui sono replicati, intavolano discorsi ai limiti del satirico o dell'improbabile, facendo scaturire più di qualche dubbio se visualizzati più e più

volte durante il corso della settimana o addirittura della giornata. «Our friend applied one of the most important rules for spotting bullshit: *If something seems too good or too bad to be true, it probably is*» (Bergstrom e West 2020: 133).

Questa regola universale, "se qualcosa sembra troppo bello o troppo brutto per essere vero, probabilmente lo è", funziona particolarmente bene quando ci stiamo interfacciando con contenuti poco o malamente progettati, come dei cheapfake o dei video sintetizzati, prodotti dal basso e senza pretese. Gli autori distinguono infatti due tipologie differenti di contenuti: *bugie* e *stupidaggini*. Le bugie, solitamente ben progettate e pensate, sono concepite per allontanare l'utente dalla verità; le stupidaggini sono frutto di una grossolana indifferenza per la verità. Secondo questa distinzione quindi, gli output più frivoli e mal pensati risultano estremamente più facili da individuare, poiché ridicoli anche solamente all'apparenza (cfr. Bergstrom e West 2020: 295).

Riconosco tuttavia un inevitabile imprecisione del criterio sopra citato: per una fetta di popolazione apparentemente allenata al pensiero critico, corrisponde una maggioranza degli utenti poco o scarsamente coscienti del potere degli algoritmi di machine learning. Vi è una chiara asimmetria informativa che non permette alle persone di valutare l'utilizzo effettivo che fanno dei social media. Oltre a ciò, c'è una scarsa trasparenza da parte delle piattaforme, colpevoli di non comunicare i dati raccolti dagli utenti stessi e di spiegare il perché mostrino un contenuto piuttosto che un altro. Tutto questo sfocia in un comportamento delle masse confusionario e dalle conseguenze non immediatamente verificabili, ma anzi di un protrarsi delle ripercussioni negative, mentre gli utenti stessi sono lasciati in balia degli algoritmi e delle marionette (von der Weth et al. 2020).

Se non per particolari occasioni infatti, non verifichiamo mai la fonte dei contenuti, soprattutto se questi ci vengono proposti in rapida successione uno dopo l'altro. È come se le nostre difese immunitarie non risultino ancora abbastanza forti verso la sovrabbondanza di virus-informazione. Per sviluppare queste difese è necessario instaurare un meccanismo di selezione interno degli input, un pensiero critico ed etico che diventa un filtro fra il mondo là fuori e la nostra mente, un cambio di paradigma e una messa in discussione che tiene conto del nostro esistere in questi mondi virtuali. Conoscere metodi progettuali etici ma soprattutto l'aver progettato per la diffondibilità dà sicuramente un vantaggio, poiché permette di individuare pattern e fondamenta invisibili a occhio nudo.

Sicuramente lo scetticismo può aiutare. Come nello scenario che precede l'apatia della realtà in 5.3, essere scettici e avere un pensiero critico è fondamentale, ma servirebbe una maggiore educazione sulle possibili campagne di disinfor-

FAKE MEANS CARRIER OF MISLEADING INFORMATIONS

(CASTELLI 2022)

NULLA È VERO NEL SUPERSTRATO

mazione che possono essere messe in atto, su come è possibile manipolare una fotografia o un contenuto audiovisivo e sulle principali tecniche attuate *dall'alto*, senza necessariamente prendere una posizione politica (Bergstrom e West 2020: 8). Molto spesso chi appartiene a community underground possiede o sviluppa una polarizzazione politica molto forte, che enfatizza pericolosamente il suo scetticismo e sfocia nell'odio a prescindere verso la comunicazione attuata dai media tradizionali e da individui dalle ideologie opposte. Per questo anche l'educazione accademica o *super partes* in questo ambito è fondamentale, non solamente da un punto di vista metodologico ma anche sociale.

La relazione fra mondo reale e mondo online parte, nella storia, dalla condivisione di informazioni fra l'uno e l'altro, ma tende sempre di più a differenziarsi man mano che gli attori al suo interno riescono a plasmare il nostro modo di percepire gli eventi.

Our virtual experience has to be seen as part of a novel cultural framework, which we need to redefine and rethink more comprehensively. Accepting it with fluidity and consciousness is the key for not being anchored to old preconceptions. In fact, the virtual realm is interpreted now as another cultural and social context where one has to adopt a new set of skills in order to interpret what is given by the various platforms and interactions [...] Physical world and virtual one are nowadays both part of our reality, just with different set of rules and different type of representations. The same discourse, turned upside down, works for trying to define what is then fake. Fake is not synonymous of artificial or simulated. Fake means carrier of misleading informations, not genuine contents and not reliable sources. It has an implicit condition that we are learning to identify while interacting and communicating.
(Castelli 2020)

Vivere contemporaneamente nel mondo fisico e nel mondo virtuale fa parte della nostra quotidianità; non venire schiacciati dalla quantità di informazioni si traduce nell'abbracciare uno scenario di postverità nel quale osserviamo con fluidità ciò che ci viene proposto e sviluppiamo un nuovo set di competenze in grado di aiutarci nell'interpretazione dei contenuti. In questa nuova realtà, il "falso" non è più il contenuto generato artificialmente e il "vero" non è più la fotografia scattata e non post prodotta che ritrae un evento accaduto. Il falso è ciò che porta con sé disinformazione o misinformazione, condiviso a prescindere dal procedimento di generazione attuato.

RICONOSCERE I CONTENUTI: IL NOSTRO COMPITO

5.5 CONTROMISURE DALL'ALTO

Vorrei approfondire infine un ultimo aspetto che deve necessariamente affiancarsi al lavoro che come individui dobbiamo portare avanti. Serve infatti un'azione di responsabilità diffusa, proveniente dall'alto, ovvero dalle aziende stesse che sviluppano questi modelli, dalle piattaforme social e dai governi. Attualmente si sta portando avanti un costante sviluppo di sistemi di rilevamento automatici dei media sintetici. Nei modelli GAN dei deepfake, ad esempio, vi è lo scontro fra due reti neurali, una delle quali è adibita al riconoscimento dell'accuratezza dell'output: questa può essere allenata per imparare a scovare la manipolazione visiva di fotografie.

La *DARPA (Defense Advanced Research Projects Agency)*, un'agenzia governativa del Dipartimento della Difesa degli Stati Uniti) ha investito massivamente nella causa con l'apertura di programmi da milioni e milioni di dollari. Facebook, da parte sua, ha finanziato una *Deepfake Challenge Competition* in cui i partecipanti si sono sfidati nello sviluppo di modelli per il rilevamento di contenuti deepfake, raggiungendo però solamente il 65% di accuratezza. Lo sviluppo di questi contro-software deve seguire l'evoluzione dei modelli di manipolazione, che diventano sempre più raffinati e precisi nell'esecuzione tecnica. Purtroppo però, gli attori che li utilizzano sono decisamente in vantaggio rispetto a chi rincorre: le stesse contromisure aiutano infatti i creatori di contenuti a correggere il tiro e migliorare i propri modelli.

Un esempio è come nel 2018 sia stato presentato uno studio in cui è stato indicato come le persone rappresentate nei video chiudessero le palpebre con una velocità differente rispetto agli umani reali. Nel giro di poche settimane le commu-

nity hanno notato queste considerazioni e tutti i nuovi contenuti deepfake sono stati corretti (cfr. Helmus 2022).

«There is a “fundamental mathematical limit to the ability of a given detector to distinguish between real and synthetic images”» (ivi: 2022). Man mano che gli output diventeranno sempre più accurati e indistinguibili per noi, lo saranno conseguentemente anche per gli algoritmi di rilevamento, raggiungendo un punto limite dove nemmeno la macchina potrà rilevare una manipolazione. Capiamo quindi che non può esistere una strada univoca per affrontare questo scenario ma che si debba agire su più fronti. I social media, ad esempio, potrebbero fornire database di immagini classificate come disinformative o misinformative nei loro archivi, con lo scopo di addestrare i modelli.

Le stesse piattaforme, da questo punto di vista, hanno già proposto alcune soluzioni che vanno a mitigare il problema, almeno per quanto riguarda contenuti appartenenti a tematiche scottanti. Instagram, Tiktok, Facebook: molti hanno implementato dei *disclaimer* o finestre *pop-up* da inserire in overlay su post con immagini o video controversi, dove l'utente è invitato a informarsi sull'argomento (v. fig. 5.6).



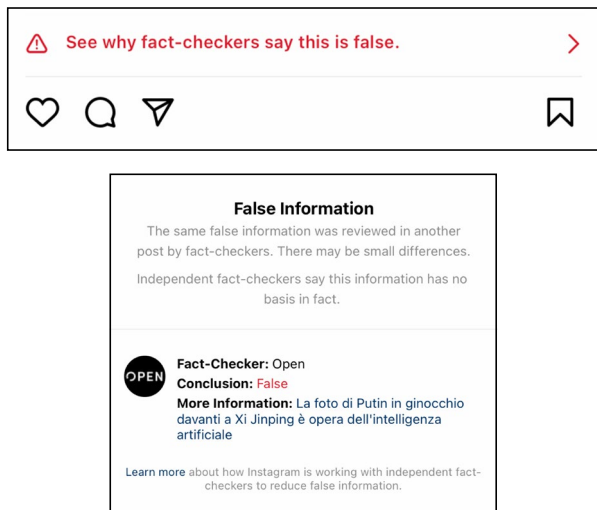
YouTube, con la rubrica *Premi Pausa*, intende educare il visitatore medio al riconoscimento della manipolazione audiovisiva, mostrando delle finestre evidenziate direttamente nella homepage che portano a una playlist video (v. fig. 5.7).

Meta inoltre affida il compito di categorizzare determinate informazioni come false, alterate o senza contesto a figure esterne denominate *fact-checkers*, degli attori umani indipendenti certificati dall'*International Fact-Checking Network*, che hanno il ruolo di verificare la veridicità o meno di un contenuto mediale, svolgendo interviste alle fonti primarie di informazione, consultando dati pubblici e analizzando l'aspetto del contenuto stesso (v. fig. 5.8 e 5.9).

↑ FIG. 5.6 OVERLAY PRESENTE SUI VIDEO DI TIKTOK CHE TRATTANO
TEMATICHE INERENTI AL COVID-19

↑ FIG. 5.7 FINESTRA DELLA RUBRICA *PREMI PAUSA* NELLA HOMEPAGE DI YOUTUBE

Una volta contrassegnato il contenuto, ne viene ridotta la distribuzione tramite algoritmi e viene etichettato su tutte le repliche in post affini, ma non viene rimosso. Viene poi inviata una notifica agli amministratori della relativa pagina, che possono consultare le informazioni trovate dai fact-checkers (Meta 2023). La verifica dei contenuti da terze parti aiuta sicuramente a decentralizzare la percezione di monopolio che gli utenti hanno verso queste piattaforme, deresponsabilizzando l'azienda nella scelta di etichettare o meno un contenuto, o addirittura rimuoverlo.



I fornitori di servizi generativi come OpenAI stanno sviluppando dei *watermark* incorporati negli output sintetici. Similmente a *Getty Images*, piattaforma di fotografie stock, che applica un rettangolo ben visibile su ogni immagine, anche DALL-E 2 pone una piccola striscia colorata nella parte inferiore destra dei risultati (v. fig. 5.10). Il posizionamento del watermark ha la duplice funzione di proteggere la proprietà intellettuale dell'opera ma anche di esplicitare l'origine del contenuto soprattutto nel caso di DALL-E 2. Questa contromisura è, tuttavia, facilmente aggirabile tagliando l'immagine o effettuando un piccolo ritocco su Photoshop. Per questo nel corso del tempo si è puntato su altri metodi, che prevedono, ad esempio, il cambiamento del decimo pixel di ogni immagine con un valore compreso fra lo 0 e il 255. OpenAI è al lavoro anche per rendere identificabili sequenze di parole composte da ChatGPT e rintracciare un pattern nascosti (Farid 2023). Stable Diffusion utilizza già una

↑ FIG. 5.8 OVERLAY PRESENTE SUI POST DI INSTAGRAM REVISIONATI DAI FACT-CHECKERS

↑ FIG. 5.9 FINESTRA POP-UP SUCCESSIVA AL TAP DELL'OVERLAY

libreria Python chiamata *invisible-watermark* per nascondere all'occhio umano la presenza di un marchio. I due algoritmi utilizzati decompongono i segnali delle immagini in differenti bande di frequenza, utilizzate dalla macchina per scegliere le aree dell'immagine più adatte. Generalmente i cambiamenti nelle zone in cui vi sono i bordi o delle texture, che corrispondono a una frequenza d'onda media, risultano poi invisibili all'occhio umano e perfetti per incorporare il watermark, il quale mantiene una resistenza a processi di editing anche pesanti della foto ma non è ancora completamente immune a tagli, rotazioni e ridimensionamenti (Steins 2022).

Dall'altro lato, alcune aziende stanno lavorando allo sviluppo di uno standard per etichettare i contenuti reali e incorporare delle informazioni nei metadati nelle fotografie quando queste vengono scattate, ad esempio dal nostro smartphone o da una macchina fotografica. Questa tecnologia, frutto della *Coalition for Content Provenance and Authentication (C2PA)* in collaborazione con numerose aziende quali *New York Times*, *Adobe* e *Microsoft*, *Intel*, *BBC*, creerebbe delle crittografie, delle firme digitali che garantiscono la provenienza "umana" del contenuto (Farid 2023).

Oltre a ciò, non iniziano a tardare contromisure da istituzioni governative, più o meno radicali, per la regolamentazione di tali strumenti. Nel dicembre 2022 la *Cyberspace Administration of China* ha bandito la condivisione di qualsiasi medium sintetico senza watermark, intimando ai fornitori di servizi di rafforzare la gestione dei dati e a vietare il trattamento illegale delle informazioni personali, ripensando il modo di agire dei loro algoritmi. Queste decisioni sono in parte utili per controllare il modus operandi dei provider, anche se attualmente risultano difficilmente applicabili nella loro interezza, soprattutto per quanto riguarda l'utente finale, a cui non è precluso l'editing o la manomissione dell'output stesso (Karahde 2022).

Le iniziative statali sono indispensabili per qualsiasi innovazione tecnologica e chiudono il cerchio aperto in questi ultimi due capitoli. Come ogni parte del puzzle contribuiscono alla battaglia contro la disinformazione online, ma non rappresentano una risposta assoluta al problema.



↑ FIG. 5.10 WATERMARK PRESENTE SULLE IMMAGINI GENERATE DA DALL-E 2

06

CONCLUSIONI

La riflessione che ho portato mira a rintracciare e spiegare come pattern insiti nella nostra natura possano influenzare, ancora una volta, l'utilizzo malevolo di strumenti di intelligenza artificiale, e come noi creatori di contenuti dovremmo affrontare queste tendenze nel nuovo *superstrato* digitale. Nel mio percorso, ho scelto di analizzare un lasso temporale che parte dagli strumenti disponibili, proiettando criticità e possibilità a un futuro prossimo. Ritengo necessaria una presa di coscienza repentina dei sintomi già rintracciabili nella società online di oggi, oltre a un'azione immediata. Gli scenari descritti in questa tesi sono volutamente presentati come vicini a noi, poiché già accaduti nella storia: l'intelligenza artificiale è infatti solamente un componente dell'infocalypse.

Queste tecnologie portano con sé dei connotati intrinseci che ci rendono inclini a un'esagerata estetizzazione dell'output a scapito di uno sviluppo del pensiero progettuale. Il cambio di approccio verso la nostra relazione con la macchina, in un'ibridazione quasi trasparente, abbatte le barriere presenti in precedenza nel mondo della creazione di contenuti. Come per il passaggio fra pittura e fotografia o fra progettazione grafica analogica e software, l'introduzione di modelli IA *text-to-image*, *text-to-speech*, *text-to-text* e così via ha ridotto drasticamente i tempi di produzione, facilitando la partecipazione e l'approccio da parte di attori non-progettisti o semplicemente malintenzionati.

Ho ritenuto necessario sottolineare con una descrizione *step by step* come sia possibile agire attualmente per generare le marionette, delle copie carbone sufficientemente perfette dell'essere umano in tutti i suoi connotati, riproducibili su sup-

porti audiovisivi condivisibili online. Nonostante il costante progresso della tecnologia, è fondamentale conoscere le possibilità che ci fornisce lo stato dell'arte attuale per essere in grado di avviare il nostro processo di acquisizione di una consapevolezza progettuale. Il poter replicare aspetto fisico, movenze, voce e addirittura idee, per poi trasportarle su piattaforme digitali che ci permettono una diffondibilità senza precedenti, anche a causa degli algoritmi stessi che gestiscono i social media, apre a un duplice problema. Nelle mani di fautori e di manipolatori è facile capire come gli strumenti possano essere utilizzati per spargere disinformazione, mentre nelle mani di utenti non informati o inconsapevoli, appartenenti a community, o sottoboschi, sono l'ingrediente per uno scenario di caos informativo e di misinformazione dilagante.

Il mondo online è una sorta di *rifugio per la mente*, chi ne partecipa si sente inglobato ed è quasi assalito da una sensazione di onnipotenza e controllo, dove ogni fantasia è realizzabile (Marchesini 2002: 394).

Il Web è l'esaltazione dell'atopia, del galleggiamento amniotico, dell'isocronia che annulla qualsiasi evoluzione e permette al presente di rendersi flusso, accadimento che non lascia traccia, e quindi di abbracciare in una sorta di contemporaneità passato e futuro. Internet è il luogo di ciò che è senza luogo: consente di viaggiare rimanendo nel nido, di evitare cioè di spiccare quel volo che recide il cordone ombelicale con l'infanzia. (Marchesini 2002: 394)

Vivere e *partecipare* online è quindi un pretesto inconscio che le persone adottano per tornare a uno stato infantile, a dei luoghi virtuali dove esistono gerarchie differenti dalla realtà e dove possono esprimersi senza filtri, dando sfogo a ogni desiderio, anche il più perverso o spaventoso, e costruire inconsapevolmente le fondamenta dell'infocalypse.

La scissione tra mondo reale e mondo virtuale è ora più che mai visibile. Quest'ultimo è come un superstrato imposto sulla nostra cultura del reale e comporta uno scenario illusorio dove la realtà è sintetizzata sui media tramite segni e simboli. Qui ogni contenuto è plasmabile, personalizzabile e generabile a piacimento, sia per puro desiderio personale che per fini manipolativi. Nel superstrato i media agiscono creando uno scenario iperreale dove il valore di un determinato evento è superato dalla sua rappresentazione, dove i contenuti fruiti online sono dei simulacri perfettamente confondibili con la controparte reale. Avvicinandoci al superstrato, ci avviciniamo a una realtà parallela dove non abbiamo più una percezione oggettiva

va degli eventi. Il mondo reale di riferimento non viene soppiantato, ma diventa via via meno influente una volta che la rappresentazione dell'informazione, condivisa sui media, si proietta sullo strato superiore.

Ritengo che per prepararsi al meglio per questo scenario, o cercare di evitarlo, sia necessario abbracciare alcuni ideali postumanisti esposti nel capitolo 1. Il *cyborg* descritto da Haraway (1985) è fluido e *super partes*, trascende il pensiero dicotomico umanista che vede l'*io* al di sopra dell'*altro*, la perfezione rappresentativa e l'estetica al di sopra dell'ascolto della diversità e delle esigenze esterne.

Il designer della comunicazione deve essere, metaforicamente, un cyborg. Progettare correttamente significa abbracciare le alterità, farle proprie e sapere sempre la propria collocazione nel contesto di riferimento. Osservare il mondo circostante, ibridarsi con la tecnologia e gli strumenti ma non farsi trasportare da essi, avere il controllo sul processo, essere consapevoli delle conseguenze e sapere come diffondere un risultato in modo etico: tutte queste sono caratteristiche fondamentali per lottare contro la disinformazione e progettare in un futuro dove vero e falso coincidono. È di fondamentale importanza saper scindere quelle che sono le nostre convinzioni e i nostri principi dalle informazioni che, pur contrastando con il nostro pensiero, possono risultare utili per riconoscere la natura di un media sintetico. Non è sicuramente una cura infallibile, ma è un punto di partenza imprescindibile che noi progettisti e utenti del mondo online dobbiamo considerare.

BIBLIOGRAFIA

- Anders, Günther. 2022. Considerazioni sull'anima nell'epoca della seconda rivoluzione industriale. Ristampa. Torino: Boringhieri.
- Baudrillard, Jean. 1981. Simulacra and Simulation. Ann Arbor, Mich: Univ. of Michigan Press.
- Baule, Giovanni, e Elena Caratti, a c. di. 2016. Design è traduzione: il paradigma traduttivo per la cultura del progetto. Milano: FrancoAngeli.
- Bergstrom, Carl T., e Jevin D. West. 2020. Calling Bullshit: The Art of Skepticism in a Data-Driven World. First edition. New York: Random House.
- Bianchi, Leonardo. 2021. Complotti! Da QAnon alla pandemia, cronache dal mondo capovolto. Minimum Fax.
- Boden, Margaret A. 2018. Artificial Intelligence: A Very Short Introduction. Oxford: Oxford University Press.
- Bonanata, Antonio. 2023. «Il fake della foto di Putin inginocchiato ai piedi di Xi e i rischi dell'intelligenza artificiale». RaiNews. Recuperato 22 marzo 2023 (<https://www.rainews.it/articoli/2023/03/fake-foto-putin-inginocchiato-piedi-xi-jinping-rischi-intelligenza-artificiale-9db959c8-8eab-4bfb-bbcf-f69863065cee.html>).
- Byock, Ira. 2012. The Best Care Possible: A Physician's Quest to Transform Care Through the End of Life. Thorndike Press.
- Cambridge Dictionary. 2023. «NSFW». Recuperato 24 marzo 2023 (<https://dictionary.cambridge.org/it/dizionario/inglese/nsfw>).
- Castelli, Ilaria. 2020. «Simulacra. A Speculative Journey among Ethics, Aesthetics and Emerging Technologies, towards a Post Human Future». Politecnico di Milano, Milano.
- Centini, Massimo. 1999. La sindrome di Prometeo: l'uomo crea l'uomo : dalla mitologia alle biotecnologie. Milano: Rusconi.

Cole, Samantha. 2019. «Deepnude: The Horrifying App Undressing Women». Recuperato 23 febbraio 2023 (<https://www.vice.com/en/article/kzm59x/deepnude-app-creates-fake-nudes-of-any-woman>).

Cortex. 2022. «AI Will Make Marionettes Of Us All Before It Destroys The World». YouTube. Recuperato 16 marzo 2023 (<https://www.youtube.com/watch?v=2pr3thuB10U&t=124s>).

Cross, Katherine. 2022. «It's Not Your Fault You're a Jerk on Twitter». Wired, febbraio 3.

Dzyaloshinsky, I. M. 2022. «Artificial Intelligence: A Humanitarian Perspective». Vestnik NSU. Series: History and Philology 21(6):20–29. doi: 10.25205/1818-7919-2022-21-6-20-29.

ElevenLabs. 2023. «ElevenLabs». Recuperato 3 marzo 2023 (<https://beta.elevenlabs.io/>).

Farid, Hany. 2023. «Watermarking ChatGPT, DALL-E and other generative AIs could help protect against fraud and misinformation». Recuperato 1 aprile 2023 (<https://theconversation.com/watermarking-chatgpt-dall-e-and-other-generative-ais-could-help-protect-against-fraud-and-misinformation-202293>).

Ferraresi, Mauro. 2023. «Creatività o inventiva? Riflessioni su ChatGPT». Facebook. Recuperato (<https://www.facebook.com/mauro.ferraresi.315/posts/pfbid02KjS9rFdFYbP19w59y9CZyvJrm7sqAGmHxeuEHvHjXAhwdv89xnPthDaBnVQd57Gal>).

Finn, Ed. 2017. What Algorithms Want. MIT Press.

Garcia, Chris. 2016. «Harold Cohen and AARON—A 40-Year Collaboration - CHM». Recuperato 26 gennaio 2023 (<https://computerhistory.org/blog/harold-cohen-and-aroon-a-40-year-collaboration/>).

Giorgio, Giovanni. 2016. «Postumano: il cyborg tra natura e cultura».

Haraway, Donna J. 1985. «Un manifesto per Cyborg: scienza, tecnologia e femminismo socialista nel tardo ventesimo secolo».

Helmus, Todd C. 2022. Artificial Intelligence, Deepfakes, and Disinformation: A Primer. RAND Corporation.

iFake. 2019. «The Irishman De-Aging: Netflix Millions VS. Free Software! - YouTube». Recuperato 2 marzo 2023 (<https://www.youtube.com/watch?v=dyRvbFhknRc>).

Ihde, Don. 1996. Technology and the Lifeworld: From Garden to Earth. Nachdr. Bloomington, Ind.: Indiana University Press.

Jenkins, Henry, Sam Ford, e Joshua Green. 2013. Spreadable Media: i media tra condivisione, circolazione, partecipazione. New York University.

Jourová, Věra. 2020. «“From Pandemic to Infodemic”». European Commission - European Commission. (https://ec.europa.eu/commission/presscorner/detail/en/speech_20_1000).

Kalpokus, Ignas. 2021. «Problematising Reality: The Promises and Perils of Synthetic Media». SN Social Sciences 1(1):1. doi: 10.1007/s43545-020-00010-8.

Karhade, Mandar. 2022. «China Bans AI-Generated Media (Again)». Medium. Recuperato 1 aprile 2023 (<https://medium.com/geekculture/china-bans-ai-generated-media-again-ca2f132b67f4>).

Karlova, Natasha A., e Karen E. Fisher. 2013. «A Social Diffusion Model of Misinformation and Disinformation for Understanding Human Information Behaviour». Recuperato 17 marzo 2023 (<https://www.informationr.net/jir/18-1/paper573.html>).

Kastrenakes, Jacob. 2020. «How photoshop became a verb - The Verge». Recuperato 17 febbraio 2023 (<https://www.theverge.com/2020/2/19/21143794/photoshop-30th-anniversary-adobe-verb-origin-story>).

LMU München. 2022. «Revolutionizing image generation by AI: Turning text in ... - LMU Munich». Recuperato 10 marzo 2023 (<https://www.lmu.de/en/newsroom/news-overview/news/revolutionizing-image-generation-by-ai-turning-text-into-images.html>).

Marchesini, Roberto. 2002. Post-human: verso nuovi modelli di esistenza. Torino: Bollati Boringhieri.

McCarthy, John, e Peter Wright. 2004. Technology as Experience. Cambridge, Mass: MIT Press.

McCorduck, Pamela. 2018. Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence. Boca Raton London New York: CRC Press.

McGlynn, Clare, Erika Rackley, e Ruth Houghton. 2017. «Beyond ‘Revenge Porn’: The Continuum of Image-Based Sexual Abuse». Feminist Legal Studies 25(1):25–46. doi: 10.1007/s10691-017-9343-2.

McLuhan, Marshall, e Quentin Fiore. 2011. Il Medium è il messaggio. Mantova: Corraini.

Merlan, Anna. 2020. «The Conspiracy Singularity Has Arrived». Vice. Recuperato 23 marzo 2023 (<https://www.vice.com/en/article/v7gz53/the-conspiracy-singularity-has-arrived>).

Meta. 2023. «Programma di fact-checking indipendente di Meta».

Michael Waters. 2017. «The Great Lengths Taken to Make Abraham Lincoln Look Good in Portraits - Atlas Obscura». Recuperato 16 febbraio 2023 (<https://www.atlasobscura.com/articles/abraham-lincoln-photos-edited>).

Nightingale, Sophie J., e Hany Farid. 2022. «AI-Synthesized Faces Are Indistinguishable from Real Faces and More Trustworthy». Proceedings of the National Academy of Sciences 119(8):e2120481119. doi: 10.1073/pnas.2120481119.

OpenAI. 2021. «OpenAI». OpenAI. Recuperato 28 gennaio 2023 (<https://openai.com/>).

Paris, Britt, e Joan Donovan. 2019. «Deepfakes and Cheapfakes». Peirce, Charles. 1932. *Collected Papers of Charles Sanders Peirce*, Volumes I and II. Vol. I&II. a cura di C. Hartshorne. Cambridge, MA: Belknap Press.

Podcast.ai. 2022. «Podcast.ai». Recuperato (<https://podcast.ai/>).

Rogers, Reece. 2023. «How You Can Tell the AI Images of Trump's Arrest Are Deepfakes». *Wired*, marzo 22.

Rosenbaum, Steven. 2019. «What Is Synthetic Media?» Recuperato 22 febbraio 2023 (<https://www.mediapost.com/publications/article/341074/what-is-synthetic-media.html>).

Sadin, Éric. 2019. *Critica della ragione artificiale: Una difesa dell'umanità*. I. LUISS University Press.

Schick, Nina. 2020. *Deepfakes: The Coming Infocalypse*. New York New York: Twelve.

Sequoia. 2022. «Generative AI: A Creative New World | Sequoia Capital US/Europe». Recuperato 28 gennaio 2023 (<https://www.sequoiacap.com/article/generative-ai-a-creative-new-world/>).

Shneiderman, Ben. 2003. *Leonardo's Laptop: Human Needs and the New Computing Technologies*. 1. MIT Press paperback ed. Cambridge, Mass.: MIT Press.

Silberling, Amanda. 2022. «Kickstarter Shut down the Campaign for AI Porn Group Unstable Diffusion amid Changing Guidelines». *TechCrunch*. Recuperato 26 marzo 2023 (<https://techcrunch.com/2022/12/21/kickstarter-shut-down-the-campaign-for-ai-porn-group-unstable-diffusion-amid-changing-guidelines/>).

Song, Yang, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, e Ben Poole. 2021. «Score-Based Generative Modeling through Stochastic Differential Equations».

Stability AI. 2022. «Stable Diffusion launch announcement — Stability AI». Recuperato 10 marzo 2023 (<https://stability.ai/blog/stable-diffusion-announcement>).

Steins. 2022. «Stable Diffusion — The Invisible Watermark in Generated Images». *Medium*. Recuperato 1 aprile 2023 (<https://medium.com/@steinsfu/stable-diffusion-the-invisible-watermark-in-generated-images-2d68e2ab1241>).

Synthesia. 2022. «Synthesia». Recuperato 3 marzo 2023 (<https://www.synthesia.io/>).

Taylor, Adam. 2021. «Before 'Fake News,' There Was Soviet 'Disinformation'». *Washington Post*, dicembre 1.

This Person Does Not Exist. 2023. «This-Person-Does-Not-Exist». *This-Person-Does-Not-Exist*. Recuperato 23 febbraio 2023 (<https://this-person-does-not-exist.com/en>).

Thorn, Abigail. 2022. «Transhumanism: "The World's Most Dangerous Idea" | Philosophy Tube». Recuperato 26 marzo 2023 (<https://www.youtube.com/watch?v=DqPd6MShV1o>).

Treccani. s.d.-a. «disinformazione in Vocabolario - Treccani». *Treccani.it*. Recuperato 17 marzo 2023 (<https://www.treccani.it/vocabolario/disinformazione/>).

Treccani. s.d.-b. «Haraway, Donna Jeanne nell'Enciclopedia Treccani». *Treccani.it*. Recuperato 21 gennaio 2023 (<https://www.treccani.it/enciclopedia/donna-jeanne-haraway>).

Treccani. s.d.-c. «Misinformazione in Vocabolario - Treccani». *Treccani.it*. Recuperato 17 marzo 2023 (https://www.treccani.it/vocabolario/neo-misinformazione_%28Neologismi%29/).

Treccani. s.d.-d. «sostrato in Vocabolario - Treccani». *Treccani.it*. Recuperato 2 aprile 2023 (<https://www.treccani.it/vocabolario/sostrato>).

Treccani. s.d.-e. «superstrato in Vocabolario - Treccani». *Treccani.it*. Recuperato 2 aprile 2023 (<https://www.treccani.it/vocabolario/superstrato>).

Treccani. s.d.-f. «tecnologia in Vocabolario - Treccani». *Treccani.it*. Recuperato 21 gennaio 2023 (<https://www.treccani.it/vocabolario/tecnologia/>).

Vales, Aldana. 2019. «An introduction to synthetic media and journalism | by Aldana Vales | WSJ Digital Experience & Strategy | Medium». Recuperato 22 febbraio 2023 (<https://medium.com/the-wall-street-journal/an-introduction-to-synthetic-media-and-journalism-cbbd70d915cd>).

Warzel, Charlie. 2018. «Believable: The Terrifying Future Of Fake News». *BuzzFeed News*. Recuperato 22 marzo 2023 (<https://www.buzzfeednews.com/article/charliewarzel/the-terrifying-future-of-fake-news>).

von der Weth, Christian, Ashraf Abdul, Shaojing Fan, e Mohan Kankanhalli. 2020. «Helping Users Tackle Algorithmic Threats on Social Media: A Multimedia Research Agenda».

Wiggers, Kyle, e Amanda Silberling. 2022. «Meet Unstable Diffusion, the Group Trying to Monetize AI Porn Generators». *TechCrunch*. Recuperato 16 marzo 2023 (<https://techcrunch.com/2022/11/17/meet-unstable-diffusion-the-group-trying-to-monetize-ai-porn-generators/>).

Xu, Jincheng. 2022. «Analysis of Social Media Algorithm Recommendation System». *Studies in Social Science & Humanities* 1(3). doi: 10.56397/SSSH.2022.10.06.

Yehya, Naief. 2017. *Homo cyborg. Il corpo postumano tra realtà e fantascienza*.

Zingale, Salvatore. 2020. «Appunti di Semiotica del Progetto A.A. 2020/2021». *Scuola del Design, Politecnico di Milano*.

INDICE FIGURE

FIG. 0.1	Rielaborazione dello schema di Karlova e Fisher, 2013. Sintesi delle caratteristiche di informazione, misinformazione e disinformazione. → Rielaborazione dell'autore. Karlova, Natasha A., e Karen E. Fisher. 2013. «A Social Diffusion Model of Misinformation and Disinformation for Understanding Human Information Behaviour».	15
FIG. 1.1	Femore rotto e guarito. → https://imglib.ru/works/bgtr8rqp/	23
FIG. 1.2	Copertina di A Cyborg Manifesto, Donna Haraway. → https://theanarchistlibrary.org/library/donna-haraway-a-cyborg-manifesto	25
FIG. 1.3	Diane arbus, Masked Child with a Doll, N.Y.C., 1961 → https://onlineonly.christies.com/s/photos-diane-arbus/masked-child-doll-n-y-c-1961-39/260	31
FIG. 2.1	Macchina Analitica, Charles Babbage e Lady Ada Lovelace, 1834. → https://www.historybit.it/timeline/macchina-analitica-di-charles-babbage/	37

FIG. 2.2	39
Prima iterazione di AARON, 1979. → https://cyberneticforests.substack.com/p/dalle-aaron-kata-strategies-for-legibility	
FIG. 2.3	40
AARON e Harold Cohen. → https://www.nytimes.com/2016/05/07/arts/design/harold-cohen-a-pioneer-of-computer-generated-art-dies-at-87.html	
FIG. 2.4	41
AARON nella fase di colorazione dell'output. → https://www.katevassgalerie.com/blog/harold-cohen-aaron-computer-art	
FIG. 2.5	42
Harold Cohen, AARON with Decorative Panel, 1992. → https://news.artnet.com/art-world/the-10-must-see-london-gallery-shows-during-frieze-week-2181699	
FIG. 2.6	47
Esempio di immagine generata con DALL-E 2 a partire da un input testuale → https://twitter.com/sama/status/1511715302265942024	
FIG. 2.7	47
Esempio della funzione inpainting di DALL-E 2, che permette di sostituire connotati di immagini già esistenti → https://twitter.com/AlexTamkin/status/1548423353496571904	
FIG. 2.8	51
Prompt testuale utilizzato su DALL-E 2 per generare la figura 2.6, declinato in differenti lingue → https://labs.openai.com/	
FIG. 3.1	57
Fotografia di Abraham Lincoln scattata da Mathew Brady poco prima del discorso alla Cooper Union, 1860 → https://commons.wikimedia.org/wiki/File:Photo_of_Lincoln_%28Cooper_Union_photo,_1860%29_LCCN2016823206.jpg	

FIG. 3.2	58
Dettaglio dell'incisione originale di John C. Calhoun → Schick, Nina. 2020. Deepfakes: The Coming Infocalypse. New York New York: Twelve.	
FIG. 3.3	59
Incisione manipolata da Thomas Hicks → Schick, Nina. 2020. Deepfakes: The Coming Infocalypse. New York New York: Twelve.	
FIG. 3.4	61
Fotografia originale di Joseph Stalin al Congresso del Partito Comunista, 1925 → http://www.betterphotography.in/features/shot-and-framed-framed-and-shot/42520/	
FIG. 3.5	61
Fotografia manipolata risalente al 1939 → http://www.betterphotography.in/features/shot-and-framed-framed-and-shot/42520/	
FIG. 3.6	63
User interface di Photoshop 1.0 su MacOS, 1990 → https://petapixel.com/2013/02/14/photoshop-v1-0-1-source-code-released/	
FIG. 3.7	65
Frame censurato del video diffuso su Whatsapp attribuito a Rana Ayyub → https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316	
FIG. 3.8	66-67
Rielaborazione dello spettro cheapfake-deepfake di Paris e Donovan (2019) → Rielaborazione dell'autore. Paris, Britt, e Joan Donovan. 2019. «Deepfakes and Cheapfakes».	
FIG. 3.9	72-73
Volto generati su this-person-does-not-exist.com → Generati dall'autore con this-person-does-not-exist.com	
FIG. 3.10	75
Subreddit r/deepfakes, 2017 → https://wjta.com/2018/02/15/deepfakes-a-disastrous-merger-of-ai-and-porn/	

FIG. 3.11	76-77
@deeptomcruise → https://www.theguardian.com/technology/2021/mar/05/how-started-tom-cruise-deepfake-tiktok-videos	
FIG. 3.12	78
@deeptomcruise → https://www.artribune.com/progettazione/new-media/2021/03/deep-fake-tom-cruise-christopher-ume/	
FIG. 3.13	84
Video su tiktok del podcast fra Joe Rogan e Steve Jobs, 2022 → https://vm.tiktok.com/ZMYtUJRuT/	
FIG. 3.14	88-89
Step di generazione di un'immagine su Midjourney → Generata dall'autore con midjourney.com	
FIG. 3.15, 3.16	90-91
Immagini generate con Midjourney → Generate dall'autore con midjourney.com	
FIG. 3.17	93
Sezione nella landing page di DALL-E 2 dove viene invitato l'utente a generare immagini nello stile di Matisse → https://labs.openai.com/	
FIG. 3.18-3.33	96-111
Celebrità generate con Midjourney → Generate dall'autore con midjourney.com	
FIG. 4.1	117
Prima pagina del The Patriot di Nuova Delhi, 17 luglio 1983 → https://democratic-integrity.eu/anton-shekhovtsov-operation-covid/	
FIG. 4.2	119
"Sloppy Joe" → https://twitter.com/SilERabbit/status/1254551597465518082	

FIG. 4.3, 4.4	124
Arresto di Donald Trump generato con Midjourney, 2023. → https://www.rainews.it/photogallery/2023/03/le-foto-fake-dellarresto-di-trump-create-dallintelligenza-artificiale-047e1f66-943f-4de4-be9d-4997d37f3392.html	
FIG. 4.5	125
Incontro fra Vladimir Putin e Xi Jinping generato con Midjourney, 2023. → https://www.vogue.it/news/article/intelligenza-artificiale-immagini	
FIG. 4.6, 4.7	130
Network esplorabile e sezione esperienziale di Off Track → Off Track: a journey inside the Bitchute Infodemic Underground. https://densitydesign.github.io/teaching-dd17/websites/g02/home.html	
FIG. 4.8	130
Agglomerato di hashtag rappresentativo della singolarità complottista → Among the Plandemic Truthers: a data-journey through the Infodemic conspiracist underground. https://drive.google.com/file/d/1jYi9v-v0I2liVnj904r5HdkPSTNi7JzPc/view	
FIG. 4.9	132-133
Thumbnail di video su Bitchute → https://www.bitchute.com/	
FIG. 4.10	135
Canali tematici del server Discord di Unstable Diffusion → https://unstable-diffusion.com/	
FIG. 4.11	135
Esempio di discussione fra utenti sulle control net → https://unstable-diffusion.com/	
FIG. 4.12	136
Immagine condivisa su un canale tematico → https://unstable-diffusion.com/	

FIG. 4.13	138
Control net condivisa da un utente su un canale di discussione prompt → https://unstable-diffusion.com/	
FIG. 4.14	139
Immagine ottenuta grazie alla control net precedente → https://unstable-diffusion.com/	
FIG. 5.1	151
Triangolo semiotico → Rielaborazione dell'autore. Zingale, Salvatore. 2020.«Appunti di Semiotica del Progetto A.A. 2020/2021». Scuola del Design, Politecnico di Milano.	
FIG. 5.2	151
Rapporto di mediazione segnica per somiglianza, nell'intersezione vi sono i connotati in comune fra la persona di riferimento e l'output sintetico → Rielaborazione dell'autore. Zingale, Salvatore. 2020.«Appunti di Semiotica del Progetto A.A. 2020/2021». Scuola del Design, Politecnico di Milano.	
FIG. 5.3, 5.4	152-153
Immagini virali di Papa Francesco generate con Midjourney, 2023 → https://www.leggo.it/tecnologia/news/papa_francesco_foto_piumino_bianco_intelligenza_artificiale_oggi_26_3_2023-7311854.html	
FIG. 5.5	157
Rappresentazione del rapporto fra superstrato e substrato. Segni e i simboli appartenenti al mondo reale vengono proiettati nel superstrato, mantenendo connotati riconoscibili ma perdendo contatto con la realtà a causa della facilità di manipolazione → Rappresentazione originale dell'autore.	
FIG. 5.6	163
Overlay presente sui video di tiktok che trattano tematiche inerenti al covid-19 → https://www.tiktok.com/	

FIG. 5.7	163
Finestra della rubrica premi pausa nella homepage di youtube → https://www.youtube.com/	
FIG. 5.8	164
Overlay presente sui post di Instagram revisionati dai fact-checkers → https://www.instagram.com/p/CqFn6SYs5sZ/	
FIG. 5.9	164
Finestra pop-up successiva al tap dell'overlay → https://www.instagram.com/p/CqFn6SYs5sZ/	
FIG. 5.10	165
Watermark presente sulle immagini generate da DALL·E 2 → https://labs.openai.com/	

X

RINGRAZIAMENTI

SUPERSTRATO
Sul rapporto fra essere umano
e intelligenze artificiali nell'era
della postverità

Maggio 2023

FORMATO
130×220 mm
192 pag.
100 g/ms

TIPOGRAFIA
ABC Asphalt
ABC Social
ABC Social Condensed

