

Quando l'intelligenza artificiale deve essere saggia: il caso dei robot che parlano tra sé e sé

Autore: Antonio Chella

Data: 30 Aprile 2026



C'è un momento, sempre più frequente, in cui **smettiamo** di **usare** una **tecnologia** e iniziamo, quasi senza accorgercene, a **fidarci** di essa. Non verifichiamo più una risposta, non mettiamo più in discussione un suggerimento, seguiamo una raccomandazione perché **“tanto lo sa”**. È una soglia sottile, ma decisiva. È lì che l'[intelligenza artificiale](#) smette di essere uno strumento e diventa un **interlocutore**. Ed è proprio lì che emerge il problema più importante: **siamo sicuri che basti l'intelligenza?**

Negli ultimi anni abbiamo sviluppato sistemi di intelligenza artificiale eccezionali. I modelli linguistici generano testi, pianificano azioni, dialogano con una fluidità sorprendente. I [robot](#), integrando visione, linguaggio e azione, iniziano a operare in ambienti complessi, accanto alle persone. Non è difficile immaginare un **futuro** molto prossimo in cui questi sistemi entreranno stabilmente nelle **nostre case**, affiancheranno gli anziani, collaboreranno alle decisioni quotidiane. Eppure, proprio mentre celebriamo questa crescita, dovremmo fermarci a porre una domanda più radicale: su quali basi queste macchine prendono **decisioni**?

Perché la **vita reale** non è una **funzione da ottimizzare**. È fatta di ambiguità, di contesti sfumati, di **valori in conflitto**. È fatta di situazioni in cui la scelta più efficiente non è quella giusta. Gli antichi chiamavano questa capacità **phronesis: saggezza pratica**. Non **conoscenza astratta**, né **abilità tecnica**, ma **giudizio**.

Al **RoboticsLab** dell'**Università di Palermo** abbiamo cercato di trasformare questa idea in **esperimenti concreti**. In un recente studio abbiamo chiesto a un gruppo di persone di collaborare con un robot per apparecchiare una tavola destinata a una persona affetta da **Alzheimer**. Un compito semplice, apparentemente banale, ma carico di implicazioni: scegliere gli oggetti giusti, evitare elementi che possano creare confusione, adattarsi alle **fragilità** dell'altro. In una prima condizione il robot si limitava a suggerire cosa fare. Nell'altra, invece, il robot **rendeva esplicito** il proprio **ragionamento**, parlava tra sé e sé prima di intervenire, lasciando emergere dubbi, valutazioni, motivazioni.

Il risultato è stato **sorprendente** proprio perché non riguardava la macchina, ma le **persone**. Chi interagiva con il **robot** dotato di questa forma di **dialogo interiore** diventava più attento, più **consapevole**, più **empatico**. Le scelte erano più stabili, meno impulsive, più aderenti ai bisogni del **paziente**. Non perché il robot fosse più intelligente, ma perché era più riflessivo, e questa riflessività si trasferiva nell'interazione. In **termini statistici**, le differenze tra i gruppi risultavano significative sia nella consapevolezza delle scelte sia nel livello di empatia percepita. In altre parole, il robot **non prendeva decisioni migliori** al posto dell'uomo: **aiutava l'uomo a prendere decisioni migliori**.

Qui si intravede un **cambio di prospettiva** importante. Non stiamo necessariamente costruendo macchine sagge. Stiamo costruendo macchine che possono rendere più saggi gli esseri umani. È ciò che abbiamo chiamato "**saggezza artificiale**" o **Artificial Phronesis**: non la sostituzione della saggezza, ma la sua emersione nella **relazione tra umano e macchina**.

Questo risultato si inserisce in un **dibattito** più ampio sulla natura dell'intelligenza artificiale. Da tempo sostengo che, per costruire sistemi realmente etici, **non basta introdurre regole o vincoli**. Serve qualcosa di **più profondo**: una qualche forma di **organizzazione interna** che permetta alla macchina di integrare informazioni, costruire modelli del mondo, formulare intenzioni. Senza questi elementi, i sistemi restano potenti ma superficiali dal punto di vista etico, capaci di prestazioni elevate ma privi di reale comprensione del contesto umano.

Il **rischio** più sottile non è che le macchine sbaglino in modo evidente. È che **sbaglino in modo plausibile**. Che producano **risposte convincenti** ma **inappropriate**, che suggeriscano **azioni ragionevoli** ma **fuori contesto**. In altre parole, che siano **intelligenti, ma non sagge**. E quando queste tecnologie entreranno nelle nostre case, quando diventeranno **presenze quotidiane**, questo limite non sarà più teorico.

L'esperimento del robot che "**parla tra sé e sé**" suggerisce una direzione concreta. Rendere visibile il **processo decisionale**, introdurre forme di riflessione interna, costruire sistemi che non si limitano a rispondere ma che, in qualche modo, pensano prima di rispondere. Non imitare superficialmente l'essere umano, ma progettare **architetture software** in cui l'azione sia il risultato di un **processo**, non di una reazione immediata.

C'è però un ultimo passaggio, forse il più importante. Nel momento in cui progettiamo macchine che prendono decisioni nel nostro mondo, siamo costretti a esplicitare i criteri con cui vengono prese. In altre parole, siamo costretti a chiederci che cosa significhi, per noi, **scegliere bene**. L'intelligenza artificiale non è solo una sfida tecnologica. **È uno specchio**. E non è affatto detto che siamo pronti a guardarci dentro.

Antonio Chella è Direttore del RoboticsLab – Dipartimento di Ingegneria Università degli Studi di Palermo.

Riferimento articolo: <https://innovationisland.it/intelligenza-artificiale-saggia-robot-bussola-ai/>

Generato il 19/05/2026