

Studiare la storia dell'AI per capirne il futuro

A cura di **Elisa Marasca**

Con gli interventi di:

Barbara Czarniawska, Professoressa Emerita dell'Università di Göteborg

Lorenzo Fattori, Ricercatore di Organizzazione aziendale all'Università della Valle d'Aosta

Luigi Maria Sicca, Professore ordinario di Organizzazione aziendale all'Università degli Studi di Napoli Federico II, Coordinatore del Dottorato di ricerca in *Law and Organizational Studies for People with Disability* della Scuola Superiore Meridionale.

Gli studi organizzativi offrono materiali importanti per approfondire i problemi dell'economia e della società contemporanea. Questa rubrica commenta i libri recenti che danno un contributo in questo senso.

*La voce artificiale.
Un'indagine media-archeologica
sul computer parlante*
di Domenico Napolitano



Ridefinire il senso della comunicazione in uno scenario socio-tecnico e postumanistico come quello in cui la voce artificiale si inserisce, e che essa stessa contribuisce a produrre. È questo l'obiettivo del libro *La voce artificiale. Un'indagine media-archeologica sul computer parlante* (Editoriale Scientifica, 2022) scritto da Domenico Napolitano, Assegnista di Ricerca in Organizzazione aziendale presso la Scuola Superiore Meridionale di Napoli nell'area *Law and Organizational Studies for People with Disability* e Visiting Researcher presso l'Humboldt Universität di Berlino. Per l'autore infatti la voce artificiale è uno dei motivi immaginari più ricorrenti di personificazione della tecnologia. Non soltanto perché la voce umana è pensata come il naturale che fa da modello per l'artificiale, ma anche perché le rappresentazioni sociali della 'macchina parlante' la ritraggono continuamente come qualcosa che può agire al posto di un'altra persona umana, piuttosto che riconoscerla per sé. Ma questo tipo di tecnologia non è solo imitazione da parte del computer, perché mette in campo una serie di saperi, pratiche, operazioni e norme che ridefiniscono il concetto stesso di voce, così come la portata di cosa si possa fare attraverso di essa. Il suo primo effetto è quello di 'detronizzare' l'umano come unico ed esclusivo essere parlante. "La tecnologia non è una semplice protesi. Media e natura si determinano a vicenda con una grande varietà di interessi socio-economici e valori culturali", scrive Napolitano. Alla voce di queste considerazioni, la voce artificiale per l'autore finisce per essere vista come l'espressione di una condizione che si verifica tra esseri umani e tecnologia, oppure una modulazione del confine mobile che li mette in relazione.



LA VOCE È SEMPRE STATA E SARÀ UN MEZZO DA USARE CON CAUTELE

DI LUIGI MARIA SICCA

La ricerca che qua si presenta si iscrive nel milieu che interessa la dialettica tra studi (e pratiche) culturali e studi organizzativi, quindi dei modi di pensare la comunicazione intersoggettiva. In questo clima, torna feconda l'assonanza con la lunga marcia, avviata agli inizi degli Anni 90, che ha intrecciato gli studi organizzativi con quelli musicali. Se è vero che questi studi appaiono spesso abbastanza chiari e rassicuranti alle tribù di riferimento, è altresì assodato che l'uno e l'altro si prestano sistematicamente a dubbi ontologici, dunque messa in discussione di cosa sia tanto l'organizzazione quanto la musica. "Musicare", per esempio, può non funzionare un granché perché significa proprio un'altra cosa, per dire del mestiere dei musicisti. E lo stesso vale anche per "coltivare" con riferimento alla parola cultura. Musica, organizzazione e coltivazioni, insomma, prese isolatamente, possono significare molto o troppo, in ultima istanza anche poco o nulla. In questa partita la voce è (sempre stata e sarà) medium. Da maneggiare con cura.

Quella marcia di avvicinamento tra suonare e organizzare, agli albori considerata come esperienza esotica o comunque eccentrica rispetto all'allora pensiero prevalente attestato sulla tradizione dell'*industrial organization* o altri costrutti di marca razionalistica, nel corso degli ultimi 30 anni si è istituzionalizzata con risultati riconoscibili nel dibattito internazionale e una generazione di studiosi solidi che dedicano tempo, risorse intellettuali e impegno alla

comprensione proprio di quell'intreccio, che è per definizione una narrazione, un plot. Domenico Napolitano è uno di loro e ci offre anche lui un intreccio, il suo: una riflessione dall'interno della letteratura sottesa alla co-costituzione e co-determinazione di saperi bilanciata con una fenomenologia, quella della voce artificiale della e nella macchina, fatta di storie, desideri, interessi, fantasie, immaginari.

UN'INDAGINE CHE PARTE DAGLI ALBORI DEL 'COMPUTER PARLANTE'

Dalle prime aziende che iniziarono a lavorare all'integrazione di algoritmi (l'autore ci ricorda della Texas Instruments, che nel 1978 commercializzò *Speak&Spell*, in Italia noto come "Grillo Parlante"), alla messa a profitto da parte delle piattaforme digitali attraverso quel capitalismo della sorveglianza caro alla Zuboff, Napolitano evidenzia i crescenti cambiamenti che il mondo dei Big data sta imprimendo attraverso strategie di produzione di valore economico e finanziario, senza trascurare il ruolo del Knowledge management nei rapporti tra ricerca pubblica universitaria e quella privata. In questo lavoro, l'autore da un lato recluta le più qualificate impostazioni di metodologia della ricerca qualitativa in seno alle scienze umane e sociali, dall'altro sollecita posizioni di frontiera nella ricerca organizzativa. Alimentando, in questo modo, la dialettica tra la profondità di un sano investimento intra-disciplinare e rigore del confronto inter-disciplinare. Il focus di questo volume (il computer parlante, declinazione informatica della voce artificiale che ha preso avvio negli Anni 60 del secolo scorso per imporsi più di recente, specie attraverso la messa in campo di tecniche di Machine learning e Intelligenza Artificiale) è opportunamente circoscritto in verticale secondo i canoni della competizione accademica internazionale.

Lo apprezzeremo sfogliandolo, lungo la dorsale che inizia con una ampia introduzione e attraversa quattro capitoli fino alle conclusioni e all'appendice con efficaci precisazioni terminologiche. Parallelamente, non sfuggirà al lettore analogo attenzione alla dimensione idiosincratica di questo (come qualunque altro, il che dà merito all'onestà intellettuale del ricercatore) specifico dominio che, per essere trattato correttamente, invoca necessariamente una articolata famiglia di concetti.

Napolitano ci propone quindi sia un oggetto sia un metodo: si muove da un tema ben definito e curvandolo con la forza della argomentazione, che esita anche nella capacità di dimostrazione, intercetta ideali regioni di confine, dove si incontrano e si intersecano saperi organizzativi, musicali, filosofici e sociologici. Un'indagine

media-archeologica sul computer parlante, insomma, che getta le premesse per un'analisi critica del paradigma abilista che ha orientato il dogmatismo della tradizione economica neoclassica prima, e di quella aziendale poi. Uno sguardo sul mondo che non può non fare i conti con le evidenze socioeconomiche di questo complicato passaggio di millennio e delle conseguenze associabili ai facili ottimismo che hanno caratterizzato la seconda metà del Novecento. Parola di Alexa. Chiedetelo a Siri.

ANALIZZARE LA VOCE ARTIFICIALE CON UN APPROCCIO SOCIALE

DI LORENZO FATTORI

La trasformazione digitale dei processi produttivi e delle pratiche socio-culturali è uno dei principali poli attorno ai quali, negli studi di Organizzazione aziendale, si sono coagulate la riflessione e il dibattito, già prima dell'insorgere della pandemia da Covid-19. Questa poi, a propria volta, ha ribadito la centralità della riflessione sui processi mediali e sullo sviluppo della comunicazione digitalizzata in un'ottica interdisciplinare, che sappia quindi tener conto degli aspetti epistemologici, produttivi e delle pratiche che sono implicati in queste trasformazioni.

Come sottolinea Domenico Napolitano nel suo *La voce artificiale*, la ricerca sullo sviluppo delle tecnologie digitali e dei media chiama dunque in causa molteplici discipline, arrivando a lambire la riflessione sulla nuova condizione antropologica della contemporaneità.

La pubblicazione di questo testo è caduta opportunamente proprio mentre la sintesi vocale, veicolata dall'accelerazione delle ricerche nell'Intelligenza Artificiale (AI), sta conoscendo un sempre più rapido sviluppo, con nuove tecnologie che consentono di generare voci più realistiche e naturali. La proposta di Napolitano è di studiare la voce artificiale integrando l'archeologia dei media con un approccio costruttivista sociale: è con questo sguardo, approfondendo la storia della sintesi vocale e le successive trasformazioni e ibridazioni a essa riconducibili, che l'autore fa esplodere numerose questioni sottese al complesso rapporto tra l'accelerazione tecnologica e i processi di riconfigurazione organizzativa.

I VARI UTILIZZI DEI MODELLI DI GENERAZIONE DI LINGUAGGIO

I pilastri su cui si appoggia la riflessione dell'autore sono, in estrema sintesi, due. Innanzitutto, l'approccio media-archeologico, che non dirige lo sguardo verso una mera descrizione delle tecnologie, bensì orienta la riflessione in

direzione dell'insieme di saperi, valori e pratiche culturali che si coagulano attorno alle trasformazioni tecnologiche. Ciò vuol dire che, lontano da un'impostazione deterministica, il lavoro di Napolitano cerca di tenere in considerazione le complesse reti di trasformazione e di significazione che vengono in essere tra attori umani e non umani, processi in cui i mutamenti sociali aprono spazi a sviluppi tecnologici che, a loro volta, veicolano nuovi mutamenti sociali. E dunque, e veniamo al secondo punto di appoggio, cioè che la voce artificiale viene considerata un fenomeno socio-tecnico: qualcosa che, come sottolinea l'autore, attraversa le dimensioni della tecnologia, del sapere, dell'immaginario, delle relazioni sociali e delle strutture antropologiche.

Napolitano ha cura di precisare che, nel riferirsi alla voce artificiale, la sua attenzione è rivolta innanzitutto alla "capacità del computer di generare parlato in forma sonora attraverso processi computazionali", il che da un lato circoscrive, e dall'altro espande, il campo di interesse rispetto ai processi di riproduzione vocale. Dopo una storia della sintesi vocale, l'autore fa notare che il vero salto di qualità c'è stato negli ultimi anni: questi sistemi utilizzano algoritmi di apprendimento automatico per generare la voce sintetica. I modelli di generazione di linguaggio implementati grazie a questi sistemi producono una varietà significativamente più ampia di voci e toni rispetto a tutti i sistemi precedenti e, soprattutto, hanno la possibilità di generare parole, frasi o discorsi completamente nuovi rispetto perfino ai dati utilizzati per il loro addestramento.

TANTE POSSIBILITÀ QUANTE DOMANDE ETICHE

Possibilità e rischi, nell'utilizzo di tali tecnologie, si intrecciano. Da un lato, ad esempio, una rilevante potenzialità di assistenza alle persone con disabilità fonco-articolatoria (si pensi alla SLA), le quali possono "recuperare" la loro voce, con tutte le caratteristiche timbriche e prosodiche che la rendono unica, attraverso dispositivi di "clonazione vocale"; oppure la possibilità di creare interfacce vocali sempre più naturali e intuitive, rendendo più facile per gli esseri umani comunicare con loro e utilizzarle in una vasta gamma di contesti. Dall'altro la contraffazione, la possibilità che le voci sintetiche vengano utilizzate per impersonare altre identità o per diffondere false informazioni, con la produzione di *deepfake*.

Sono soprattutto le possibilità offerte dagli sviluppi tecnologici più recenti a porre questioni profonde sulla natura stessa della voce artificiale: la voce, nella storia dell'umanità, aveva finora goduto dello statuto di forma espressiva più immediatamente riconducibile alla manifestazione di una natura umana. Questo legame viene, se non spezzato,

perlomeno messo in discussione dall'irrompere della possibilità che si producano eventi sonori, aventi la forma della voce umana, ma che diversamente da essa non sono legati a una coscienza e a un'intenzionalità: sottolinea infatti Napolitano che "attraverso la sintesi vocale o il riconoscimento automatico del parlato la macchina simula gli effetti dell'intelligenza senza tuttavia accedere al regno del semantico e della comprensione". La riflessione, insomma, continua anche dopo la pubblicazione di questo volume, nella convinzione che il dibattito da coltivare ancora, in seno agli studi organizzativi, ha terribilmente bisogno di ulteriori approfondimenti e sguardi.

COME DEVONO PARLARE I ROBOT?

DI BARBARA CZARNIAWSKA

In effetti, "la voce artificiale" si sente da per tutto e il libro di Napolitano non potrebbe essere più attuale. Nella storia che egli descrive, assistiamo al passaggio da macchine che in passato potevano riprodurre musica e ripetere qualche parola come dei pappagalli, a macchine che adesso possono parlare per comunicare! È questo un passo decisivo per una strada che ci porterà a Frankenstein? Nel romanzo di Kazuo Ishiguro *Klara e il Sole* (Einaudi, 2021) un robot parla senza nessun problema, mentre i problemi sono connessi alla coscienza (quella dei robot e degli umani) e alla socialità che attraversa i confini delle specie. Davvero siamo già così lontano? Tra le tantissime nuove informazioni contenute nel libro, mi sono posta anche una domanda, dettata dai miei interessi, ma riconosciuta anche dall'autore: come devono parlare i robot? Come uomini, oppure come robot: "macchinicamente", "roboticamente" ecc.?

Come nota Napolitano, le discussioni sui lati buoni e cattivi del digitale inizialmente si concentravano soprattutto sugli aspetti visuali: adesso si porge sempre più attenzione a quelli sonori, producendo molte nuove domande, tra le quali forse la più importante è, appunto, quella della sorveglianza acustica. È anche rilevante che nell'ambito sonoro si inizi a parlare di "zona per-turbante" (la cosiddetta *Uncanny valley*), concetto inventato nel 1970 da Masahiro Mori, un giapponese, il che è particolarmente interessante perché proprio i giapponesi, con la loro fede nell'animismo, non vedono niente di strano nei robot che assomigliano alle persone umane.

UNA LETTURA OBBLIGATORIA PER GLI STUDENTI DI TUTTE LE DISCIPLINE

Eppure, in molte altre culture è proprio questo a essere 'perturbante'; perciò i robot devono avere le sembianze di aspirapolveri (si veda il contrasto tra il robot buono e



intelligente, R2-D2, e quello sciocco, C-3PO nella saga *Star Wars*). Mentre Alexa e Siri sono percepite solo come delle voci, quelle voci potranno essere incarnate in robot parlanti - e un robot che parla come una persona può fare paura. Da quello che ci descrive l'autore, gli scienziati e i robotisti fanno di tutto affinché la voce dei robot suoni come quella degli umani (con l'eccezione di artisti come Paul DeMarinis): "Il risultato considerato più importante è la maggiore naturalezza che può essere ottenuta in una voce sintetica", si legge nel libro.

Se i giapponesi, in maniera non sorprendente, sono d'accordo, cosa ne pensiamo noi altri? Il fenomeno più terrificante sembra il *voice cloning*, con la macabra idea di imitare la voce dei morti. Credo, invece, che la voce del navigatore Gps ci piaccia proprio perché è artificiale: in questo modo, noi umani ci sentiamo meno turbati. Sbaglio? Sia Michel Serres sia l'autore sembrano essere d'accordo: "La voce senza imperfezioni non ottimizza l'informazione". Questo libro presenta molti modi nuovi di vedere gli assemblaggi umano-macchinici e per questo dovrebbe diventare una lettura obbligatoria per gli studenti di tutte le discipline. Così gli umanisti capiranno meglio la tecnologia dentro la quale vivono, mentre i tecnici impareranno di più degli aspetti filosofici e sociali dei propri prodotti. Questa lettura li aiuterà a capire meglio le promesse e le sfide della nostra epoca, ma anche a capire se stessi. Così potranno aggiungere anche le loro voci al dibattito generale.