Linux Day Palermo 2026

I robot possono ‘capire’ cosa proviamo?

Mente, corpo e algoritmi: l’orchestrazione del dialogo nella robotica sociale

Alessandro Saracco

Dottorando in Intelligenza Artificiale
Università di Pisa · Università Cattolica di Milano
Internship all’ICAR-CNR di Palermo

Di cosa parliamo oggi?

Di cosa parliamo oggi?

01

HRI e robotica sociale

Concetti e fondamenti

02

Integrazione dei modelli linguistici

Uso e limiti degli LLM nella HRI

03

La regia del dialogo

Oltre le parole: voce, volto e fisiologia per guidare la risposta

04

Verso un utilizzo open e locale

Obiettivi e prospettive

HRI e robotica sociale

Concetti e fondamenti

Persone e robot, dal compito alla relazione

Human–Robot Interaction

Studiare, progettare e valutare l’interazione fra persone e robot

Chi fa che cosa?

Ruoli, controllo e collaborazione

Come ci si capisce?

Comandi, feedback, gestione degli errori

Funziona per le persone?

Usabilità, fiducia e sicurezza

Robotica sociale

Robot che interagiscono attraverso segnali sociali

Voce, sguardo e gesti

Mostrare attenzione e intenzioni

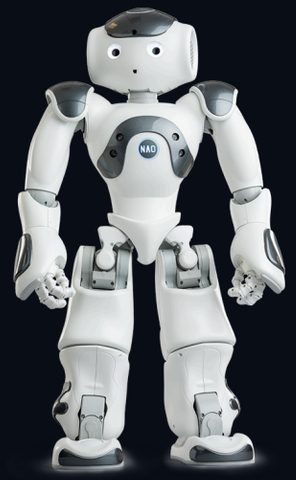
Tempi e turni

Rispondere al momento giusto

La situazione

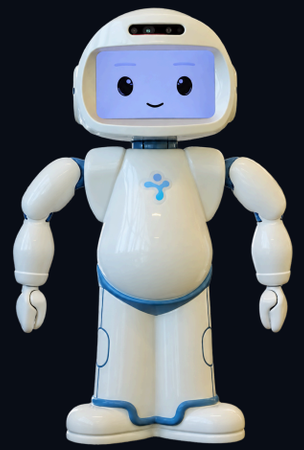
Adattarsi alla persona e al compito

Tre robot, tre compiti



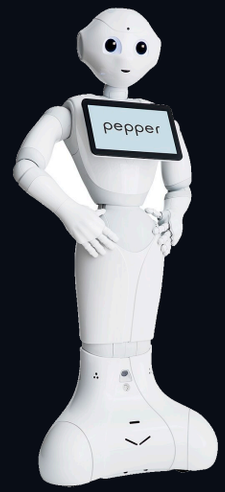
NAO

Didattica e ricerca
su gesti e dialogo



QTrobot

Esercizi guidati di
comunicazione e apprendimento

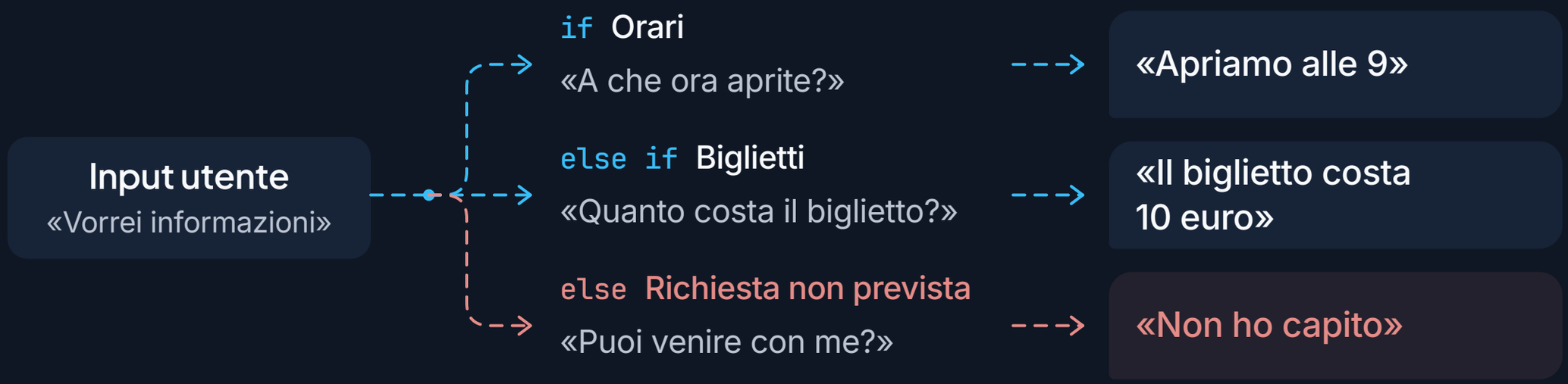


Pepper

Accoglienza, informazioni
e orientamento

Prima degli LLM: dialoghi a regole

ESEMPIO Un robot al banco informazioni di un museo



Il dialogo funziona finché la richiesta segue un percorso previsto

Integrazione dei modelli linguistici

Uso e limiti degli LLM nella HRI

Integrazione dei modelli linguistici

Più naturali, non più affidabili

Cosa migliora

- Richieste formulate liberamente
- Risposte legate al contesto
- Dialoghi su più turni

Cosa può andare storto

Allucinazioni	Inventa fatti o dettagli
Disallineamento	Ignora obiettivi o vincoli
Accondiscendenza	Dà ragione anche a chi sbaglia
Verbosità	Si dilunga quando basta poco

Integrazione dei modelli linguistici

«Chiarissimo»: conferma o ironia?

SCENARIO Il robot ha appena dato indicazioni per l'aula dell'esame

La persona dice

«Chiarissimo»

VOCE Tono ironico

VOLTO Fronte corrugata

---→

Il modello legge

`Chiarissimo`

Solo la trascrizione
Voce e volto non arrivano

---→

Il robot risponde

«Ottimo,
buon esame!»

Una risposta possibile

Servono altri indizi, o una domanda di chiarimento

La regia del dialogo

Oltre le parole: voce, volto e fisiologia per guidare la risposta

La regia del dialogo

SDialog, un toolkit Python open source

Costruire

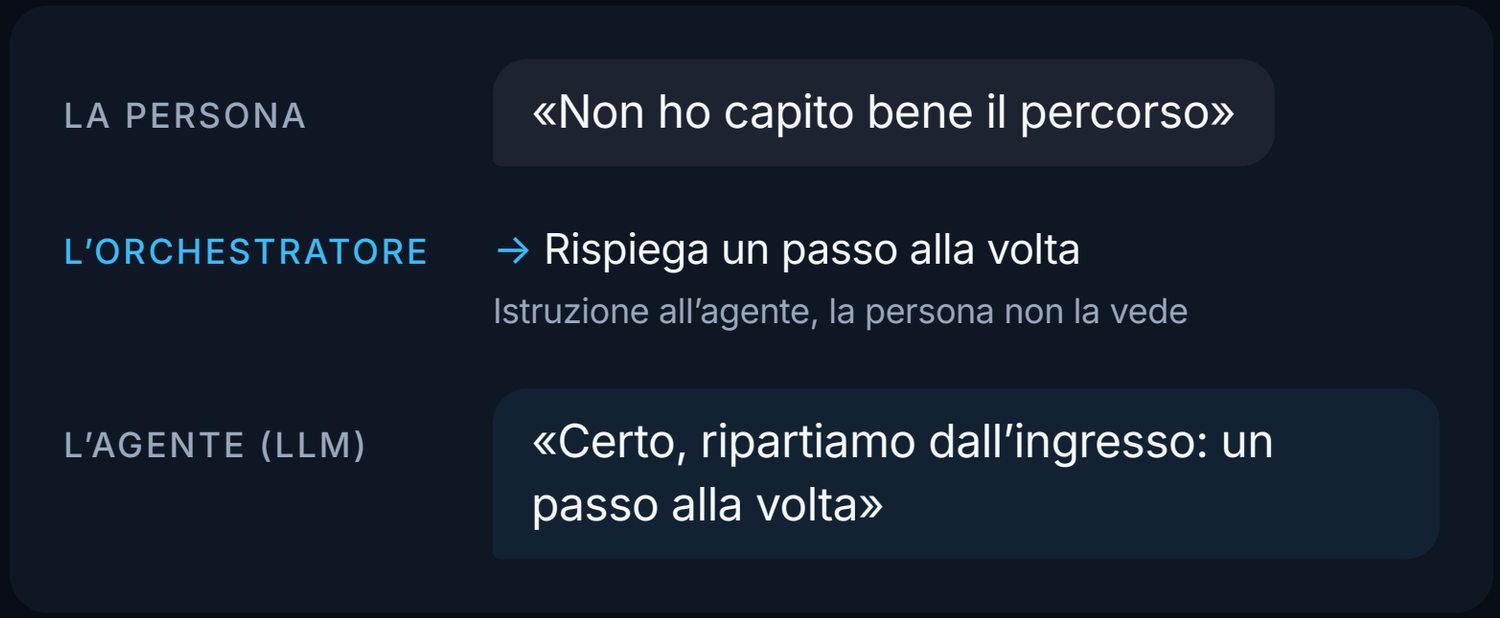
Agenti con ruolo, obiettivi e contesto

Simulare

Utenti diversi, per esempio incerti o di fretta

Orchestrare

Istruzioni all’agente al momento giusto



La regia del dialogo

Con i segnali, il robot chiede

STESSO SCENARIO Il robot ha appena dato indicazioni per l'aula dell'esame

La persona dice

«Chiarissimo»

VOCE Tono ironico

VOLTO Fronte corrugata

---→

Il motore del dialogo legge

`Chiarissimo`

`+ voce ironica`

`+ fronte corrugata`

C'è un contrasto

→ Chiedi se ripetere il percorso

---→

Il robot risponde

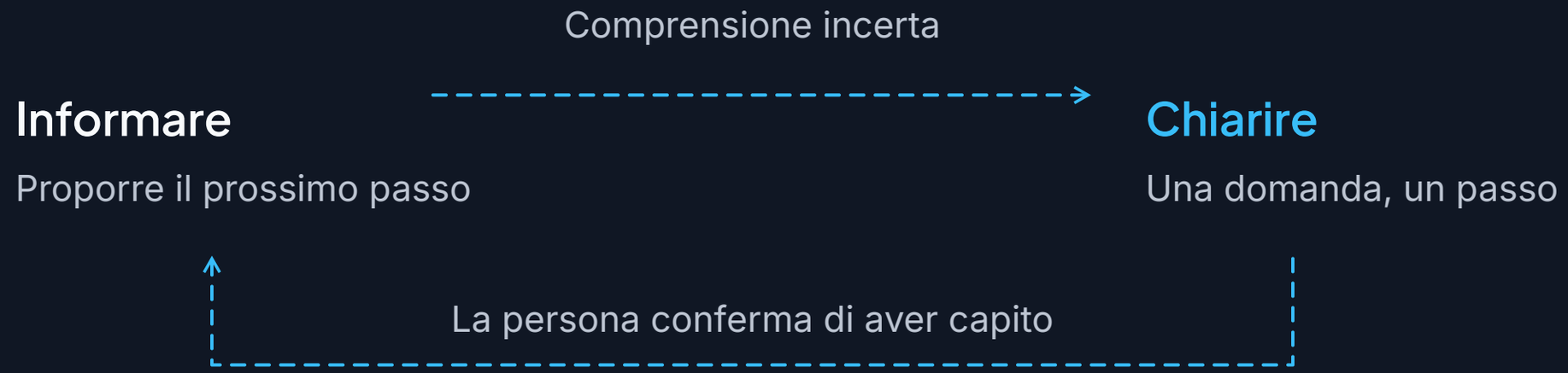
«Vuoi che ripeta il percorso?»

Una risposta possibile

Il motore prepara le istruzioni, l'LLM formula la risposta

La regia del dialogo

Ogni stato dà al modello istruzioni diverse



La regia del dialogo

Quattro ingressi, nessuno infallibile

Testo

Intento della richiesta, indizi di urgenza o frustrazione

Voce

Prosodia: tono, ritmo, intensità e pause

Volto

Movimenti, espressioni e direzione dello sguardo

Fisiologia

HR: frequenza cardiaca
EMG: elettromiografia

La regia del dialogo

Una stima della situazione

Theory of Mind: ipotesi sulla persona, da aggiornare e verificare

COMPRENSIONE

Ha seguito le istruzioni?

Che cosa sembra aver capito e dove può essersi fermata

OBIETTIVO

Che cosa vuole fare?

Il compito e le intenzioni nel contesto del dialogo

BISOGNO

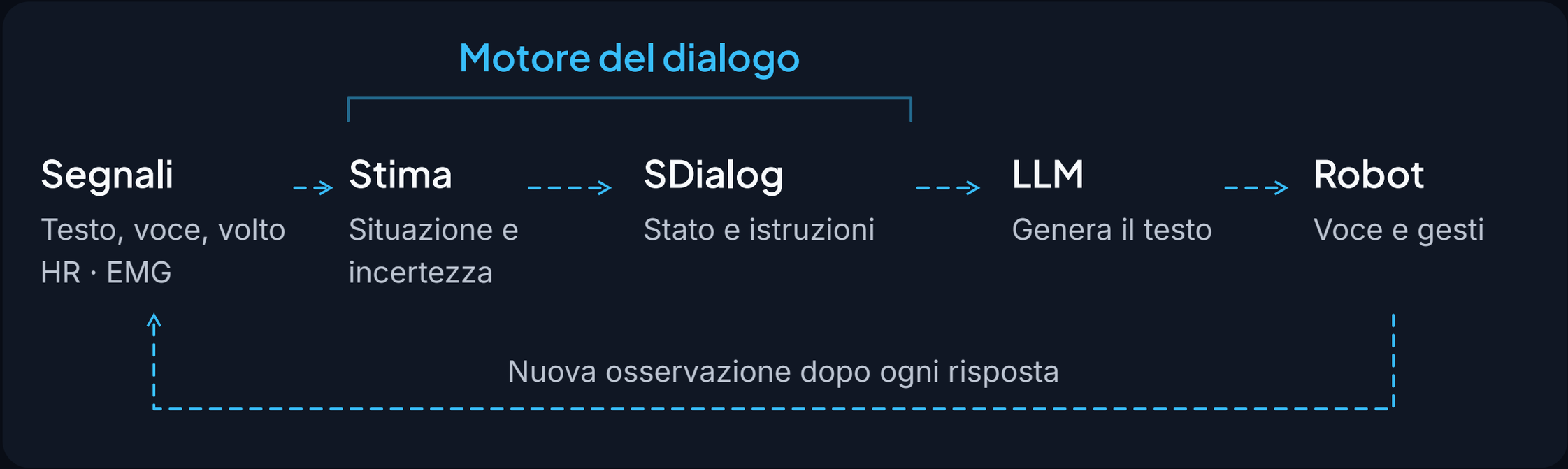
Quale aiuto può servire?

Una pausa, una verifica o un passaggio più semplice

DENTRO IL MOTORE Segnali + contesto + affidabilità → ipotesi da aggiornare

La regia del dialogo

Il ciclo continua dopo ogni risposta



Verso un utilizzo open e locale

Obiettivi e prospettive

Verso un utilizzo open e locale

Il motore prepara, il modello risponde



Ipotesi: un modello piccolo basta, se il compito è ben definito

Verso un utilizzo open e locale

Perché un modello locale, anche piccolo

Privacy

Dati in locale

Conversazioni e segnali restano sul robot

Offline

Senza API a consumo

Nessun costo per chiamata, dialogo anche senza Internet

Edge

Modelli compatti

Esecuzione direttamente sull’hardware del robot

I robot possono ‘capire’ cosa proviamo?

Non come noi.
Ma possono stimarlo, chiedere e adattare la risposta.

Grazie per l’ascolto

Domande o osservazioni?