



UNIVERSITÀ
DEGLI STUDI
DI TORINO



Loughborough
University

Centre for Research in
Communication & Culture

Macchine ingannevoli: L'Intelligenza Artificiale dopo il test di Turing

DR SIMONE NATALE

PROFESSORE ASSOCIATO, UNIVERSITÀ DEGLI STUDI DI TORINO

PRINCIPAL INVESTIGATOR, CIRCUITS OF PRACTICE PROJECT (AHRC),
LOUGHBOROUGH UNIVERSITY, UK

138° MERCOLEDÌ DI NEXA, POLITECNICO DI TORINO, 10 MARZO 2021





Michael Faraday e il dilemma delle sedute spiritiche

- 1853-54: lo scienziato britannico Michael Faraday decide di investigare il fenomeno dello spiritismo



Le conclusioni di Faraday

- I fenomeni osservati durante le sedute spiritiche sono prodotti dagli stessi partecipanti, che provocano, consciamente o inconsciamente, i movimenti del tavolo spiritico
- L'origine dello spiritismo non va ricercata tanto negli spiriti **quanto nei partecipanti alla seduta** – noi umani

Da Faraday all'Intelligenza Artificiale



DECEITFUL MEDIA



Artificial Intelligence and
Social Life after the Turing Test

SIMONE NATALE

*Deceitful Media: Artificial Intelligence
and Social Life after the Turing Test*
(Oxford University Press, 2021)

- Traiettorie storica, dal test di Turing a Siri ed Alexa
- Quando l'Intelligenza Artificiale (IA) è programmata per entrare in comunicazione con essere umani, le loro percezioni e reazioni sono importanti quanto le caratteristiche tecniche dell'IA
- Centralità della *banal deception* ("inganno banale") per comprendere la storia e il presente dell'AI

Preludio: Comunicazione e Intelligenza Artificiale

MACCHINE INGANNEVOLI



“Communicative AI”



- Programmi e interfacce di IA che entrano in **comunicazione** con gli utenti (esempi: chatbot, assistenti virtuali come Siri o Alexa, customer care, ma anche sistemi industriali che includono elementi di comunicazione)
- Rifiuto di ogni distinzione rigida tra “interazione” e “comunicazione” (Guzman, 2018; Gunkel, 2020)

Comunicazione e Intelligenza Artificiale

- Comunicazione come elemento primario delle definizioni di “Intelligenza Artificiale” fin dalle origini
- Eppure, scarsi tentativi di considerare IA a partire da un punto di vista non solo ingegneristico sulla comunicazione
- Human-Machine Communication (Guzman 2018): applicazione delle categorie e degli approcci applicati nello studio della comunicazione umana

Implicazioni per il campo dell'AI

- I meccanismi di comunicazione innescati dall'IA devono essere situati in un contesto sociale e culturale

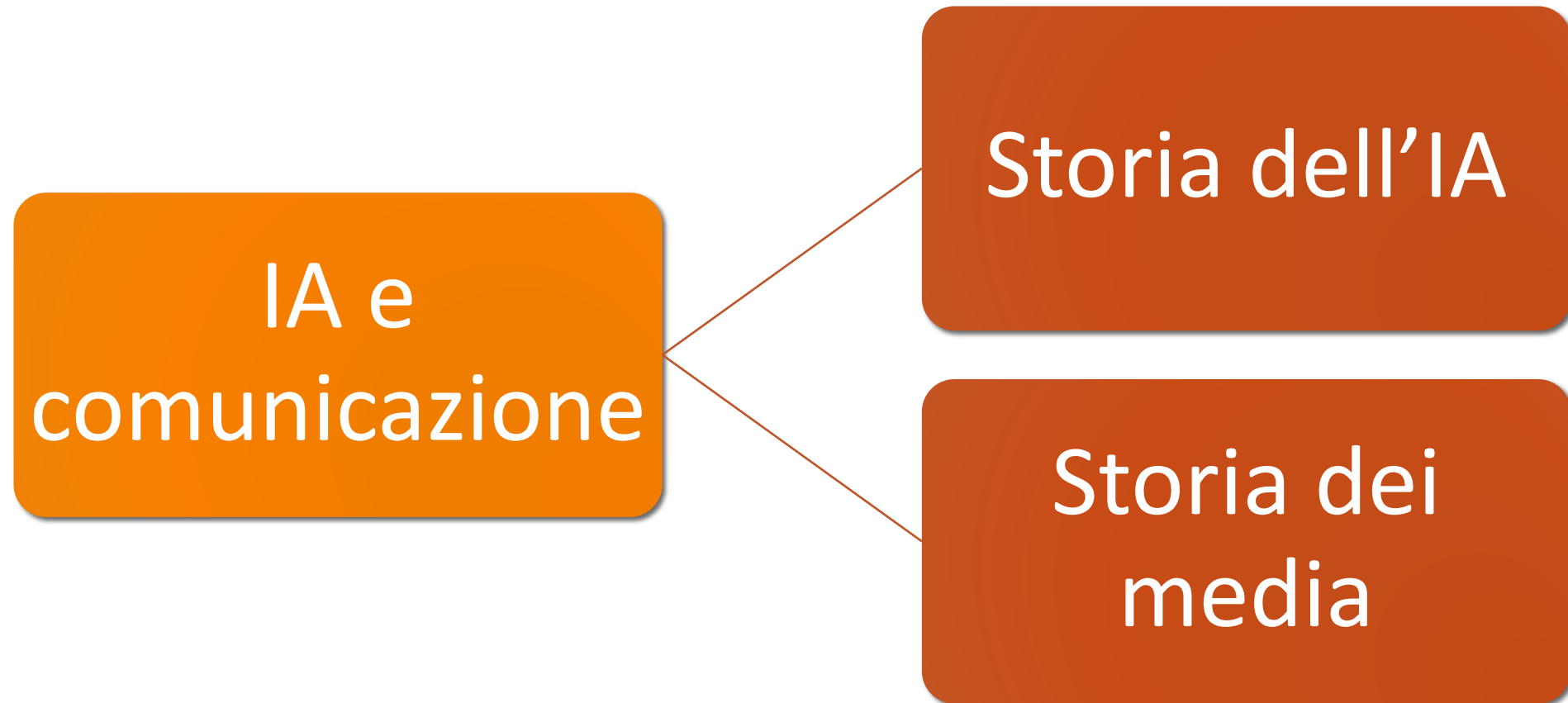




Oltre il concetto di “medium”?

- Medium come ciò che sta “nel mezzo” del processo comunicativo, ovvero il **canale**
- IA come canale ma al tempo stesso come **comunicatore**

Implicazioni per lo studio dell'Intelligenza Artificiale



1. Prima del digitale: media e inganno

MACCHINE INGANNEVOLI

Stretto rapporto tra media e *deception* (inganno)

- Inganno come centrale per il funzionamento dei media
- Tutti i media moderni, almeno a partire dal diciannovesimo secolo e dall'emergere della fisiologia, utilizzano forme 'banalizzate' di inganno (*deception*) per raggiungere determinati effetti estetici e pratici



“Banal deception” (inganno banale)

- Il concetto di inganno (deception) è usato di solito in senso negativo e per descrivere eventi eccezionali
- Nelle scienze cognitive, studiosi come Hoffman (2019) sottolineano come l'inganno rappresenti una componente funzionale e inseparabile dalla percezione
- “Banal deception” come inganno normalizzato e integrato nel funzionamento stesso delle tecnologie medial

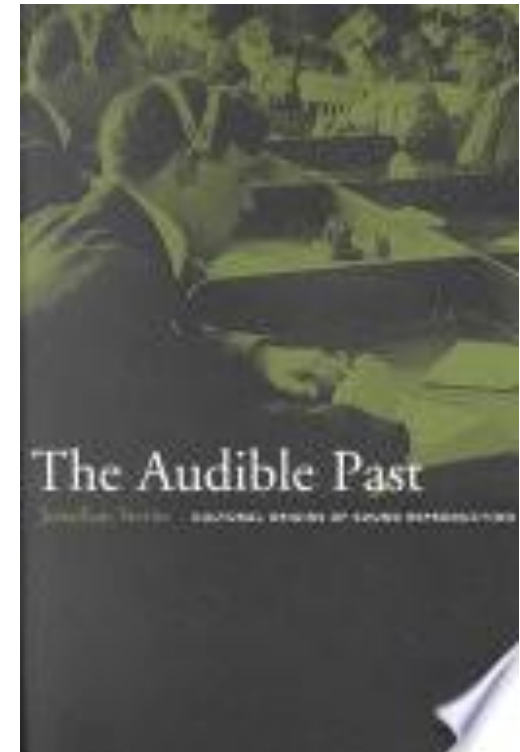
Primo esempio: il cinema

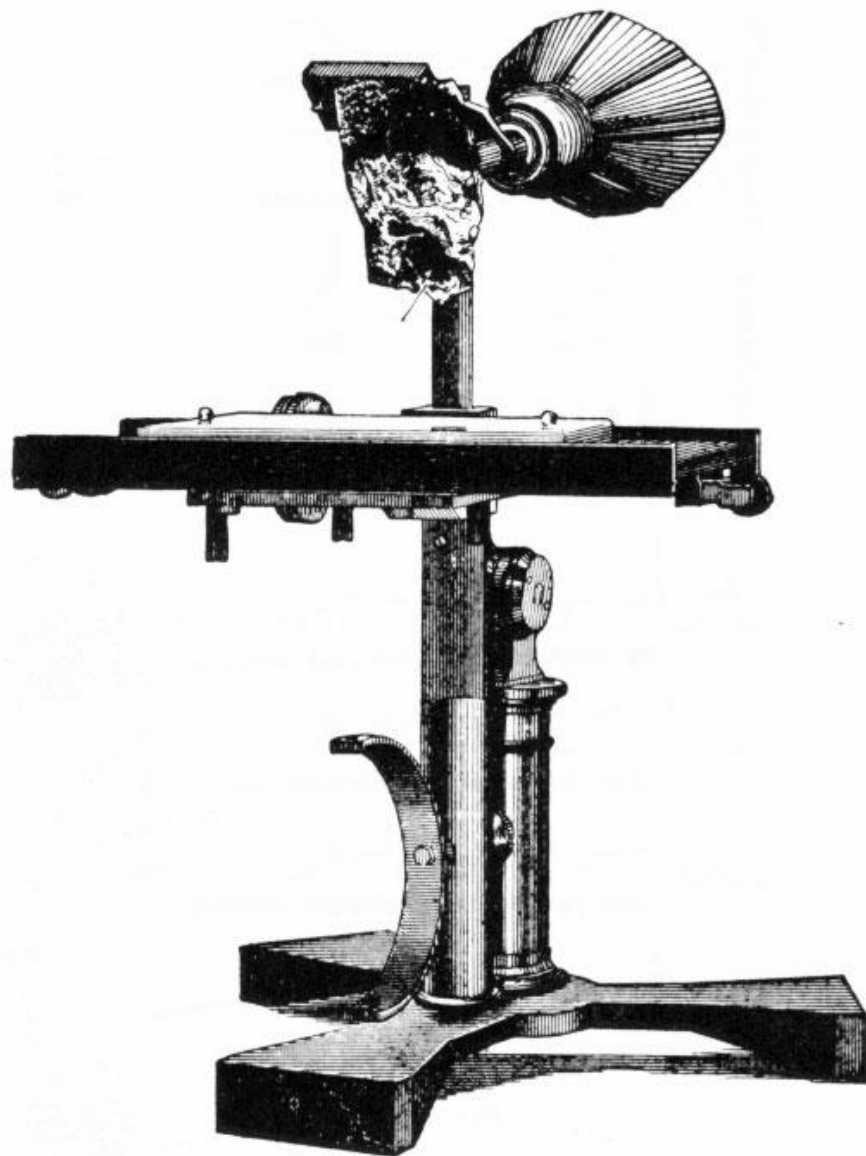
- Invenzione del cinema come risultato di studi sulla percezione umana nell'ottocento e anche in precedenza
(vedi ad esempio Doane, *The Emergence of Cinematic Time*, Harvard University Press 2002)
- Il cinema è costruito “su misura” dei sensi e dell'intelletto umani, e della nostra facilità a cadere nell'illusione



Secondo esempio: i media sonori

- I media di riproduzione e registrazione sonora emergono a partire da studi sull'udito umano (Sterne 2003, 2012)
- Il formato delle riproduzioni sonore di natura analogica, come quelle successive di natura digitale, viene calibrato sulle 'orecchie' dell'ascoltatore umano





Il fonografo di Alexander Graham Bell incorpora l'orecchio e parte del cranio di un cadavere

Marshall McLuhan
(1964): media come
“estensioni
dell’uomo”
(e della donna)

Cinque caratteristiche della “banal deception”

1. **Carattere quotidiano e ordinario**: “scompare” nel banale e nel quotidiano
2. **Funzionalità**: ha sempre un valore potenziale per l’utente
3. **Inconsapevolezza**: di solito non viene considerata come tale, ma data per scontata
4. **“Bassa definizione”**: stimola la partecipazione di utenti e spettatori (che non sono da considerare passivi o ingenui, ma profittano in maniera attiva la propria capacità di essere ingannati)
5. **Programmabilità**: può essere anticipata e ‘programmata’ (anche se mai completamente) da sviluppatori e designer

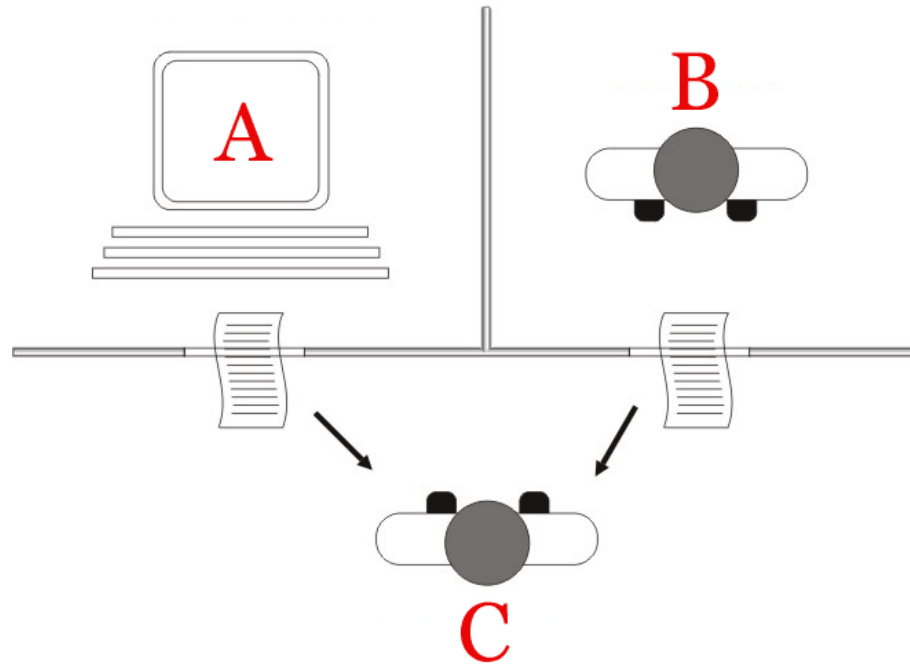
2. Intelligenza Artificiale e 'banal deception'

MACCHINE INGANNEVOLI



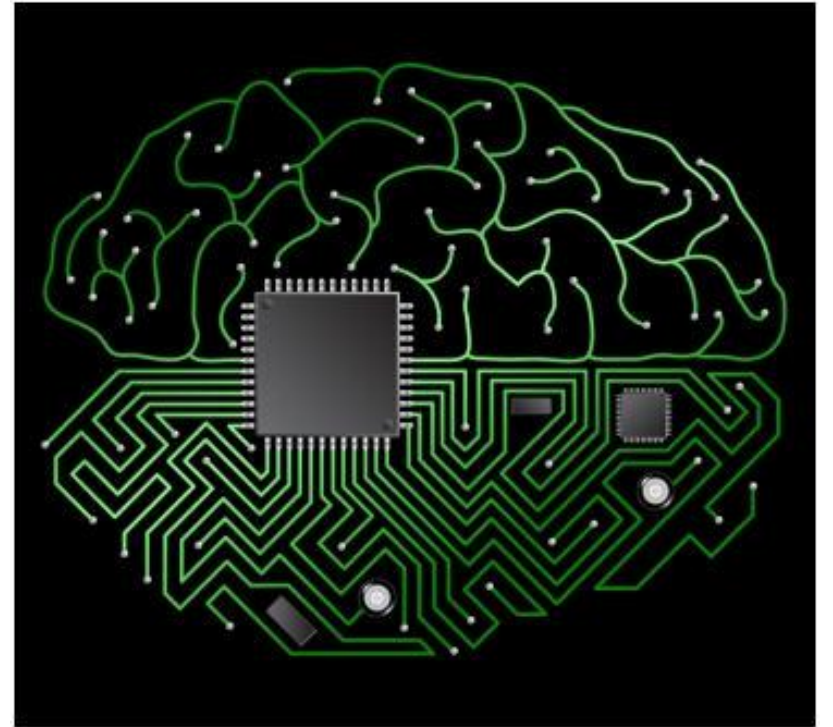
Il test di Turing

Proposto da Alan Turing nel 1950 come **alternativa** alla domanda “le macchine possono pensare?”



Implicazioni del test di Turing

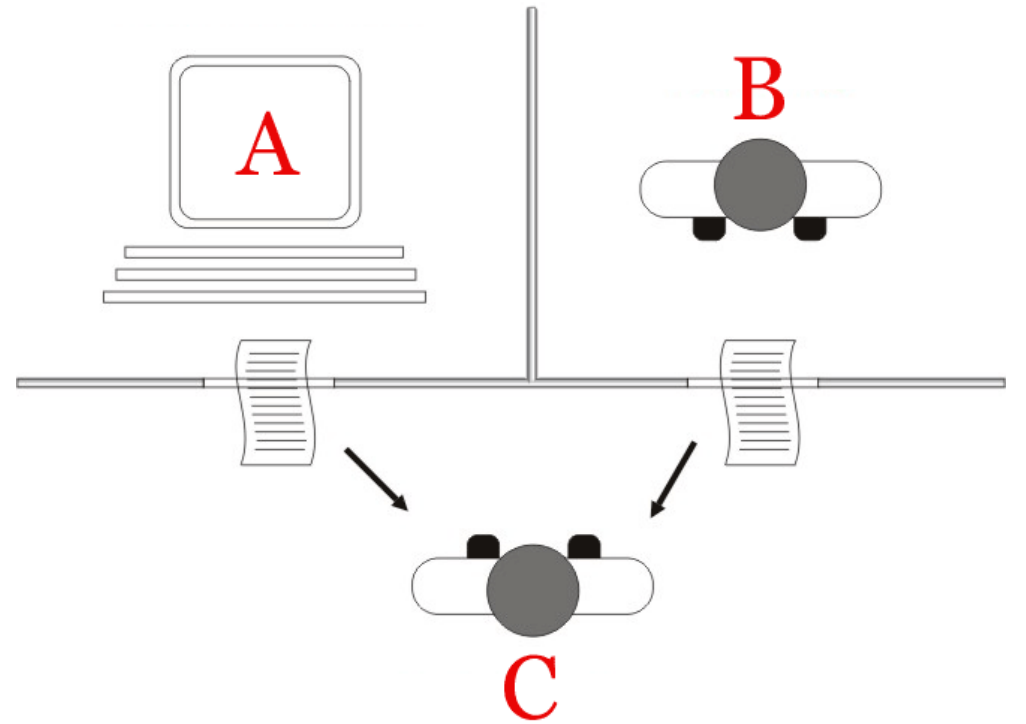
- Definizione di **intelligenza** (enfasi sul comportamento verbale)
- **Approccio operativo**: obiettivo pratico per la ricerca sull'Intelligenza Artificiale



Ruolo dell'inganno nel test di Turing

- Il computer 'vince' se riesce a ingannare l'interrogatore umano
- L'"Intelligenza Artificiale" ha a che fare con la nostra percezione; una percezione che può venire manipolata attraverso pratiche ingannevoli

Turing come Faraday?



Limiti del test di Turing

- Scelta rigida e binaria: **umano o computer?**
- Approccio che non riflette la grande maggioranza delle interazioni con l'IA

Joseph Weizenbaum ed ELIZA

1964-66:
Joseph Weizenbaum al
Massachussets Institute
of Technology crea il
primo “chatbot”, ELIZA



Welcome to

```
EEEEEE LL      IIII ZZZZZZZ  AAAAA
EE      LL      II     ZZ     AA  AA
EEEEEE LL      II     ZZZ     AAAAAA
EE      LL      II     ZZ     AA  AA
EEEEEE LLLLLL IIII ZZZZZZZ  AA  AA
```

Eliza is a mock Rogerian psychotherapist.

The original program was described by Joseph Weizenbaum in 1966.

This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?

YOU: Men are all alike.

ELIZA: What is the connection, do you suppose ?

YOU: They're always bugging us about something or other.

ELIZA: Can you think of a specific example ?

YOU: Well, my boyfriend made me come here.

ELIZA: Is it important to you that your boyfriend made you come here ?

YOU: He says I'm depressed much of the time.

ELIZA: I am sorry to hear that you are depressed.

YOU: It's true. I am unhappy.

ELIZA: Can you explain what made you unhappy ?

YOU: █

“... la magia si dissolve”

Weizenbaum predisse che ELIZA avrebbe ingannato alcuni utenti, ma solo temporaneamente:

“... una volta che un particolare programma informatico viene smascherato (“unmasked”), ovvero che il suo funzionamento è spiegato in maniera sufficientemente chiara, la sua magia si dissolve” (Weizenbaum, 1966)

Banal deception ed ELIZA

- Anecdote about Weizenbaum's secretary
- “Eliza effect” (Turkle): cerchiamo di fare in modo che ELIZA ci risponda, anche se sappiamo che è solamente un programma

Cinque caratteristiche della “banal deception”

1. **Carattere quotidiano e ordinario**: “scompare” nel banale e nel quotidiano
2. **Funzionalità**: ha sempre un valore potenziale per l’utente
3. **Inconsapevolezza**: di solito non viene considerata come tale, ma data per scontata
4. **“Bassa definizione”**: stimola la partecipazione di utenti e spettatori (che non sono da considerare passivi o ingenui, ma profittano in maniera attiva la propria capacità di essere ingannati)
5. **Programmabilità**: può essere anticipata e ‘programmata’ (anche se mai completamente) da sviluppatori e designer

3. Ecce Alexa (e altri assistenti vocali)

MACCHINE INGANNEVOLI



Three WiFi Call

13:56



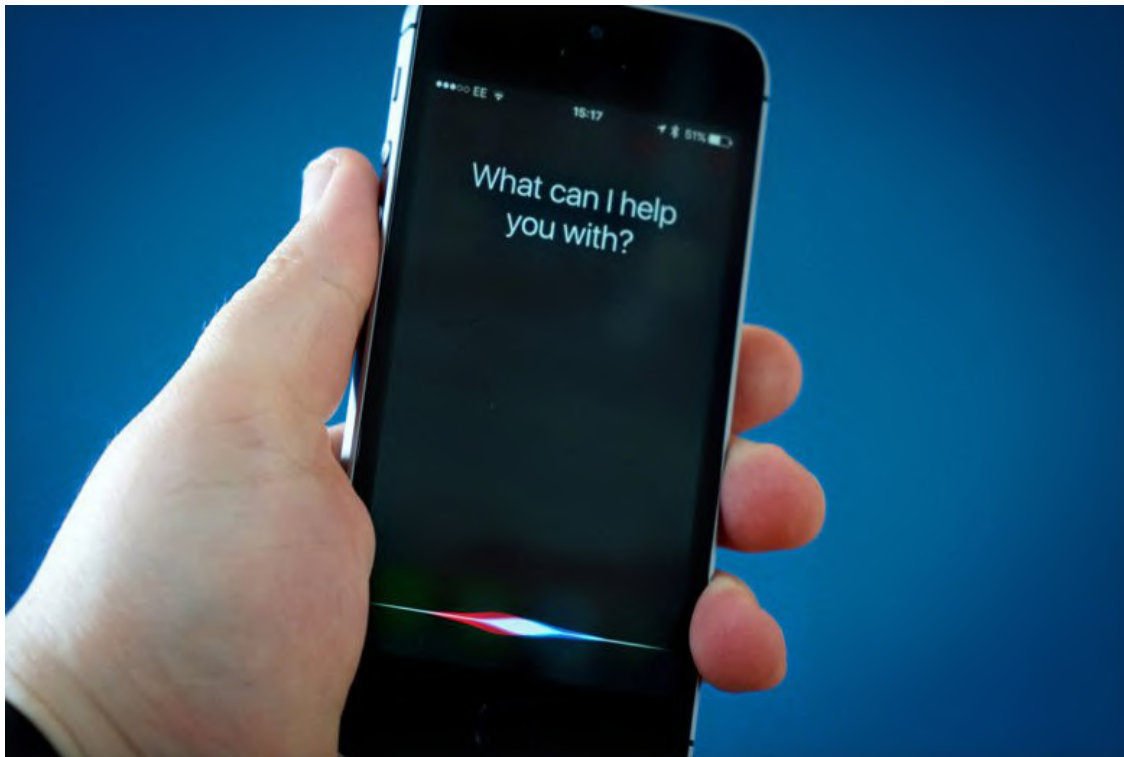
Do you know Eliza

Tap to Edit >

ELIZA is my good friend. She was a brilliant psychiatrist, but she's retired now.

Search the web for ELIZA

Assistenti vocali come Siri e Alexa



Assistenti vocali e “banal deception”

- Chi usa strumenti come Alexa e Siri mostra di comprendere le distinzioni tra umano e macchina, ma allo stesso tempo adotta comportamenti sociali per ‘dialogare’ con queste tecnologie (Nass and Moon 2000; Guzman 2019; ...)
- Superamento dell’approccio binario (umano o computer) proprio del test di Turing

Modi in cui la deception viene banalizzata dagli assistenti vocali

- Voce sintetica “umana”
- Genere (spesso femminile)
- Nome
- Caratterizzazione e uso dell’ironia

(vedi anche: Natale, S. (2020). To believe in Siri: A critical analysis of AI voice assistants. *Communicative Figurations Working Papers*, 32, 1–17)

“Banal deception” di Alexa e Siri

- L’obiettivo **non** è di far credere agli utenti che questi strumenti siano “umani”
- Piuttosto, la voce umana, il genere, e le altre forme di caratterizzazione hanno l’obiettivo di creare un **senso di continuità** nella relazione con l’assistente vocale, aiutandoci a integrarlo nel vivere quotidiano e nell’ambiente domestico

Dinamiche di inganno “banale” negli assistenti vocali

- **Collaborazione** con gli utenti che contribuiscono alla produzione di senso e attivando le proprie interpretazioni e rappresentazioni
- **Valore** per l'utente: rende più facile incorporare questi strumenti nella vita quotidiana
- **Normalizzazione** dell'inganno, che non viene considerato come tale ma è invece dissolto nel quotidiano

Conclusioni

MACCHINE INGANNEVOLI





NASA's Perseverance Mars Rover

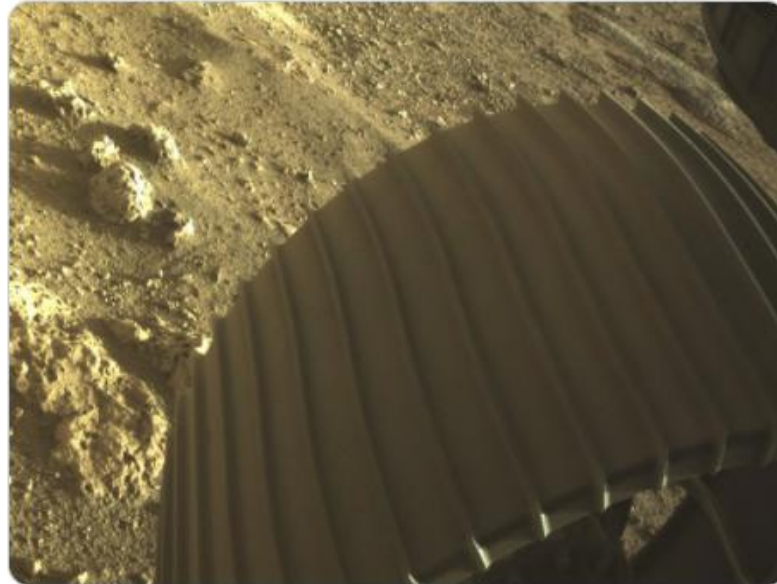
@NASAPersevere



I love rocks. Look at these right next to my wheel. Are they volcanic or sedimentary? What story do they tell? Can't wait to find out.

[#CountdownToMars](#)

mars.nasa.gov/mars2020



7:29 PM · Feb 19, 2021 · Twitter Web App

10.5K Retweets 1,556 Quote Tweets 86.4K Likes



Joanna J Bryson

@j2bryson



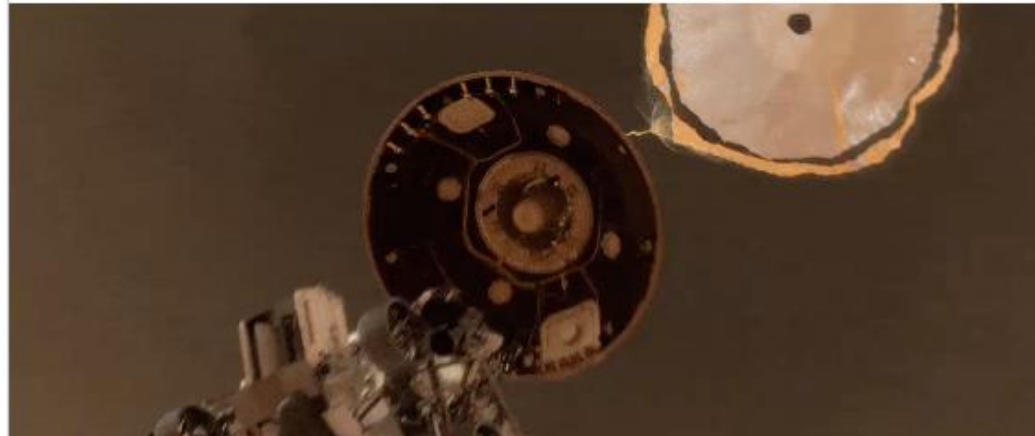
More [#anthropomorphised](#) AI. You're a human on earth lying to children about what robots can do / "think", instead of exposing some of our greatest engineering as what it actually is, an awesome feat of design.



NASA's Perseverance Mars Rover @NASAPersevere · Feb 18

Here goes! Lighting the engines on my "jetpack" for final descent. Wheels down in less than a minute.

[#CountdownToMars](#)



Superare la dialettica autenticità/inganno

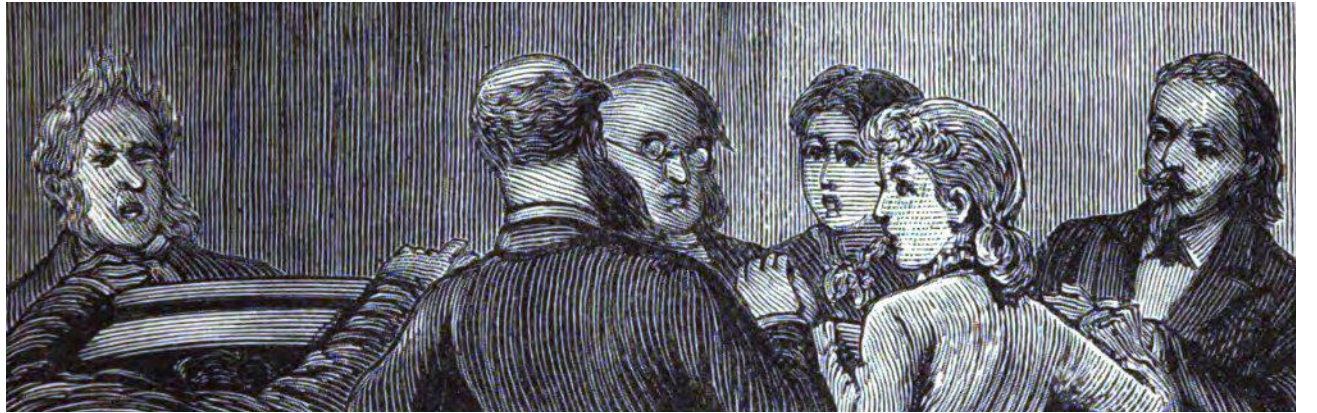
- Approccio alternativo: l'inganno per l'IA è una componente **strutturale**, così come i circuiti, il software e i dati che la fanno funzionare
- Il punto non è di capire se ci sia inganno, ma quale sia l'impatto potenziale in specifiche situazioni e con specifiche tecnologie

Conclusioni

- Intelligenza Artificiale come parte di una lunga storia dei media e della “banal deception”
- Domande urgenti da porre: non tanto sull’invasione dei robot o sul raggiungimento di forme di coscienza artificiale, quanto sull’impatto dell’IA **nel tessuto sociale e nella vita quotidiana**

Perchè urgente

- Possibili usi della tecnologia come “*strong*” *deception*
- L’abitudine alla *banal deception* nel quotidiano può influenzare le nostre percezioni su (e interazioni con) robot e computer come attori sociali
- Non solo, può anche influenzare i modi in cui interagiamo con umani (es. operatori di call center, stereotipi di genere)
- Necessità di sviluppare un approccio critico e informato alla base delle nostre interazioni con tecnologie di IA





10:44

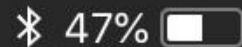


Hey Siri do you believe in spirits

Tap to Edit >



10:44



Hey Siri do you believe in spirits

Tap to Edit >

I don't believe that I have beliefs.

DECEITFUL MEDIA



Artificial Intelligence and
Social Life after the Turing Test

SIMONE NATALE

Grazie!

simone.natale@unito.it



UNIVERSITÀ
DEGLI STUDI
DI TORINO