

Il riconoscimento vocale arriva ovunque, e per la privacy è anche peggio di quello facciale



Quando, nel corso di una chiamata a un servizio clienti, per un acquisto, un reclamo o una prenotazione, ascoltiamo il messaggio preimpostato “Questa telefonata potrebbe essere registrata per il controllo qualità” dovremmo farci qualche domanda in più. Perché non saranno solo gli operatori a essere monitorati nel loro lavoro – il che già apre da tempo fronti piuttosto problematici – ma anche i clienti. Le attività di marketing guidato in tempo reale da profilazioni della voce di milioni di clienti e di tutte le loro sfumature sono e saranno sempre di più una realtà, spiega su [The Conversation](#) Joseph Turow, docente di sistemi dei media all’università della Pennsylvania che sul tema ha condotto un’approfondita ricerca per il suo libro in uscita battezzato *The Voice Catchers: How Marketers Listen In to Exploit Your Feelings, Your Privacy, and Your Wallet*.

In base a tutte le interviste, le analisi e le letture che ha svolto, Turow spiega che “appare chiaro come siamo ai primi passi di una rivoluzione basata sulla profilazione della voce che le società vedono come parte integrante del futuro del marketing”. Ed è piuttosto curioso che questo accada,

aggiungiamo noi, proprio in concomitanza con l'esplosione dei social network basati sulla voce, da Clubhouse alle novità di Facebook passando per Twitter Spaces. Non solo, ovviamente: "Grazie alla larga diffusione degli smart speaker, dei display intelligenti delle automobili e degli smartphone da gestire sempre con la voce, insieme alle capacità di analisi dei call center, i responsabili di questi settori si dicono in grado di poter sfruttare tecnologie di analisi vocale per ottenere dati e informazioni senza precedenti sulle identità e le inclinazioni dei clienti" spiega l'autore. Verso una pubblicità sempre più personalizzata, ben più che nel recente passato che pure su questo punto ha visto scontrarsi senza sosta regolamentazioni continentali, specie europee, e interessi di ogni genere da parte di piattaforme, colossi della tecnologia e cosiddetti "over the top".



Il risultato? Le persone potrebbero essere profilate non solo in base agli "speech pattern", cioè ai loro schemi di conversazione e dialogo, ma anche valutate e categorizzate in virtù del timbro della voce. Un elemento per molti esperti unico e in grado di svelare, insieme a una serie di altri aspetti, sensazioni, sentimenti, tratti della personalità e perfino alcune peculiarità fisiche. Secondo i responsabili del marketing intervistati, questo tassello della profilazione sarà di fatto uno standard nel giro di una decina di anni. La

ragione è paradossalmente la poca fiducia verso il sistema attuale: fra i blocchi delle inserzioni sui dispositivi e del tracciamento pubblicitario (come accade sui prodotti Apple con la soluzione App Tracking Transparency lanciata da poco tempo), la confusione nella raccolta dei dati, l'uso di più profili e utenti su diversi dispositivi e molti altri fattori, quella profilazione – contrariamente a quanto pensano gli utenti – è tutt'altro che precisa. Fare leva su un dato biometrico come la voce cambierebbe ovviamente ogni cosa.



Ovviamente Amazon e Google, i principali fornitori di smart speaker fuori dalla Cina e dei relativi sistemi di assistenza, insieme a Siri che equipaggia i dispositivi di Apple (che però sulla privacy ha una policy praticamente blindata e su cui fa leva un bel pezzo della propria credibilità), garantiscono di non utilizzare alcuna registrazione delle conversazioni e delle richieste degli utenti sui loro miliardi di device. Per ora, spiega Turow, convinto che il futuro porterà necessariamente con sé nuovi strumenti e soluzioni. Alcuni brevetti registrati negli scorsi anni svelano in effetti qualcosa delle prospettive. Un progetto di Amazon, per esempio, illustra un dispositivo con integrato l'assistente virtuale Alexa in grado di individuare le irregolarità nelle frasi di una donna in modo da valutare se abbia un raffreddore “utilizzando un'analisi del tono, del ritmo, della voce, del

tremolio e/o dell'armonia della voce, come determinato dall'elaborazione dei dati vocali". A cosa serve? Facile: Alexa potrebbe domandare all'utente se voglia ascoltare una ricetta per una tisana, una zuppa o ordinare un certo farmaco in farmacia.



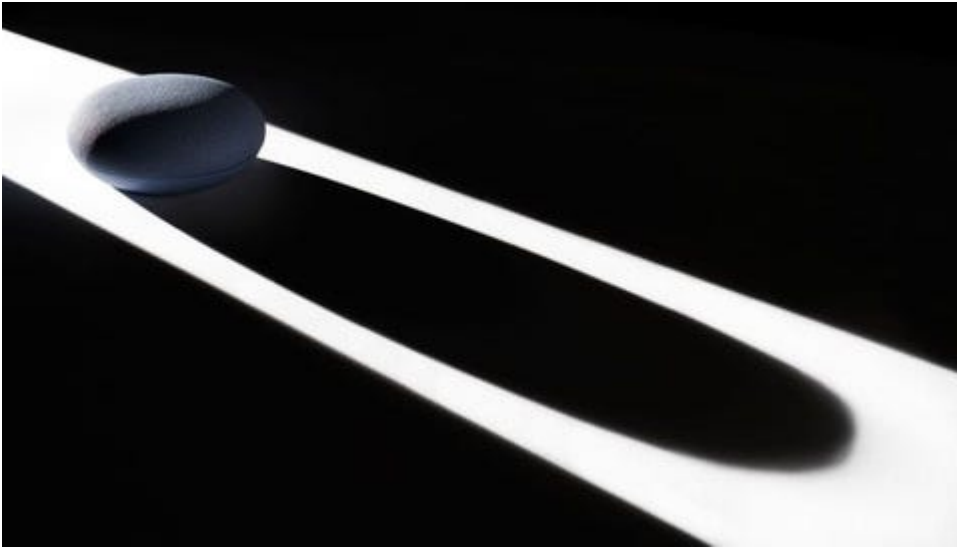
Un altro brevetto racconta invece di un'applicazione in grado di aiutare un commesso a decifrare la voce di un cliente per ricavare reazioni inconsce di quest'ultimo rispetto a un certo prodotto. E dunque variare di conseguenza la propria strategia di vendita. Uno di Google, invece, prevede di tracciare i membri di una famiglia in tempo reale sfruttando una serie di microfoni sparsi per casa: in base agli aspetti univoci delle diverse voci, Big G si dice in grado di ricavarne età e genere e categorizzarli in modo distinto così da registrare e comparare nel tempo le abitudini in casa (a che ora si mangia, quando si sta di fronte alla tv, quando si gioca ai videogame, quando si lavora e così via) e offrire agli utenti i suggerimenti più giusti per una corretta routine familiare. Inquietante, vero?

Per Turow ci si arriverà quando la frittata sarà già fatta, con buona pace delle sempre più esigenti legislazioni sulla riservatezza. Nel senso che, un po' come è accaduto con Google e Facebook e l'incredibile mole di utenti e informazioni

raccolte nel tempo, quel genere di marketing diventerà un pezzo fondamentale delle strategie di promozione e vendita proprio quando non potremo più fare a meno degli assistenti vocali. Quando, cioè, ancora più di oggi saranno dappertutto, integrati in ogni genere di dispositivo e in numeri che supereranno i miliardi di unità.



Archiviando per un momento l'aspetto più importante, quello appunto alla riservatezza, il punto più assurdo è che in realtà anche questa strada è ricca di imprecisioni. Quando si parla di emozioni e propensioni non è mica chiaro quanto la voce sia in effetti affidabile. Se è vero, secondo Rita Singh, esperta del settore della Carnegie Mellon, che l'attività dei nervi vocali è collegata allo stato emotivo, "con l'ampia disponibilità di pacchetti di apprendimento automatico personale con competenze limitate sarà tentato di eseguire analisi delle voci di livello scadente, conducendo a conclusioni dubbie tanto quanto i metodi". Senza contare una serie di altri fattori pregiudizievoli, di tipo culturale o fisiologico, che possono inficiare l'analisi.



“Sebbene alcuni di questi progressi promettano di rendere la vita più facile – conclude l’autore – non è difficile vedere come la tecnologia vocale possa essere abusata e sfruttata. E se la profilazione vocale indicasse a un potenziale datore di lavoro che sei un grave rischio per un impiego di cui hai disperatamente bisogno? E se comunicasse a una banca che sei un pessimo pagatore quando stai chiedendo un prestito? E se un ristorante decidesse che non accetterà la tua prenotazione perché sembri di bassa classe o troppo esigente?”.