

Dall'analogico al digitale. Su inconscio e linguaggio nell'era dei Big Data

Deborah De Rosa

La psicoanalisi non ha che un medium: la parola del paziente [...]. Ora, ogni parola chiama risposta.

Jacques Lacan (1956), p. 240

Il linguaggio è il modo più comune e affidabile per le persone di tradurre i propri pensieri ed emozioni interiori in una forma che gli altri possano capire.

Yla R. Tausczik, James W. Pennebaker
(2010), p. 25

Sulla scia di numerose conversioni al digitale, che negli ultimi anni hanno riguardato i più svariati aspetti della vita quotidiana, si è aggiunta di recente la proposta teorica di Derrick de Kerckhove volta a evidenziare la possibilità che, in questo processo, sia stato incluso anche il più importante concetto psicoanalitico: l'inconscio. Il sociologo ritiene che la variante virtuale dell'elemento cardine della disciplina fondata da Sigmund Freud sarebbe connessa, a sua volta, al «gemello digitale», una sorta di “replica” – sotto forma di dati – del soggetto in carne e ossa. Queste due definizioni spiccano in modo icastico nel quadro di un'ampia critica che de Kerckhove rivolge alla

direzione ormai consolidata secondo cui si evolve lo sviluppo tecnologico, specificamente nei suoi aspetti di pervasività capillare e nei suoi esiti di tracciamento e profilazione degli utenti.

Parallelamente alle riflessioni a carattere sociologico e al monito lanciato dallo studioso sugli effetti del temibile “inconscio digitale”, negli ultimi anni le ricerche sulla gestione automatica dei dati e sull'estrazione di conoscenza (*data mining*) stanno tentando di incrociare gli studi sulla psiche in modi inediti. Sono ormai disponibili diverse ricerche relative a tentativi di analisi della personalità, abbozzi di diagnosi di disturbi psichiatrici, persino analisi automatiche di resoconti onirici.¹ Lo strumento principale di questa sfida, e in generale di tutti gli ambiti nel campo del *Big data analytics*, è costituito dall'affinarsi delle tecniche di elaborazione automatica del linguaggio naturale. *Natural Language Processing* (NLP) è il nome del processo tramite cui si cerca di “insegnare” alle macchine a “comprendere” quanto più possibile del linguaggio naturale. A differenza del linguaggio formale e di quello numerico, su cui - prevalentemente in forma di sistema binario - si basa il funzionamento dei computer, le lingue usate per comunicare tra individui sono contraddistinte da un'irriducibile ambiguità semantica che le rende, almeno allo stato attuale, solo parzialmente gestibili da sistemi automatici.

Questo complesso quadro ci invita a porre a confronto due triadi: al nesso inscindibile tra soggetto, inconscio e linguaggio naturale,

¹ Qualche esempio: Nadeau *et al.*, 2006; Wald *et al.*, 2012; Razavi *et al.*, 2013; Youyou, Kosinski, Stillwell, 2015; Niederhoffer *et al.*, 2017; Eichstaedt *et al.*, 2018; Birnbaum *et al.*, 2020; Fogli *et al.*, 2020; Paquet *et al.*, 2020; Hänsel *et al.*, 2021.

già presente in Freud e reso esplicito da Jacques Lacan, corrisponderebbero, in ambiente virtuale, gemello e inconscio digitale supportati da NLP e basati su codici formali.

Le citazioni poste in apertura rappresentano due posizioni teoriche contrapposte in materia di linguaggio: quella psicoanalitica e quella che potremmo definire “ingegneristica” (cfr. Shannon, Weaver, 1949, p. 6), sottesa alle ricerche in campo di NLP.

Scopo di questo contributo è riflettere sulle effettive condizioni di possibilità dei due “equivalenti digitali” di soggetto e inconscio paventati da de Kerckhove e, di conseguenza, sui limiti dell’elaborazione automatica del linguaggio nello specifico ambito costituito da tutto ciò che riguarda la psiche.

1. Lake View

«Benvenuto nell’Upload!»: questa è la formula d’accoglienza riservata agli ospiti di Lake View, amena cittadina bagnata da un esteso e meraviglioso lago e arricchita da una folta vegetazione. Il clima è piacevole, l’Hotel è fornito di ogni leccornia, il personale è cortese e disponibile 24 ore su 24, l’arredamento è classico ed elegante. Un quadro decisamente invitante; se non fosse che Lake View è un’ambientazione interamente digitale, allestita dalla casa produttrice Horizen per ospitare versioni virtuali di persone in fin di vita, fin tanto che possono permettersi di pagare l’affitto alla speciale banca dati.

Upload (Daniels, 2020) è una serie-tv girata per la prima stagione nel 2019, e andata in onda, ironia della sorte, proprio nell’anno in cui la pandemia da Covid-19 ha costretto gran parte del mondo

ad accelerare la digitalizzazione per ogni aspetto della vita quotidiana. La serie, pubblicata per Prime Video da Amazon Studios, non è di particolare ricercatezza e non offre contenuti o riflessioni impegnate; tuttavia risulta interessante per i nostri fini in quanto propone, in chiave di commedia più che drammatica, un ipotetico scenario ambientato nel 2033 in cui il progresso tecnologico, incontrando la creatività e l'umano timore della morte, consentirebbe di conservare sé stessi in un aldilà digitale. Nella storia, "Upload" è il nome della tecnologia che permette di convertire una persona, che sia sul punto di morire ma non ancora deceduta, in un flusso di dati trasferibile in una delle realtà virtuali offerte da varie compagnie.

In un'apposita sala ospedaliera il corpo viene privato fisicamente della testa e ibernato, in attesa che lo sviluppo tecnologico prima o poi consenta di ripristinarne l'uso. Uno speciale scanner estrae i "dati di coscienza" che verranno integrati, in un'ambientazione virtuale, a un *avatar* che riproduce l'aspetto dell'ospite e che è in grado di interagire con oggetti e altri *avatar*, simulando perfettamente i *feedback* percettivi e sensoriali. A Lake View, tra le più costose piattaforme alternative alla morte "tradizionale", il livello di comfort e verosimiglianza è impressionante. Nathan, il protagonista, "sente" il sapore del cibo, la temperatura esterna e la brezza, la piacevole pressione dell'acqua della doccia; può abbracciare e stringere la mano, può persino simulare un rapporto sessuale con la sua fidanzata viva tramite apposite tute da indossare. Quando l'inevitabile sensazione che non si tratti di una vita reale si fa troppo forte, e Nathan rischia di impazzire o tentare il "suicidio" lanciandosi in un *torrent* di dati che lo distruggerebbe, Nora, la sua assistente (dipendente di Horizen in

carne e ossa) gli ricorda di concentrarsi su di sé: «Pensa all'io nella frase “Io penso dunque sono”».

Il complesso e fantascientifico assetto combinato tra “dati di coscienza”, simulazioni percettive e un *avatar* iperrealistico in un ambiente altrettanto verosimile costituirebbero, ipoteticamente, le condizioni per poter parlare di “gemello digitale” di un essere umano in maniera paragonabile all'originale definizione coniata per dispositivi e oggetti. Proviamo a spiegarci meglio.

La definizione di *Digital Twin* si deve a Michael Grieves (Florida Institute of Technology): lo studioso la introdusse nel 2003 nel suo corso di *Product Lifecycle Management (PLM)* presso l'Università del Michigan. Grieves intende il gemello come il «concetto di un equivalente virtuale, digitale, di un prodotto fisico» (Grieves, 2014, p. 1): questi ultimi due termini non si riferiscono, dunque, al corpo di un individuo, ma a materia inanimata. *Digital Twin* è il nome di un metodo, utilizzato da alcune tra le aziende maggiormente all'avanguardia, che consente di portare lo studio e la progettazione di dispositivi dal livello della modellistica 3D a quello di vere e proprie simulazioni, per ottenere dati su reazioni e funzionamenti al variare delle condizioni ma senza dover servirsi di prototipi fisici. Tra le maggiori aziende che utilizzano il modello *Digital Twin* emerge la General Electric; sul sito web della multinazionale statunitense questa procedura è definita come «una rappresentazione software di una risorsa materiale, un sistema o un processo, progettata per rilevare, prevenire, prevedere e ottimizzare, attraverso l'analisi in tempo reale».² Un'altra importante azienda ad avvalersi di gemelli digitali è Siemens, che li definisce come «una rappresentazione virtuale di un prodotto o processo,

² <https://www.ge.com/digital/blog/what-digital-twin>.

utilizzata per comprendere e prevedere le caratteristiche prestazionali della relativa controparte fisica». ³ Così, per esempio, sarebbe possibile conoscere in anteprima la precisa reazione in termini di usura o funzionalità di un motore o di un elettrodomestico a determinate condizioni, abbattendo costi e tempi necessari a condurre test su prototipi fisici.

In alcuni recenti scritti, de Kerckhove propone un nuovo utilizzo di questo termine, trasponendolo dal suo originale contesto ingegneristico a quello della sorveglianza digitale, con particolare attenzione ai risvolti sociali ed economici del fenomeno dei Big Data, così denominati per l'immensa disponibilità in termini di volume, velocità, varietà e veridicità (cfr. Zgurovsky, Zaychenko, 2020).

Secondo il sociologo, le molteplici pratiche di tracciamento e profilazione degli utenti della rete sarebbero in procinto di convergere verso la creazione di veri e propri gemelli digitali, come altrettanti «avatar» che equivarrebbero alla «nostra vita raccontata dai dati» (Rossignaud, de Kerckhove, 2020). Pur prendendo atto che il «nostro gemello digitale personale non è ancora nato», de Kerckhove ne dà per certa l'imminente realizzazione e si spinge a descriverne le caratteristiche: si tratterebbe di «un'unica entità» (*iv*) che raccolga dati e profili di una medesima persona, provenienti dalle piattaforme e sorgenti più disparate.

Una simile replica dovrebbe assolvere alla medesima funzione del *digital twin* ingegneristico: dare modo alle aziende di disporre non più soltanto di flussi di dati da interpretare, ma di 'modelli' di persone reali da sottoporre a simulazioni per anticiparne il

³<https://www.plm.automation.siemens.com/global/it/our-story/glossary/digital-twin/24465>.

comportamento. Ciò permetterebbe di affinare le tecniche commerciali di promozione e pubblicità in maniera tale da orientare scelte e acquisti con meno dispendio e un sempre minor margine di incertezza.

Il problema che si pone circa l'effettiva imminenza di simili "gemelli" sembra, innanzitutto, legato all'unificazione di informazioni che richiederebbero. I sistemi di rilevamento dati in cui ciascuno di noi è inserito, principalmente in quanto utente della rete internet, rispondono a due macrocategorie: una definita da finalità di sicurezza nazionale e internazionale, l'altra a carattere commerciale. Siamo fonte continua di tracce digitali, ormai del tutto consapevolmente: accordiamo il consenso all'utilizzo di *cookies* e dati per utilizzare servizi come Netflix, Amazon, IBS, dispositivi di domotica e *home assistants*, *social network*, *app* di vario tipo, finanche per usare *smartphone* e *smartwatch* che integrano microfoni, dispositivi di geolocalizzazione, sensori di movimento, rilevazione del battito cardiaco e conseguente monitoraggio delle fasi di veglia e sonno. De Kerckhove lo ricorda in passi come il seguente: «con Alexa, Siri, Google Assistant, Cortana, Bigsby ecc., tutte le nostre attività (e in futuro sarà potenzialmente così anche per i nostri pensieri) analizzate dall'IA sono tracciate, catalogate per servire al profitto delle *imprese* e dei *governi*» (*ibidem*). In questa ultima distinzione, che abbiamo corsivato, risiede – almeno attualmente – una differenza fondamentale. Sul piano delle imprese si verificano raccolte dati operate da agenti differenti, ciascuno secondo i propri interessi commerciali. Per esempio, a Netflix potrà interessare se trascorriamo più tempo davanti a serie tv di genere fantascientifico o a documentari, per poterci suggerire i prodotti più adatti ai nostri gusti e mantenerci così fidelizzati

come clienti. Lo stesso dato sarà irrilevante, per esempio, per Leroy Merlin, grande azienda di vendita al dettaglio di materiali per il fai da te, che sarà invece interessata al nostro comportamento di navigazione per capire se ci preme arredare il terrazzo o se compriamo spesso vernici.

Queste categorie così differenti di informazioni, però, potrebbero essere entrambe rilevanti a fini di sicurezza nazionale. Se da un medesimo indirizzo IP si rileva una scelta frequente di documentari sull'attacco alle Torri Gemelle e acquisti di materiali che, combinati, consentirebbero di costruire ordigni, il quadro complessivo potrebbe condurre a un'allerta su quell'individuo a fini antiterroristici. Questa convergenza di dati provenienti da più sorgenti, che rende possibile un'ampia profilazione di un individuo, richiede autorizzazioni ad accessi e monitoraggi trasversali che almeno per ora – e auspicabilmente in maniera duratura – non sarebbero possibili che per finalità di sicurezza. Le rilevazioni per motivi commerciali e pubblicitari costituiscono *database* spesso anche molto estesi, ma articolati su interessi diversi, organizzati secondo *cluster* specifici e non progettati per essere integrati tra loro; ragionevolmente, anche per garantire a ciascun *competitor* i vantaggi che riesce a guadagnarsi.

Per funzionare come il *digital twin* industriale, e dunque giustificare questo nome, il gemello dovrebbe essere *uno* per ciascuno di noi, tale da consentire a chi lo interroghi di simulare i precisi comportamenti e le esatte reazioni che avremmo se collocati in precise situazioni. Dovremmo trasformarci in abitanti di Lake View e “caricarci” in *avatar* perfettamente verosimili. È mai possibile, per una struttura complessa come una persona?

2. Per un inconscio analogico

Secondo de Kerckhove, l'era dei Big Data sarebbe interpretabile come un cammino verso una sorta di trasposizione digitale integrale, una situazione in cui «tutto ciò che costituisce il nostro “io”, ovvero memoria, intelligenza e giudizio, in altre parole, il nostro “contenuto”, sta emigrando in un database» (*ibidem*). A questa riflessione si lega l'altro provocatorio concetto proposto dal sociologo: l'«inconscio digitale», che si comporrebbe di «tutto ciò che si sa su di te e che tu non sai» (de Kerckhove, 2016). Paragonandolo all'inconscio freudiano, e specificando che si tratta di qualcosa di «forse ancora più potente» (Rossignaud, de Kerckhove, 2020), egli lo definisce come una «presenza di dati potenzialmente estraibili su ciascuno di noi»:

con Freud, la psicoanalisi fonda la teoria della mente a partire dal concetto di inconscio. La dimensione inconscia è caratterizzata da istinti e desideri che non si manifestano a livello razionale e che, pertanto, non sono immediatamente controllabili. Tuttavia, tale dimensione guida e indirizza i comportamenti individuali. L'inconscio digitale si caratterizza per la sua portata globale, per la straordinaria velocità attraverso cui consente l'accesso alle informazioni, per la possibilità istantanea di raccogliere e far emergere a livello cosciente una considerevole collezione di dati, correlati in diverse configurazioni in tempo quasi reale (de Kerckhove, 2015).

L'inconscio digitale funzionerebbe, pertanto, sulla base di un archivio formato dall'ipotetico accorpamento delle diverse tracce che lasciamo in rete, le quali costituirebbero uno storico

utilizzabile da terzi per «determinare i nostri comportamenti» (Rossignaud, de Kerckhove, 2020) futuri.

Un primo problema in questo parallelismo è sintetizzato nel concetto di «dati potenzialmente estraibili». A differenza delle tracce che lasciamo in rete, l'inconscio freudiano non può essere ridotto a un deposito di comportamenti, depositati nella forma in cui si sono oggettivamente svolti, e interamente richiamabili da terze persone (o persino da intelligenze artificiali). Questa visione ignora una serie di questioni di capitale importanza poste da Freud e riproposte dai successivi teorici della psicoanalisi, tra cui la questione dell'intricata temporalità dell'inconscio, il meccanismo della rimozione, il materiale traumatico talvolta irrecuperabile, il delicato lavoro dell'analista che non è lontanamente paragonabile a una *query* su un *database*. La riflessione di Sarantis Thanopoulos sul tema condensa alcuni aspetti del problema in maniera molto efficace:

de Kerckhove sovrappone due cose diverse. La prima è l'inconscio come modalità di funzionamento mentale, che produce dati inaccessibili a una loro conoscenza diretta. La seconda è il processo di accumulazione di dati di cui non si è consapevoli, perché fanno parte della "spazzatura" di uno scambio informativo che consuma i suoi dati molto in fretta o perché si producono come informazioni potenziali che restano in attesa di una loro estrazione e uso. [...] sono dati [...] caso mai "preconsci": sono ai margini della nostra coscienza ma possono essere richiamati in essa e usati (Thanopoulos, 2015).

Lo psicoanalista centra un importante bersaglio rilevando che, spesso, i «teorici della tecnologia digitale» tendono a trattare approssimativamente il «fattore umano», con cui pure sono

chiamati a rapportarsi, per via di una «concezione ideologica del pensiero che lo configura come rete di connessioni neurali assimilabile al linguaggio computazionale». Questa prospettiva dissocia il «pensiero umano» dalla «corporeità/gestualità e dall'affettività» (*ivi*), compromettendo ogni possibilità di trattarlo adeguatamente.

A nostro avviso, il problema può essere colto riflettendo sul significato dell'aggettivo "digitale". L'etimo affonda le proprie radici nel latino *digitus*, "dito"; nel campo elettronico e informatico, il termine viene ricondotto all'utilizzo in lingua inglese di *digital*, a sua volta da *digit*, "cifra". La qualifica si attribuisce «ad apparecchi e dispositivi che trattano grandezze sotto forma numerica, cioè convertendo i loro valori in numeri di un conveniente sistema di numerazione (di norma quello binario, oppure sistemi derivati da questo)», ed è dunque sinonimo di «numerico».⁴

Il suo contrario è espresso dall'aggettivo "analogico", con chiaro legame al concetto di "analogia"; infatti, esso qualifica «dispositivi, apparecchi, strumenti dove si svolgono fenomeni che sono modelli fisici di altri fenomeni»; si utilizza anche in fisica per definire «due fenomeni diversi ma retti dalle stesse leggi formali e che perciò possono essere assunti l'uno come modello dell'altro».⁵

Provando a restare nel gioco della similitudine proposto da de Kerckhove, evidenziamo che il meccanismo dell'analogia detiene un ruolo di primo piano nella formulazione dell'inconscio freudiano: per esempio nella formazione dei sintomi, nel metodo delle libere associazioni, nella formazione

⁴ <https://www.treccani.it/vocabolario/digitale2/>

⁵ <https://www.treccani.it/vocabolario/analogico/>

dei ricordi di copertura, per le idee deliranti e le fobie. Tutti questi aspetti vengono ricondotti da Lacan al piano del simbolico, suggellando la definizione dell'inconscio tramite un'analogia fondamentale: esso è «strutturato come un linguaggio» (Lacan 1956, p. 262). Ma si tratta del linguaggio naturale, che consente metafore, metonimie, ellissi e altre figure retoriche, permette di mentire, non garantisce l'univocità del messaggio poiché si fonda su ambiguità e polisemia.

In questo senso figurato potremmo dire che l'inconscio della psicoanalisi è, semmai, analogico piuttosto che digitale, poiché può essere (parzialmente) avvicinato e compreso tramite analogia con le leggi del linguaggio.

Proseguendo la similitudine, l'ipotesi dell'inconscio digitale riguarderebbe una "riduzione" al sistema binario su cui si fonda il funzionamento dei calcolatori: se di linguaggio si tratta, sarebbe in questo caso piuttosto quello formale. Come in ogni conversione elettronica da analogico a digitale, anche qui avremmo un margine di perdita, che impedisce di ricostruire il quadro completo.

3. Parole piene e vuote

La questione del legame tra inconscio digitale e linguaggio è, in effetti, fondata e anche particolarmente urgente in ambito di gestione dei dati. I contenuti testuali che produciamo e immettiamo in rete in varie forme costituiscono una tra le risorse più importanti di informazioni sul nostro conto. Nasce così la specifica tecnica del *Text Mining*, che si avvale dell'Intelligenza Artificiale per elaborare i testi prodotti dagli utenti per i più

svariati scopi (come *email*, *post* sui *social network*, *pagine web*) e convertirli sotto forma di dati strutturati adatti ad analisi automatiche. La branca di ricerca fondamentale per questo scopo è denominata *Natural Language Processing* (NLP), ed è volta a raffinare progressivamente le possibilità di trattamento automatico del linguaggio naturale (cfr. Prakash *et al.*, 2011).

Vincere questa sfida equivarrebbe a ottenere un'interazione uomo-macchina completa e consentirebbe di varcare la soglia verso la realizzazione, tuttora chimerica, di un'IA forte (*AI-complete* o *AI-hard*); finora, l'unico obiettivo in merito che sia risultato effettivamente perseguibile consiste nel procedere a raffinare la tecnica, ponendosi nell'ottica di una continua approssimazione a un limite. Infatti, data la complessità semantica e, soprattutto, l'ambiguità caratteristiche del linguaggio umano, non si può non preventivare che si tratti di qualcosa che le macchine non potranno mai maneggiare allo stesso modo delle persone, a prescindere dal livello di evoluzione tecnica raggiunta. Tuttavia è necessario riconoscere che ciò non ha impedito alla ricerca sull'IA di mettere a punto dispositivi e funzionalità straordinari, capaci di migliorare la qualità del lavoro e della vita in numerosissimi ambiti, dalle "semplici" funzioni di supporto vocale degli *smartphone* fino al campo ingegneristico e medico.

La sfida sul linguaggio ha compiuto passi da gigante dalla fine degli anni Sessanta e i primi promettenti programmi, come SHRDLU (cfr. Winograd, 1970). Negli ultimi tempi, nella fitta ramificazione di questo vasto campo di studi, si assiste alla proliferazione di ricerche in cui si tenta di utilizzare la NLP persino in ambito psichico. Utilizzando tecniche di *text mining* sono stati condotti studi su analisi della personalità, tentativi di

diagnosi di disturbi psichiatrici, rilevamento di segnali di volontà di suicidio, finanche analisi automatiche dei sogni. In quest'ultimo caso, per cui non si teme di esagerare considerandolo una tra le più eclatanti provocazioni per la psicoanalisi, sono state utilizzate delle banche dati contenenti resoconti testuali di sogni; tra le più cospicue si annoverano il *database* dreambank.net e il social DreamsCloud.

Uno dei *software* maggiormente utilizzati e diffusi per l'elaborazione del linguaggio naturale con finalità di questo tipo è LIWC (da pronunciarsi "Luke"), acronimo per Linguistic Inquiry and Word Count. Il *software* è stato ideato dallo psicologo James W. Pennebaker con il supporto della programmatrice Martha E. Francis; la definizione dell'operato di LIWC è riportata nel loro sito web: esso

legge un dato testo e conta la percentuale di parole che riflettono diverse emozioni, stili di pensiero, preoccupazioni sociali e persino parti del discorso. Poiché LIWC è stato sviluppato da ricercatori con interessi in psicologia sociale, clinica, sanitaria e cognitiva, le categorie linguistiche sono state create per catturare gli stati sociali e psicologici delle persone.⁶

L'idea di base su cui questo approccio NLP si fonda risiede nella distinzione, tra le parole che compongono i testi, in vocaboli «di contenuto» (nomi, verbi, aggettivi, avverbi) e vocaboli «di stile» (pronomi, preposizioni, articoli, congiunzioni, verbi ausiliari; cfr. Tausczik, Pennebaker, 2010, p. 29). Pennebaker ritiene che siano proprio le particelle apparentemente meno decisive sul piano del senso a veicolare maggiori informazioni su chi scrive.

⁶ <https://liwc.wpengine.com/how-it-works/>

Pertanto i vocaboli di contenuto, che consentirebbero di cogliere dati su *cosa* viene detto, non sarebbero tanto importanti quanto quelli di stile, che ne chiarirebbero invece il *come*. Il funzionamento del *software* è complesso, ma è possibile farsene un'idea parziale combinando questo assunto teorico con un metodo quantitativo di conteggio delle occorrenze delle parole (cfr. Pennebaker 2013).

Questi studi, in varia misura, riescono a mostrare anche qualche risultato e a suffragare in qualche modo il proprio operato, collocandosi in continuità con i metodi psicologici di tipo quantitativo e le impostazioni di tipo psicometrico. Tuttavia, sia da una prospettiva psicoanalitica che da un punto di vista filosofico, gli assunti sul linguaggio e, a monte, sul soggetto su cui questi studi si basano, non possono consentire in alcun modo di avvicinarsi in maniera appropriata a questioni riguardanti il campo psichico.

Per poter essere elaborato automaticamente, il linguaggio naturale deve essere trasformato in una fonte di dati strutturati e quantificabili, passando per un impoverimento inevitabile di senso e perdendo ogni sfumatura e valore soggettivo. Se il processo può garantire un buon margine di efficacia per alcuni scopi, come la sicurezza o la pubblicità, esso è incompatibile invece con gli assunti di base della psicoanalisi, la quale, come recita la citazione lacaniana in apertura, «non ha che un medium: la parola del paziente».

Anche in questo caso, per mantenere la similitudine proposta nel paragrafo precedente, la “conversione” da analogico a digitale comporterebbe la perdita di aspetti del registro del simbolico indispensabili alla diagnosi prima ancora che alla cura. L'assunto lacaniano per cui il funzionamento dell'inconscio è indagabile -

anche se sempre e solo parzialmente - per la via privilegiata dell'analogia con la struttura d'uso del linguaggio naturale, azzerando qualsiasi possibilità di intersezione tra la NLP e il mondo psichico. Oltre alle articolate dinamiche cui l'autonomia del significante dà vita, come i lapsus o le figure retoriche, è possibile individuare almeno altri due fattori di cui il discorso verrebbe privato sottoponendolo a elaborazione automatica. Uno di questi è costituito dal valore dell'assenza di parola, che può manifestarsi nel silenzio del paziente come misura, consapevole o meno, di resistenza, o come limite posto dallo psicoanalista con la sospensione dell'incontro (consentita dalla durata variabile della seduta prevista nella pratica lacaniana). Un altro aspetto che non potrebbe passare al vaglio automatico è costituito dalla distinzione - capitale almeno per una fase del pensiero lacaniano - tra una «parola piena», compiuta espressione simbolica tramite cui il soggetto riesce ad assumere il proprio desiderio, e una «parola vuota», legata all'immaginario e che perde la funzione di domanda di riconoscimento. Quest'ultima rappresenta un elemento delicato e importante da riconoscere; si tratta per Lacan della «via più ingrata», che impone di capire quando «il soggetto sembra parlare invano di qualcuno che, gli somigliasse anche fino a trarre in inganno, mai si unirà all'assunzione del suo desiderio» (Lacan 1956, p. 247). Simili aspetti non solo risultano fuori dalla portata di un'elaborazione automatica, ma sono difficilmente accessibili anche per intelligenze umane che non siano state adeguatamente formate alla professione dell'analisi. L'ipotesi di condurre delle diagnosi in ambito psichico o di analizzare resoconti onirici sottoponendo il linguaggio a tecniche automatiche di elaborazione richiederebbe che le parole, i costrutti sintattici, i giochi linguistici (qualora si trovi il modo di

riconoscerli efficacemente tramite NLP) abbiano il medesimo valore per ogni persona, e che non siano condizionati dal destinatario e dalla possibilità o meno di ricevere una risposta. Per esempio, un legame così determinante nella veglia e nel sonno, come quello che, nel caso clinico dell'uomo dei lupi, coinvolge la lettera V, il numero romano 5 con la medesima grafia e il sogno dello strappare le ali a una vespa, costituisce materia irriducibile a rilevamenti quantitativi o comuni ad altri individui.

Un ipotetico inconscio digitale non avrebbe modo di essere strutturato come un linguaggio naturale, e non potrebbe rimuovere (e tantomeno forcludere) alcunché, poiché nel digitale ogni traccia è sempre reperibile, a condizione di interrogare adeguatamente il *database*. Ma soprattutto un'eventuale istanza simile e, a monte, un ambiente di dati informatici, non possono consentire diagnosi psichiche o analisi di contenuti onirici perché in essi non c'è spazio per la domanda. «Ogni parola chiama risposta», come si legge nell'esergo; ma l'interrogazione di un *database* testuale non è un'operazione paragonabile a una domanda e una risposta tra soggetti. Infatti, quando affidiamo del testo alla rete, possiamo avere molti destinatari diversi o anche nessuno in particolare, ma non ci rivolgiamo né al *software* che tratterà quei dati né, ovviamente, alla persona del programmatore che ha configurato il sistema. I dati non consentono interazione diretta, ma solo la possibilità di un'interpretazione a posteriori e *in absentia*; per di più, dovendo funzionare secondo la maggior estensione possibile, essi costringono a cercare parametri oggettivamente validi, che si trovano in posizione ossimorica rispetto ai fondamenti della psicoanalisi.

Rispetto alla complessità, ricchezza e variabilità di cui ogni soggetto è costituito, in effetti anche i Big Data, a dispetto del nome, costituiscono comunque un campione relativamente ristretto per dare conto del comportamento, delle scelte, dei problemi e dei desideri in maniera piena.

Se consideriamo un soggetto come un intero gestaltico,⁷ pur disponendo di enormi quantità di tracce di ciò che egli fa, dove va e ciò che dice, non avremo altro che una quantità di *parti* insufficiente a ricreare la figura. La soglia mereologica dell'intero non è accessibile che ad altri esseri umani: le macchine possono essere programmate ad abbinare la definizione di "figura" a una gamma di *input*, ma - almeno per ora - il momento dell'*insight* (cfr. Köhler, 1917) è per loro inaccessibile.

Per questi motivi, possiamo rilevare che la seconda citazione d'apertura sottende una concezione del linguaggio, e dunque del soggetto, che si pone a una distanza irriducibile rispetto a quella psicoanalitica. La psicoanalisi insegna che, in particolar modo quando si tratta di «emozioni», desiderio e rapporto con gli altri, il nostro utilizzo della parola è spesso tutt'altro che lineare ed efficace allo scopo, anzi risulta spesso largamente fuorviante rispetto a ciò «che gli altri possano capire». Se da un lato la complessità del soggetto e dell'inconscio ci rassicura rispetto alla possibilità di una loro temibile edizione "digitale", d'altra parte i recenti tentativi di automatizzare parte del lavoro in campo psichico devono sollecitare all'impegno, filosofico e clinico, di ribadire l'importanza di un'appropriata, e non riduzionista, concezione del soggetto.

⁷ Nel nostro De Rosa (2020) argomentiamo l'opportunità di utilizzare questo paragone in merito al soggetto lacanian: cfr. *ivi*, pp. 21-46.

Bibliografia

- Birnbaum, M.L. *et al.* (2020), *Identifying emerging mental illness utilizing search engine activity: A feasibility study*, in *Public Library of Science One*, 16 October.
- de Kerckhove, D. (2015), articolo in *La Repubblica*, 28 giugno, <https://ricerca.repubblica.it/repubblica/archivio/repubblica/2015/06/28/inconscioidigitale42.html>
- Id. (2016), *La rete ci renderà stupidi?*, Castelveccchi, Roma, ebook.
- De Rosa, D. (2020), *Intrecci. Fenomenologia e psicoanalisi tra Lacan e Merleau-Ponty*, Mimesis, Milano.
- Eichstaedt, J.C. *et al.* (2018), *Facebook language predicts depression in medical records*, in *Proceedings of the National Academy of Sciences of the United States of America*, vol. 115, n. 44.
- Fogli, A. *et al.* (2020), *Our dreams, our selves: automatic analysis of dream reports*, in *Royal Society Open Science*, vol. 7, n. 8.
- Grieves, M. (2014), *Digital Twin: Manufacturing Excellence through Virtual Factory Replication*, whitepaper, https://www.researchgate.net/publication/275211047_Digital_Twin_Manufacturing_Excellence_through_Virtual_Factory_Replication.
- Hänsel, K. *et al.* (2021), *Utilizing Instagram Data to Identify Usage Patterns Associated With Schizophrenia Spectrum Disorders*, in *Frontiers in Psychiatry*, vol. 12.
- Köhler, W. (1917), *L'intelligenza nelle scimmie antropoidi*, tr. it., Giunti-Barbera, Firenze, 1968.
- Lacan, J. (1956), *Funzione e campo della parola e del linguaggio in psicoanalisi*, tr. it., in Id. (1966), pp. 230-316.

- Id. (1966), *Scritti*, 2 voll., tr. it., Einaudi, Torino 1974.
- Nadeau, D. *et al.* (2006), *Automatic Dream Sentiment Analysis*, in *Association for the Advancement of Artificial Intelligence Workshop - Technical Report*.
- Niederhoffer, K. G. *et al.* (2017), *In your wildest dreams: the language and psychological features of dreams*, in *Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology*, Association for Computational Linguistics, Vancouver, Canada, pp. 13-25.
- Paquet, C. *et al.* (2020), *A Quantitative Text Analysis Approach to Describing Posttrauma Nightmares in a Treatment-Seeking Population*, in *Dreaming*, vol. 30, n. 1, pp. 54-67.
- Pennebaker, J.W. (2013), *The Secret Life of Pronouns. What Our Words Say About Us*, Bloomsbury Press, New York.
- Prakash, M.N. *et al.* (2011), *Natural language processing: an introduction*, in *Journal of the American Medical Informatics Association*, n. 18, pp. 544-551.
- Razavi, A.H. *et al.* (2013), *Dream sentiment analysis using second order soft co-occurrences (SOSCO) and time course representations*, in *Journal of Intelligent Information Systems*, vol. 42 n. 3, pp. 393-413.
- Rosignaud, M.P.; de Kerckhove, D. (2020), *Oltre Orwell. Il gemello digitale*, Castelvecchi, Roma, ebook.
- Shannon, C.E.; Weaver, W. (1949), *La teoria matematica delle comunicazioni*, tr. it., Etas Kompass, Milano 1971.
- Tausczik, Y.R.; Pennebaker, J.W. (2010), *The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods*, in *Journal of Language and Social Psychology*, n. 29, pp. 24-54, traduzione nostra nelle citazioni.

- Thanopoulos, S. (2015), *L'inconscio e la rete*, in *Psychiatry on line*, 20 ottobre, <http://www.psychiatryonline.it/node/5882>
- Wald, R. et al. (2012), *Using Twitter Content to Predict Psychopathy*, in *Proceedings of the 11th International Conference on Machine Learning and Applications*, vol. 2, pp. 394-401.
- Winograd, T. (1970), *Procedures as a representation for data in a computer program for understanding natural language*, MIT AI Technical Report 235, <https://dspace.mit.edu/handle/1721.1/7095>.
- Youyou, W., Kosinski, M., Stillwell, D. (2015), *Computer-based personality judgments are more accurate than those made by humans*, in *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, n. 4.
- Zgurovsky, M.Z.; Zaychenko, Y.P. (2020), *Big Data: Conceptual Analysis and Applications*, Cham, Springer Nature Switzerland.

Abstract

From Analog to Digital. On unconscious and language in the era of Big Data

In the wake of numerous conversions to digital technology, which in recent years have affected the most diverse aspects of daily life, the theoretical proposal of Derrick de Kerckhove has recently been added to highlight the possibility that, in this process, the most important psychoanalytic concept has also been included: the unconscious. The sociologist believes that the virtual variant of the pivotal element of the discipline founded by Sigmund Freud would be connected, in turn, to the “digital twin”, a sort of

“replica” - in the form of data - of the subject in the flesh. The purpose of this paper is to reflect on the actual conditions of possibility of the two “digital equivalents” of subject and unconscious feared by de Kerckhove and on the limits of automatic language processing (NLP) in the specific field of everything related to the psyche.

Keywords: Digital Unconscious; Digital Twins; NLP; Analog Unconscious; Big Data.