INTRODUZIONE

Obiettivo: contestualizzare la tesi, chiarire il tema, il metodo e le motivazioni personali.
Contenuti:

- Introduzione personale al tema (origine dell'interesse per le interfacce).
- Il contesto: perché oggi parlare di interfacce è cruciale.
- Obiettivi e domande di ricerca.
- Metodo e approccio sistemico utilizzato.
- Struttura della tesi.

1. CONTESTO E CORNICE TEORICA

Obiettivo: costruire la base culturale e storica del discorso.
Contenuto:

1. **Breve storia delle interfacce uomo–macchina**

- L'evoluzione delle interfacce uomo–macchina: dal comando alla manipolazione diretta
- Dall'interfaccia esperienziale all'interfaccia cognitiva: verso l'era dell'intelligenza artificiale

2. **Storia dell'intelligenza artificiale: dall'automazione alla co-intelligenza**
    - Origini cibernetiche e visioni fondative (Turing, Wiener, McCarthy).
    - L'AI simbolica e la stagione dei sistemi esperti.
    - La crisi del paradigma logico e la nascita del connessionismo.
    - L'AI pervasiva e predittiva: machine learning, big data e algoritmi quotidiani.
    - L'AI generativa e dialogica: verso la co-evoluzione uomo–macchina.

3. **Il ruolo del design delle interfacce**
    - Design dell'interazione, UX e design sistemico: intersezioni.
    - Le interfacce come mediatori cognitivi e culturali.

4. **Approccio sistemico all'interfaccia**

- L'interfaccia come nodo in una rete di relazioni tra umani, dati e dispositivi.
- Interfacce come sistemi

2. L'INTELLIGENZA ARTIFICIALE COME AGENTE DI CAMBIAMENTO

Obiettivo: capire come l'AI stia ridefinendo il concetto di interfaccia.
Contenuto: Il capitolo 2 si propone di capire analiticamente e quindi spiegare come il concetto di interfaccia si stia ridefinendo dall'arrivo dell'AI come piattaforma per far funzionare il sistema, che cerca di completare l'ibridazione tra realtà e digitale. Se l'interfaccia fin'ora è sempre stata considerata solo attraverso la sua funzione, ossia di punto di collegamento tra due sistemi,

come cambia questo concetto quando l'interfaccia stessa non permette la chiarezza delle informazioni, facendo così da filtro?

- **2.1.1 Dalla macchina reattiva all'agente cognitivo**
- **2.1.2 Anatomia dell'intelligenza artificiale nelle interfacce contemporanee**
- **2.1.3 Nuovi paradigmi di interazione nell'era dell'AI**
- **2.1.4 Bias mitigation**
- **2.1.5 Il gap tra accountability tecnica e responsabilità etica**
- **2.1.6 Concretizzare l'etica: dalla policy al design**
- **2.1.7 Ripensare la "buona interazione": nuovi criteri di qualità**
- **2.1.8 Modelli mentali e comprensibilità: progettare per l'opacità**
- **2.1.9 Dimensione etica e responsabilità progettuale**

2.2 Questioni etiche generali.

- **2.2.1. Armi autonome e deumanizzazione della guerra**
- **2.2.2. Disuguaglianza algoritmica e discriminazione**
- **2.2.3. Privacy, controllo dei dati e autonomia digitale**
- **2.2.4. Disinformazione e polarizzazione**
- **2.2.5. Impatto ambientale dell'IA**
- **2.2.6. Tutela dei minori nell'era dell'IA**

3. ANALISI DELLO STATO DELL'ARTE

Obiettivo: mappare e interpretare le principali tendenze contemporanee.
Contenuto:

1. **Interfacce digitali:**

- Evoluzione di UI e UX con l'AI (es. sistemi multimodali, adaptive UI, agenti embedded).

2. **Interfacce fisiche:**

- Oggetti e dispositivi emergenti (wearable intelligenti, assistenti fisici, nuovi hardware post-smartphone).

3. **Interfacce vocali e conversazionali:**

- Da Alexa a ChatGPT: verso un'interazione più umana o più artificiale?

4. **Casi studio rilevanti:**

- Analisi critica di 4–6 casi (es. Humane Ai Pin, Rabbit R1, Rewind Pendant, Apple Intelligence, ecc.).

4. INTERPRETAZIONE SISTEMICA E SCENARI FUTURI

Obiettivo: capire come le informazioni dette finora possano ampliare la visione del sistema delle interfacce.
Contenuto:

1. **L'interfaccia come ecosistema**

- Attori, flussi informativi, comportamenti e valori in gioco.

2. **Scenari di transizione**

- Cosa accade quando l'interfaccia diventa "trasparente"?

- Dispositivi senza schermo e interazione ambientale.

3. **Implicazioni sociali, etiche e cognitive**

- Autonomia, fiducia, responsabilità, dipendenza.

4. **Prospettive per il design**

- Come cambia il ruolo del designer di interfacce.

- Verso un design relazionale e adattivo.

5. CONTRIBUTI PROGETTUALI

Obiettivo: proporre una visione progettuale speculativa basata sulle conclusioni.
Contenuto:

- Concept o scenari visualizzati (prototipo concettuale, interfaccia speculativa,).
- Diagrammi sistemici per rappresentare nuove relazioni utente–AI.
- Possibili linee guida per designer del futuro.

6. CONCLUSIONI

Obiettivo: sintetizzare i risultati e le riflessioni.
Contenuto:

- Risposte alle domande di ricerca.
- Principali insight teorici e progettuali.
- Limiti del lavoro e possibili sviluppi futuri.

BIBLIOGRAFIA E SITOGRAFIA

Introduzione

L'interesse per il tema di questa tesi nasce dalla consapevolezza di come le innovazioni tecnologiche stiano progressivamente rimodellando le dinamiche dell'interazione, non solo tra esseri umani e oggetti, ma anche tra gli esseri umani stessi.
Durante il mio percorso in Design Sistemico ho maturato l'idea che il ruolo del designer non si limiti a dare forma a soluzioni funzionali, ma implichi la responsabilità di **dare forma al futuro**, ancor prima che quel futuro si manifesti.
In questo senso, ogni interfaccia non è soltanto un punto di contatto con la tecnologia, come vuole la sua definizione, ma un dispositivo che inesorabilmente plasma comportamenti, percezioni e valori umani e quindi il nostro stesso futuro come specie. Basti pensare a come, in pochi anni, la diffusione di Internet e dello smartphone abbia trasformato radicalmente i nostri modi di comunicare, apprendere e percepire il mondo. Per questo ritengo urgente interrogarsi su come progettare oggi l'interazione con quella che molti considerano la prossima grande soglia evolutiva della tecnologia: l'intelligenza artificiale, forse il primo vero esempio di intelligenza aliena con cui, come sostiene Ethan Mollick (2023) della Wharton University, siamo chiamati a convivere. Una forma di co-intelligenza, capace di interrogarci sui nostri comportamenti e valori, e di rifletterli indietro verso di noi come elementi che influenzeranno in modo determinante il futuro della progettazione e della società.

Il contesto: perché è importante parlare oggi di interfacce

Negli ultimi decenni, le interfacce hanno assunto un ruolo centrale nella definizione del rapporto tra esseri umani e tecnologia.
Da semplici superfici di comando, concepite per tradurre input in output all'interno di un modello computazionale lineare (come nel *command line interface* con Arpanet), si è passati alle interfacce grafiche e tattili, che hanno reso la comunicazione con la macchina più immediata, visiva e intuitiva. Oggi, con l'intelligenza artificiale, stiamo assistendo a una terza grande trasformazione: quella dall'interfaccia *reattiva* all'interfaccia *cognitiva*, capace di interpretare, apprendere e dialogare. Le interfacce diventano quindi spazi relazionali complessi, dove la comunicazione tra umano e macchina avviene in modo fluido, multimodale e contestuale.

Come osserva Lev Manovich (2001), l'interfaccia è diventata il vero linguaggio della cultura digitale: *"Sia la parola stampata sia il cinema acquisirono delle forme stabili che subirono solo dei minimi ritocchi nel tempo. Poiché il linguaggio del computer è affidato al software, teoricamente potrebbe continuare a cambiare in eterno. Ma c'è una cosa di cui possiamo essere sicuri, stiamo assistendo all'emergere di un nuovo linguaggio meta culturale, un nuovo fenomeno che sarà significativo almeno quanto il cinema e la parola stampata." (p.127)*. È attraverso l'interfaccia che il computer diventa un medium, e la cultura si traduce in software.
Oggi, con l'integrazione dell'intelligenza artificiale nei sistemi digitali, questo linguaggio si è trasformato: l'interfaccia non è più soltanto un canale di mediazione, ma un agente cognitivo in grado di interpretare, apprendere e reagire in modo autonomo ai comportamenti dell'utente.

Viviamo in un'epoca in cui le interfacce non si limitano più a mostrare informazioni, ma modellano attivamente percezioni, decisioni e relazioni sociali.
Shoshana Zuboff (2019) ha mostrato come, nel capitalismo della sorveglianza, i sistemi digitali non si limitino a rappresentare il comportamento umano, ma lo anticipino e lo orientino attraverso feedback continui. Gli algoritmi di raccomandazione di piattaforme come TikTok, Netflix o Spotify sono esempi di interfacce che apprendono preferenze implicite e le restituiscono come esperienze personalizzate, generando cicli di retroazione comportamentale.

Parallelamente, Sherry Turkle (2011) ha evidenziato come gli assistenti vocali e i robot sociali introducano nuove forme di intimità e dipendenza emotiva, mettendo in discussione la distinzione tra comunicazione umana e artificiale. Allo stesso modo, i social agiscono come interfacce di identità, mediando non solo ciò che comunichiamo, ma *chi siamo* nel mondo digitale.

Queste trasformazioni rivelano un cambiamento profondo: l'interfaccia non è più un filtro neutrale, ma una presenza pervasiva e invisibile che struttura la nostra esperienza del reale. Come afferma Donald Norman (2013), *"design is really an act of communication, which means having a deep understanding of the person with whom the designer is communicating"* (p. 221).

Oggi, questa comunicazione è sempre più delegata a sistemi intelligenti che apprendono e reagiscono autonomamente.

Sul piano materiale, assistiamo a una progressiva dissoluzione dei dispositivi: gli schermi si miniaturizzano, le interazioni si distribuiscono nello spazio e il concetto di interfaccia tende a fondersi con l'ambiente.
L' *ambient computing* — come mostrano i progetti di Google o di Humane Ai Pin — rende l'interazione diffusa, contestuale e predittiva. Già Mark Weiser (1991) aveva anticipato questo scenario parlando di *ubiquitous computing*, un'informatica "calma", invisibile ma onnipresente, capace di fondersi con il contesto fino a sparire dallo sguardo.
In questo scenario, parlare di interfacce significa interrogarsi non tanto su ciò che appare, ma su come interagiamo, con chi stiamo comunicando e quali sistemi di valore mediano ogni gesto, ogni sguardo e ogni decisione.

Dall'interfaccia statica al sistema adattivo: nuove sfide per il design

Per il design, questo scenario rappresenta una svolta. Le interfacce non sono più entità stabili e finite, ma organismi dinamici e adattivi: si modificano nel tempo in base ai dati, apprendono dai comportamenti degli utenti e si evolvono insieme ai contesti tecnologici e sociali. Questa transizione può essere interpretata attraverso la teoria dei sistemi complessi, secondo cui ogni elemento di un sistema è interdipendente e soggetto a retroazioni continue. Come scrive Fritjof Capra (1996), "le proprietà essenziali di un sistema vivente non risiedono nelle sue parti, ma nelle relazioni tra esse" (p. 29). In questa prospettiva, l'interfaccia può essere letta come un nodo vivo in una rete di relazioni tra umani, algoritmi, infrastrutture e dati, che ne ridefiniscono continuamente forma e significato.

Se per Capra la vita di un sistema emerge dalle sue interrelazioni, Luciano Floridi (2014) amplia questa visione descrivendo l'attuale fase dell'evoluzione umana come la "quarta rivoluzione", in cui la distinzione tra agenti umani e artificiali si fa sempre più sfumata. Progettare un'interfaccia oggi significa quindi progettare relazioni tra entità cognitive ibride, capaci di apprendere reciprocamente e di co-evolvere nel tempo.

In questo contesto, il pensiero sistemico del design diventa lo strumento per comprendere e agire nella complessità. Se ogni progetto implica una rete di relazioni materiali, sociali e informative, allora l'interfaccia va intesa non come un oggetto che risponde meccanicamente a input e output, ma come un ecosistema di interazione autopoietico, cioè capace di rigenerare continuamente i propri comportamenti in risposta all'ambiente.

Il ruolo del designer cambia di conseguenza: da autore di forme e funzioni a facilitatore di relazioni, capace di orchestrare processi di apprendimento e adattamento tra sistemi intelligenti e utenti umani.

Carl DiSalvo in *Adversarial Design (2012)* invita a riconoscere il design come pratica critica e politica ed ogni progetto come gesto politico, mai neutro, capace di far emergere tensioni e scambi tra attori umani e non umani. Seguendo la prospettiva di DiSalvo, il design delle interfacce può essere interpretato come pratica politica: ogni scelta di configurazione, accesso o trasparenza implica una negoziazione tra diversi attori: utenti, dati, algoritmi, istituzioni.

L'interfaccia diventa così un terreno di confronto e di possibile conflitto, in cui si decidono le modalità della relazione uomo–macchina e siamo chiamati a definire la responsabilità reciproca.
Parlare oggi di interfacce significa quindi interrogarsi su quale idea di relazione vogliamo progettare, su che grado di trasparenza e autonomia intendiamo concedere ai sistemi intelligenti, e su come (e se) preservare la comprensibilità umana in un mondo mediato da algoritmi.

Progettare interfacce, in questo senso, equivale a progettare forme di collaborazione cognitiva.
Significa ripensare l'interazione come campo di confronto etico e cognitivo, e il design come disciplina capace di rendere visibili le logiche invisibili dei sistemi.
Oggi, discutere di interfacce non è quindi unicamente un tema tecnico, ma una questione politica e culturale: il modo in cui interagiamo con l'AI definisce, in ultima analisi, chi stiamo diventando come specie.

Domande di ricerca e metodo

A partire da queste considerazioni, la tesi si propone di indagare come l'intelligenza artificiale stia trasformando il concetto stesso di interfaccia — da semplice strumento di mediazione tra umano e macchina a soggetto attivo di relazione, capace di apprendere, anticipare e adattarsi.
L'obiettivo è comprendere come cambia la progettazione dell'interazione in un contesto in cui l'AI diventa un agente cognitivo, e quali implicazioni questo comporti sul piano etico, sociale e progettuale.

Le principali domande di ricerca che orientano il lavoro sono:

1. In che modo l'intelligenza artificiale ridefinisce la funzione e il significato delle interfacce digitali, fisiche e vocali?

2. Quali nuove modalità di interazione emergono quando la tecnologia diventa adattiva, predittiva e co-creativa?

3. Come può il design sistemico supportare una comprensione più ampia delle interfacce come ecosistemi di relazioni tra umani, dati e dispositivi?

4. Quali competenze e responsabilità emergono per il designer di interfacce nell'era dell'AI?

Il metodo adottato unisce analisi teorica, osservazione critica e interpretazione sistemica.
La prima fase consiste nella costruzione di una cornice storico-culturale, che ripercorre l'evoluzione delle interfacce uomo–macchina e del pensiero sull'interazione.
Segue una ricerca esplorativa volta a mappare lo stato dell'arte delle interfacce contemporanee, attraverso l'analisi comparativa di casi studio significativi (come Humane Ai Pin, Rabbit R1, Rewind Pendant, Apple Intelligence).
Infine, l'approccio sistemico viene utilizzato come strumento interpretativo per individuare pattern relazionali, dinamiche emergenti e possibili scenari futuri di co-evoluzione tra umani e intelligenze artificiali.

Questo percorso permette di unire riflessione teorica e prospettiva progettuale, con l'intento di contribuire alla definizione di un nuovo modo di concepire le interfacce: non più come superfici di controllo, ma come spazi di relazione e apprendimento reciproco.

1. CONTESTO E CORNICE TEORICA

1.1 Breve storia delle interfacce uomo–macchina

L'evoluzione delle interfacce uomo–macchina: dal comando alla manipolazione diretta

Le interfacce uomo–macchina rappresentano la soglia in cui si articola la relazione tra esseri umani e sistemi tecnologici. La loro evoluzione storica riflette non solo il progresso dell'informatica, ma anche il modo in cui cambiano i paradigmi cognitivi e culturali dell'interazione.
Dalla nascita dei primi calcolatori elettronici, l'interfaccia si è progressivamente trasformata da **mezzo di controllo** a **spazio di esperienza**, accompagnando la transizione del computer da strumento specialistico a oggetto quotidiano.

1. Le origini: l'interfaccia come linguaggio di comando (anni '50–'70)

Le prime interfacce digitali si sviluppano in ambito militare e accademico, dove l'interazione con la macchina avveniva esclusivamente attraverso linguaggi formali.
Negli anni Cinquanta, i sistemi informatici come l'ENIAC o l'IBM 704 erano accessibili solo a ingegneri e programmatori: la comunicazione con la macchina avveniva mediante schede perforate o linee di testo.
In questa fase l'utente doveva conoscere il linguaggio della macchina, una forma di codice binario o simbolico, e inserire istruzioni precise affinché il sistema eseguisse i comandi.
Negli anni Sessanta e Settanta, con l'introduzione dei terminali teletype e dei sistemi operativi UNIX, l'interazione si innova, rimanendo di tipo testuale ma diventando più diretta: l'utente può scrivere comandi da tastiera e ricevere risposte in tempo reale.
Questa innovazione da vita al paradigma di quegli anni, chiamato Command Line Interface (CLI): un modello di interazione basato sul linguaggio e sul controllo, che rappresenta la logica basata sulla funzione tipica del primo rapporto uomo–macchina: l'umano impartisce ordini, la macchina obbedisce.

Dal punto di vista del significato, queste interfacce possono essere lette come spazi di traduzione: la macchina non “parla” all’utente, ma traduce simbolicamente i suoi comandi in processi computazionali.
La relazione è quindi asimmetrica e cognitiva: il potere comunicativo risiede nella conoscenza del linguaggio di codice, non nel gesto o nella percezione.

In questa fase, il design non è ancora una disciplina esplicitamente coinvolta. L’interfaccia è considerata una questione tecnica, non esperienziale: il valore sta quindi nell’efficienza del comando, non nella qualità dell’interazione. Come osserva Lev Manovich (2001), il computer degli albori non è ancora un “medium” ma un “motore simbolico”: uno strumento per manipolare dati più che per generare esperienze.

2. La rivoluzione grafica: la nascita dell’interfaccia visiva (anni ’70–’90)

A partire dagli anni Settanta, un gruppo di ricercatori dello Xerox Palo Alto Research Center (PARC) introduce un nuovo paradigma: la Graphic User Interface (GUI).
Con la GUI, l’interazione con la macchina non avviene più attraverso comandi testuali, ma tramite metafore visive, finestre, icone, menu e cursori, che rappresentano oggetti e azioni del mondo reale.

Il concetto di *desktop metaphor* (metafora della scrivania) nasce da questa intuizione: rendere il computer uno spazio familiare, riconoscibile e manipolabile.Come racconta Alan Kay, uno dei pionieri dello Xerox PARC, l’obiettivo era “rendere l’uso del computer così intuitivo che un bambino potesse impararlo da solo”. L’idea si concretizza nei sistemi Xerox Alto (1973) e Star (1981), poi ripresa e portata al grande pubblico da Apple Lisa (1983) e Macintosh (1984), e infine da Microsoft Windows (1985).

Con l’introduzione della GUI, il computer smette di essere una macchina da controllare e diventa un ambiente da esplorare. Le metafore visive permettono all’utente di “vedere” l’azione che compie: nasce il principio di manipolazione diretta, teorizzato da Ben Shneiderman (1983) e poi elaborato da Donald Norman (1988), secondo cui l’interazione deve essere *immediata, reversibile e visivamente comprensibile*.

In questo nuovo contesto, il design assume un ruolo fondamentale: tradurre la complessità computazionale in forme percepibili e coerenti con i modelli mentali degli utenti.
Come scrive Norman, “il design è un atto di comunicazione” (2013, p. 221): l’interfaccia diventa perciò un linguaggio condiviso tra l’intenzione del progettista e l’interpretazione dell’utente.

La GUI segna così il passaggio fondamentale da un’interazione *funzionale* a una *cognitiva e percettiva*: il computer comincia a diffondersi diventando uno strumento sociale, non più solo tecnico. L’estetica dell’interfaccia non è più un orpello decorativo, ma il vero e proprio comunicatore di significato: permettendo di costruisce fiducia con l’utente.

3. Dall’interfaccia visiva all’interfaccia esperienziale (anni ’90–2000)

Negli anni Novanta, il computer smette ufficialmente di essere una macchina per pochi e diventa una presenza domestica. L’interfaccia, di conseguenza, cambia funzione: non serve più solo a far funzionare un sistema, ma a renderlo vivibile. L’interazione non riguarda più

l’esecuzione di comandi, ma la costruzione di esperienze. Con la nascita del World Wide Web nel 1991, si apre una nuova frontiera dell’interazione digitale. Browser come Mosaic (1993) e Netscape Navigator (1994) introducono l’idea di “navigazione”: l’utente non digita più istruzioni, ma esplora spazi informativi attraverso link e percorsi visivi. Ogni clic diventa un passaggio di senso, un attraversamento: la rete si comporta come un ambiente cognitivo più che come una macchina. Come nota Lev Manovich (2001), l’interfaccia del Web rappresenta “l’estetica del software”: una forma culturale che modella la percezione, l’apprendimento e il modo stesso di organizzare la conoscenza.

È in questi anni che il computer diventa anche uno strumento narrativo e sociale. Con l’avvento dei blog, delle chat e dei primi forum, l’utente non è più un operatore ma un partecipante: costruisce identità, relazioni e contenuti. L’interfaccia diventa un palcoscenico di espressione personale. Le pagine HTML, i layout di Geocities o i primi siti personali sono esempi emblematici di questa fase: spazi digitali progettati non solo per comunicare informazioni, ma per rappresentare sé stessi.

Parallelamente, la tecnologia si fa più personale e portatile. I laptop e i primi notebook permettono di spostare l’esperienza informatica fuori dall’ufficio, mentre la miniaturizzazione dei componenti apre la strada a una nuova idea di mobilità. Nascono i primi assistenti digitali personali (PDA), come il Palm Pilot (1996) o l’Apple Newton (1993), dispositivi che consentono di annotare, pianificare e sincronizzare attività: l’interfaccia comincia a dialogare con la vita quotidiana. In questa fase prende forma anche il concetto di User Experience (UX), introdotto da Donald Norman alla fine degli anni Novanta, per indicare la totalità delle percezioni, emozioni e aspettative che accompagnano l’interazione con un sistema. Non si tratta più di rendere la tecnologia usabile, ma di renderla significativa. Verso la fine del decennio, l’arrivo dei touch screen segna un nuovo passaggio evolutivo. I Tablet PC (1999) e i dispositivi mobili con input tattile introducono una grammatica dell’interazione basata sul gesto. Il contatto diretto con lo schermo — il tocco, la pressione, lo scorrimento — restituisce all’interfaccia una dimensione sensoriale.

Non è più necessario dire alla macchina cosa fare: basta toccarla. Questa umanizzazione del gesto anticipa l’estetica del design contemporaneo: un’interfaccia più naturale, empatica e multimodale, dove il confine tra azione e percezione si dissolve. Il computer, ormai, non si limita a rispondere: ascolta, accompagna, si adatta. Il design dell’interazione entra così in una nuova era, in cui l’esperienza diventa il vero linguaggio del digitale.

Dall’interfaccia esperienziale all’interfaccia cognitiva: verso l’era dell’intelligenza artificiale

4. L’era del touch e della mobilità (2000–2010)

Con l’inizio del nuovo millennio, il computer abbandona la postazione fissa e diventa dispositivo personale e portatile.
La miniaturizzazione dei processori e la diffusione di connessioni wireless trasformano il modo in cui l’utente percepisce la tecnologia: il digitale non è più un luogo separato, ma un’estensione continua del corpo e dell’ambiente.

L’avvento dello smartphone, inaugurato simbolicamente dall’iPhone (2007), segna una vera cesura nella storia delle interfacce.
Il gesto sostituisce il comando: *pinch*, *tap*, *swipe* diventano un linguaggio universale.
Il contatto diretto tra mano e schermo crea un’esperienza di immediata continuità percettiva, eliminando l’intermediazione di tastiera e mouse.
Come nota Bill Moggridge (2007), uno dei fondatori dell’interaction design, “le interfacce tattili non sono semplicemente strumenti di input, ma un’estensione sensoriale della mente e del corpo umano”.

In questa fase il design dell’interazione si concentra su usabilità, consistenza e immediatezza percettiva.
Donald Norman e Jakob Nielsen formalizzano i principi dell’User Experience (UX), spostando l’attenzione dal funzionamento tecnico alla qualità soggettiva dell’esperienza.
Nascono le prime linee guida del design mobile e del “material design”, basate su fluidità, feedback, e transizioni naturali.

Parallelamente, il digitale diventa esperienza ambientale.
Il telefono non serve più solo a comunicare, ma a orientarsi, acquistare, misurare, ricordare: ogni gesto quotidiano passa attraverso un’interfaccia.
La tecnologia si insinua nel quotidiano in modo invisibile e costante.
Come anticipato da Mark Weiser (1991) con il concetto di *ubiquitous computing*, la tecnologia più avanzata è quella che scompare, intrecciandosi nel tessuto della vita quotidiana fino a diventare indistinguibile da essa.

5. L’interfaccia diffusa e predittiva: ambient computing e machine learning (2010–2020)

Con l’esplosione dei big data, dell’Internet of Things (IoT) e dell’intelligenza artificiale applicata, l’interfaccia attraversa una nuova trasformazione: da spazio esperienziale a sistema distribuito e adattivo.
L’interazione non avviene più su uno schermo, ma attraverso una rete di oggetti intelligenti, sensori e algoritmi che mediano la relazione tra utente e ambiente.

Esempi emblematici di questa transizione sono gli assistenti vocali (Siri, 2011; Alexa, 2014; Google Assistant, 2016), che inaugurano il paradigma dell’interazione naturale, e gli ecosistemi domotici (Nest, Philips Hue, Amazon Echo), dove la tecnologia risponde al contesto più che al comando.
L’interfaccia diventa predittiva, capace di anticipare bisogni e preferenze attraverso l’analisi dei dati.

Secondo la visione di Mark Weiser e John Seely Brown, questo è il momento in cui si realizza il concetto di *calm technology*: una tecnologia che si muove tra il centro e la periferia

dell’attenzione, interagendo senza interrompere.
In questa prospettiva, il design deve affrontare un nuovo paradosso: progettare l’invisibile, cioè rendere percepibile e comprensibile un sistema che tende per natura a nascondersi.

Il paradigma dell’ambient computing pone inoltre questioni etiche e sistemiche:

- come rendere visibile ciò che agisce senza apparire?
- come mantenere agency umana in un ambiente che decide al nostro posto?

L’interfaccia si configura così come un campo di mediazione cognitiva, in cui l’utente non controlla più direttamente la tecnologia, ma coesiste con essa in un equilibrio dinamico.

6. L’interfaccia cognitiva e conversazionale (2020–oggi)

L’attuale fase, segnata dall’emergere dell’intelligenza artificiale generativa, inaugura un nuovo paradigma: quello dell’interfaccia dialogica e cognitiva.
Sistemi come ChatGPT (OpenAI, 2022), Gemini (Google, 2023) o Claude (Anthropic, 2023) introducono un’interazione fondata sulla lingua naturale, dove il linguaggio non è più un codice tecnico, ma un mezzo di pensiero condiviso tra umano e macchina.

In questa nuova condizione, l’interfaccia si comporta come un agente cognitivo capace di interpretare, apprendere e rispondere in modo autonomo.
Il modello tradizionale input–output si dissolve: la macchina diventa *partecipante* nella costruzione del significato.
Come osserva Ethan Mollick (2023), queste tecnologie rappresentano “la prima forma di co-intelligenza aliena”, un’intelligenza non umana ma cooperante, capace di amplificare e mettere in discussione le capacità cognitive dell’uomo.

Dal punto di vista progettuale, questo scenario ridefinisce radicalmente il ruolo dell’interfaccia.
Essa non è più il filtro tra umano e macchina, ma il luogo della relazione cognitiva — un dispositivo di negoziazione tra due agenti che apprendono reciprocamente.
Come già suggeriva Brenda Laurel (1991), l’interazione può essere letta come una forma di *dramma* o di *performance*: un dialogo continuo tra intenzione e risposta, tra previsione e sorpresa.
L’intelligenza artificiale accentua questa dimensione drammatica, portando l’interfaccia a diventare autrice di senso insieme all’utente.

1.2 Storia dell’intelligenza artificiale: dall’automazione alla co-intelligenza

1. Dalle macchine logiche all’idea di intelligenza artificiale (1940–1960)

Le radici dell’intelligenza artificiale affondano nella cibernetica e nella logica computazionale sviluppate nel secondo dopoguerra.

Alan Turing, nel celebre articolo *Computing Machinery and Intelligence* (1950), pone la domanda fondativa: *“Can machines think?”* — e introduce il test di Turing come criterio operativo per valutare se una macchina possa imitare il comportamento cognitivo umano attraverso il linguaggio.
Parallelamente, Norbert Wiener (1948) definisce la cibernetica come la scienza del controllo e della comunicazione nei sistemi biologici e meccanici, anticipando l’idea che l’intelligenza possa emergere da processi di *feedback* e autoregolazione.

Negli anni successivi, le prime sperimentazioni tentano di tradurre questa visione teorica in modelli computazionali concreti.
Tra i pionieri spiccano Claude Shannon, che nel 1950 dimostra come una macchina possa giocare a scacchi applicando regole logiche e probabilistiche, e Marvin Minsky con SNARC (Stochastic Neural Analog Reinforcement Calculator, 1951), un simulatore elettronico di reti neurali ispirato ai processi sinaptici del cervello.
Nel 1956, durante la celebre Dartmouth Conference, John McCarthy conia ufficialmente il termine *Artificial Intelligence*, sancendo la nascita di una disciplina autonoma e di un ambizioso programma di ricerca collettiva.

In questa prima stagione, dominata da un entusiasmo quasi utopico, si ritiene che basti formalizzare il pensiero umano in un linguaggio logico per ottenere comportamenti intelligenti. Nascono così i primi programmi in grado di risolvere problemi matematici o di manipolare simboli, come il Logic Theorist (1956) di Allen Newell e Herbert Simon, considerato la prima vera AI, puntava a replicare l’utilizzo umano della logica per risolvere problemi; si dimostrò capace di dimostrare teoremi di logica proposizionale: un sistema che serve a rappresentare e analizzare ragionamenti usando frasi che possono essere vere o false.
Le interfacce con queste intelligenze artificiali sono puramente testuali e simboliche: l’interazione avviene tramite comandi scritti, formule e regole deduttive.
In questo periodo l’intelligenza viene concepita come ragionamento computabile, un processo lineare e deterministico che riduce il pensiero a un insieme di istruzioni formali decodificabili e quindi eseguibili dalla macchina.

2. L’AI simbolica e la stagione dei sistemi esperti (1960–1980)

Negli anni Sessanta e Settanta si avvia una seconda fase in cui prende forma la cosiddetta AI simbolica, basata sull’idea che il pensiero umano possa essere rappresentato come una sequenza di simboli manipolabili da regole logiche.
Si tratta di un’intelligenza “deduttiva”, che non apprende ma ragiona per inferenza: data una serie di premesse, elabora conclusioni come un matematico o un filosofo formale, in questi anni “pensare” per le macchine significava manipolare simboli in modo coerente, senza alcuna comprensione semantica, applicando regole formali a simboli privi di significato intrinseco.

I primi esperimenti si concentrano sulla risoluzione di problemi logici o linguistici, ma presto si estendono a campi applicativi più specifici.
Un esempio emblematico di questa fase è ELIZA (1966), sviluppato da *Joseph Weizenbaum* al MIT. Il programma era concepito per mostrare come una macchina potesse simulare una conversazione umana attraverso semplici regole di riconoscimento linguistico, senza alcuna

comprensione semantica del contenuto. Uno dei suoi script più noti, *DOCTOR*, imitava il comportamento di un terapeuta ispirato allo stile non direttivo di Carl Rogers, ponendo domande aperte e riformulando le frasi dell'utente. Questa scelta non derivava da un intento clinico, ma da una precisa dimostrazione tecnica: il linguaggio terapeutico, basato sull'ascolto riflessivo, era ideale per rivelare quanto facilmente gli esseri umani potessero attribuire empatia e intelligenza a un sistema puramente sintattico.

Paradossalmente, proprio questa simulazione portò molti utenti a proiettare emozioni e confidenze sul programma, segnando il primo esempio di relazione "affettiva" uomo–macchina e anticipando i dilemmi etici ancora centrali nell'interazione con le AI contemporanee. Quando l'utente scriveva una frase la macchina rispondeva con domande riformulate ("Mi parli di più di questo…", "Perché pensa che sia così?"). ELIZA fu installata in laboratori e università, ma anche utilizzata come strumento didattico per esplorare la psicologia dell'interazione uomo–macchina.

Parallelamente, l'AI entra nei contesti professionali attraverso i sistemi esperti, programmi capaci di replicare il ragionamento di specialisti umani, uscendo quindi dai laboratori di linguistica e logica per entrare ufficialmente nei domini specialistici della scienza e dell'industria. I cosiddetti sistemi esperti: programmi progettati per replicare il processo decisionale di specialisti umani in campi circoscritti, si propongono di tradurre il sapere esperto in centinaia di regole "if–then". Uno dei primi e più importanti fu DENDRAL (1965), sviluppato alla Stanford University da Edward Feigenbaum e Joshua Lederberg.

DENDRAL analizzava dati di spettrometria di massa per dedurre la struttura molecolare di composti chimici. Fu il primo sistema a ottenere risultati scientificamente rilevanti: riuscì a proporre ipotesi corrette sulla composizione di molecole complesse, pubblicate poi su riviste di chimica. DENDRAL dimostrò che un programma poteva effettivamente contribuire alla scoperta scientifica.

Qualche anno più tardi, con l'obiettivo di diagnosticare infezioni batteriche e suggerire terapie antibiotiche, lo stesso gruppo di Stanford sviluppò MYCIN (1972), destinato al campo medico, mostrando un'accuratezza paragonabile a quella di medici esperti in test controllati. Il sistema non fu mai utilizzato clinicamente, principalmente per ragioni etiche e di responsabilità, ma introdusse un concetto che sarebbe diventato centrale: la spiegabilità dell'intelligenza artificiale. Ogni decisione di MYCIN era infatti accompagnata da una motivazione ("MYCIN consiglia X perché ha rilevato Y"), ponendo per la prima volta il problema della trasparenza algoritmica, una questione che oggi torna d'attualità con i sistemi di machine learning opachi o "black box".

Verso la fine degli anni Settanta, la ricerca si estese anche a contesti industriali. Il sistema PROSPECTOR (1978), sviluppato dal SRI International, utilizzava dati geologici per identificare aree promettenti per l'estrazione mineraria. Nel 1980, PROSPECTOR portò effettivamente alla scoperta di un giacimento di molibdeno a Mount Tolman (Washington), dimostrando l'applicabilità dell'AI come strumento operativo di decisione strategica nel mondo reale.

Dal punto di vista del design dell'interazione, la stagione dell'AI simbolica introduce per la prima volta la necessità di tradurre conoscenze complesse in forme comunicabili. Le interfacce testuali e decisionali dei sistemi esperti, spesso criptiche e riservate a specialisti,

rendono evidente quanto la comprensibilità umana diventi una componente cruciale dell'intelligenza artificiale.

Tuttavia, l'entusiasmo iniziale lasciò spazio alle prime critiche teoriche.
Hubert Dreyfus, in *What Computers Can't Do* (1972), contestò l'idea che l'intelligenza fosse riducibile a logica e simboli: il pensiero umano, sosteneva, è incarnato, situato e intrinsecamente ambiguo, e quindi non formalizzabile in un insieme di regole astratte.
Questa critica avrebbe aperto la strada alla crisi del paradigma simbolico e al successivo cambio di paradigma verso modelli più ispirati alla natura ed adattivi.

3. La crisi del simbolismo e la nascita del connessionismo (1980–2000)

Negli anni Ottanta, l'intelligenza artificiale entra in una fase di profonda crisi, nota come "AI winter", dovuta al fallimento delle promesse del paradigma simbolico.
I sistemi basati su regole logiche mostravano limiti evidenti: funzionavano solo in contesti chiusi e prevedibili, incapaci di affrontare l'incertezza, l'ambiguità e il senso comune che caratterizzano il pensiero umano.

Il filosofo Hubert Dreyfus, in *What Computers Can't Do* (1972), aveva già previsto questo esito. Criticava l'idea che l'intelligenza potesse essere ridotta a manipolazione simbolica, sostenendo che il sapere umano non è puramente astratto ma situato, corporeo e contestuale. Secondo Dreyfus, gran parte delle nostre decisioni e percezioni non derivano da regole logiche, ma da esperienze incarnate e da una comprensione tacita del mondo.
Le macchine, al contrario, operavano in assenza di corpo e contesto, producendo conoscenza formale ma non *intelligenza viva*. Queste critiche coincisero con il rallentamento dei finanziamenti pubblici e con la crescente sfiducia nei risultati dell'AI simbolica. Tuttavia, proprio in questo clima di scetticismo, si sviluppò un nuovo approccio che avrebbe rinnovato l'intera disciplina: il connessionismo. A differenza del simbolismo, il connessionismo non cerca di codificare la mente umana in regole, ma di simularne la struttura attraverso reti di unità semplici collegate tra loro: i neuroni artificiali.
Il principio di base è che l'intelligenza emerge non dalle regole logiche, ma dall'interazione dinamica tra elementi che apprendono dai dati e si adattano tramite feedback. Nel 1986, *David Rumelhart*, *Geoffrey Hinton* e *Ronald Williams* pubblicano l'articolo che segna il rilancio dell'intelligenza artificiale: *Learning Representations by Back-Propagating Errors*.
In esso viene descritto un principio chiave del connessionismo: una rete neurale può apprendere dai propri errori. Il funzionamento di questi modelli si ispira in modo astratto al cervello umano.
Una rete neurale è composta da unità elementari (neuroni artificiali) organizzate in strati:

- uno strato di input, che riceve i dati (ad esempio, pixel di un'immagine o parole di una frase);
- uno o più strati nascosti, che trasformano l'informazione attraverso connessioni pesate;
- e uno strato di output, che produce il risultato (come una classificazione o una previsione).

Ogni connessione tra neuroni possiede un peso, che rappresenta la forza della relazione tra i due nodi. Durante l'addestramento, la rete produce un risultato iniziale che viene confrontato con il risultato corretto; la differenza (errore) viene poi utilizzata per modificare i pesi delle connessioni interne, rinforzando quelle che hanno contribuito a risposte corrette e indebolendo le altre. Questo processo, chiamato retropropagazione dell'errore (backpropagation), consente al sistema di migliorare progressivamente le proprie prestazioni.

In questo modo, l'intelligenza non deriva da un insieme di regole imposte dall'esterno, ma emerge dall'esperienza. La macchina apprende riconoscendo regolarità nei dati, costruendo progressivamente una propria rappresentazione del mondo. Il sapere diventa distribuito (diffuso tra migliaia di connessioni), graduale (migliora con l'addestramento) e probabilistico (non produce certezze, ma stime).

Il connessionismo sostituisce così l'idea di mente come sistema simbolico con quella di mente come rete adattiva. Non si tratta più di dire alla macchina *cosa* fare, ma di creare le condizioni perché impari *come* farlo osservando esempi e reagendo a feedback.
Questo passaggio segna l'origine della moderna intelligenza artificiale basata sull'apprendimento (*machine learning*), e apre la strada a una nuova forma di relazione uomo–macchina fondata sulla collaborazione cognitiva. Questa svolta segna anche un cambiamento profondo nel modo di intendere l'interfaccia uomo–macchina. L'utente non è più soltanto colui che impartisce comandi, ma diventa parte integrante del processo di apprendimento: i dati che produce, i suoi errori, le sue scelte diventano materiale con cui la macchina si addestra e affina i propri modelli interni. L'interazione non è più lineare e deterministica: input, elaborazione, output, ma dialogica e retroattiva, in cui ogni azione umana genera una risposta che, a sua volta, influenza il comportamento futuro del sistema.

Negli anni Ottanta e Novanta, questo passaggio si manifesta concretamente in una serie di esperimenti e interfacce che, pur rudimentali, anticipano le logiche dell'apprendimento iterativo. Un esempio è il sistema ALVINN (Autonomous Land Vehicle In a Neural Network), sviluppato nel 1989 alla Carnegie Mellon University da Dean Pomerleau: un veicolo autonomo in grado di imparare a guidare analizzando i dati visivi raccolti durante la navigazione. L'interfaccia non era più un pannello di controllo umano, ma un sistema percettivo che si addestrava osservando l'operatore. Ogni intervento del conducente diventava un feedback che migliorava il modello neurale di guida. Analogamente, nei laboratori di ricerca sull'interazione grafica, come quelli di Xerox PARC o di Apple negli anni Novanta, cominciano a emergere interfacce capaci di adattarsi all'utente. Nei sistemi di riconoscimento vocale e scrittura manuale (ad esempio Apple Newton, 1993), la macchina correggeva progressivamente la propria interpretazione in base alle interazioni ripetute, apprendendo dallo stile e dal lessico individuale dell'utente. Anche in questo caso, l'interfaccia non si limitava a ricevere comandi, ma costruiva nel tempo una relazione di apprendimento reciproco. Il passaggio dall'automazione al co-apprendimento si compie così in modo graduale ma decisivo. Nell'automazione tradizionale, l'uomo progettava la macchina affinché eseguisse una sequenza di istruzioni predefinite; nel co-apprendimento, invece, la macchina impara dall'uomo mentre l'uomo impara dalla macchina come interagire con essa in modo più efficace. Ogni ciclo di uso diventa un momento di addestramento: l'interfaccia

osserva, adatta e modifica le proprie risposte, trasformandosi in un organismo relazionale più che in un semplice strumento operativo. Questa visione apre la strada al paradigma dell'intelligenza artificiale contemporanea: i sistemi non vengono più programmati una volta per tutte, ma crescono attraverso l'interazione collettiva. Le prime reti neurali interattive, i software di correzione predittiva e i sistemi di raccomandazione sperimentali degli anni Novanta rappresentano l'inizio di un nuovo modello cognitivo distribuito — quello in cui ogni utente, con i propri comportamenti, contribuisce a "educare" la macchina. L'interfaccia diventa così spazio di addestramento e feedback, luogo in cui l'intelligenza si forma attraverso la relazione.

4. L'AI pervasiva e predittiva (2000–2020)

Con l'inizio del nuovo millennio, l'intelligenza artificiale abbandona i laboratori di ricerca e si diffonde nel tessuto della vita quotidiana.
L'aumento della potenza di calcolo, la nascita del web 2.0 e la disponibilità massiva di dati generati dagli utenti danno origine a un nuovo ecosistema informativo: l'AI non è più un oggetto isolato, ma una trama invisibile che connette servizi, dispositivi e persone.

I motori di ricerca imparano a rispondere alle intenzioni dietro le parole chiave, modellando la conoscenza collettiva a partire dalle query di miliardi di utenti;
i social network costruiscono flussi personalizzati di contenuti, regolando ciò che ogni individuo vede, pensa e condivide;
le piattaforme di e-commerce come Amazon raffinano i sistemi di raccomandazione, anticipando desideri e bisogni; gli assistenti digitali come Siri (2011) e Google Assistant (2016) diventano mediatori quotidiani nell'accesso alle informazioni, interpretando comandi vocali e abitudini d'uso.

L'intelligenza artificiale diventa così una presenza distribuita e costante, integrata nelle interfacce che scandiscono ogni gesto della nostra esperienza digitale.
Come sottolinea Luciano Floridi (2014), siamo entrati nella *quarta rivoluzione*, quella in cui la distinzione tra agenti umani e artificiali si dissolve: viviamo ormai dentro l'infosfera, un ambiente cognitivo ibrido in cui dati, algoritmi e persone coesistono e si influenzano reciprocamente.

In questo scenario, l'interfaccia cambia radicalmente natura.
Non è più uno spazio neutro di rappresentazione, ma un filtro cognitivo e performativo: seleziona, interpreta e genera informazioni. Un'applicazione di mappe, ad esempio, non si limita a visualizzare percorsi, ma li calcola in tempo reale sulla base dei flussi di traffico e del comportamento collettivo degli utenti; una piattaforma musicale come Spotify apprende gusti e routine, suggerendo brani che anticipano stati d'animo; un social network stabilisce gerarchie di visibilità, definendo ciò che emerge e ciò che resta invisibile. L'interazione diventa dunque predittiva e adattiva: ciò che l'utente percepisce non è il mondo com'è, ma il risultato di un calcolo probabilistico invisibile, un modello costruito sul passato per orientare il futuro. L'interfaccia, in questa nuova fase, non si limita a mediare la realtà — la produce, configurando esperienze, preferenze e decisioni all'interno di un ambiente algoritmico che apprende e agisce insieme agli esseri umani.

5. L'AI generativa e la nascita della co-intelligenza (2020–oggi)

L’avvento dei modelli generativi, come GPT di OpenAI, DALL·E, Midjourney o Claude, segna un vero punto di svolta nella storia dell’intelligenza artificiale.
Per la prima volta, le macchine non si limitano a riconoscere schemi o prevedere comportamenti, ma producono attivamente contenuti nuovi: testi, immagini, suoni, video, codice. È il passaggio da un’intelligenza interpretativa a un’intelligenza creativa e dialogica, in cui la macchina partecipa alla costruzione del significato. Questo salto non nasce all’improvviso, ma è il risultato di decenni di evoluzione tecnica e concettuale.
Negli anni Duemila, la disponibilità crescente di big data e di potenza di calcolo ha reso praticabile ciò che nei decenni precedenti era rimasto solo teorico: l’addestramento di reti neurali su scala mondiale. Il passo successivo avviene intorno al 2012, con la rivoluzione del deep learning. Il successo di *AlexNet* (Krizhevsky, Sutskever & Hinton, 2012), rete neurale convoluzionale (CNN) profonda capace di riconoscere immagini con un’accuratezza mai vista prima, mostra che l’intelligenza artificiale può davvero imparare da sola analizzando grandi quantità di esempi. Questo approccio ha permesso di vedere la complessità non più come un limite, ma piuttosto una risorsa in cui la quantità di dati e la profondità del modello diventavano fattori di apprendimento e precisione. Da qui nasce la fiducia che avrebbe portato, pochi anni dopo, allo sviluppo dei modelli linguistici di larga scala.

Da lì prende forma una nuova generazione di modelli basati su architetture neurali profonde. Per molti anni, le reti neurali ricorrenti (RNN) sono state il principale modello usato per elaborare informazioni che si presentano in sequenza, come il linguaggio o i suoni. A differenza delle reti “classiche”, che analizzano ogni dato separatamente, le RNN mantengono una memoria temporanea: ogni nuovo elemento elaborato tiene conto di quelli precedenti. In questo modo, possono gestire frasi, testi o segnali audio, dove il significato di una parte dipende da ciò che viene prima. Tuttavia, questa architettura aveva due grandi limiti. Da un lato, faticava a ricordare informazioni molto lontane nella sequenza (per esempio, la relazione tra l’inizio e la fine di una frase lunga); dall’altro, l’elaborazione era sequenziale e lenta, perché le parole dovevano essere processate una dopo l’altra, impedendo l’uso efficiente del calcolo in parallelo.

Nel 2017, il modello Transformer (Vaswani et al.) superò questi ostacoli introducendo un nuovo principio: l’attenzione (attention mechanism). Questa nuova architettura supera i limiti delle reti ricorrenti introducendo il principio dell’attenzione (*attention mechanism*), che consente al modello di analizzare tutte le parole di una frase in parallelo, valutando automaticamente quali sono più importanti per il significato complessivo.
Invece di seguire la linearità del linguaggio, il Transformer osserva le relazioni interne tra le parole, cogliendo il contesto globale del discorso.
Il risultato è una capacità di comprensione e generazione del testo molto più rapida, coerente e flessibile, che ha reso possibile la nascita dei moderni Large Language Models (LLM) come GPT, BERT o Claude. In passato ci furono tentativi di creare sistemi in grado di dialogare o scrivere autonomamente, come *SHRDLU* di Terry Winograd (1970), ma si trattava di programmi basati su regole fisse e vocabolari limitati.
Funzionavano solo in contesti chiusi, incapaci di adattarsi a situazioni nuove.
I modelli di oggi, invece, apprendono correlazioni linguistiche su scala massiva, costruendo una forma di competenza emergente che consente loro di rispondere a una gamma quasi illimitata di richieste.
La loro forza non risiede in regole predefinite, ma nella quantità di esperienza accumulata durante l’addestramento.

Con questi nuovi sistemi, anche le interfacce cambiano completamente volto.
Non sono più ambienti in cui si premono pulsanti o si digitano comandi, ma spazi di dialogo, dove l'utente parla con la macchina in linguaggio naturale.
Il prompt, la frase con cui si avvia la conversazione, diventa il nuovo strumento di progettazione dell'interazione: un modo per guidare la macchina, ma anche per collaborare con essa.
Ogni risposta generata è unica, frutto di un processo di scambio e costruzione reciproca.
Come spiega Ethan Mollick (2023), siamo entrati nell'era della *co-intelligenza*: un periodo in cui umani e AI lavorano insieme, unendo capacità diverse.
L'intelligenza umana porta creatività, giudizio e contesto; quella artificiale aggiunge velocità, calcolo e memoria.
L'interfaccia diventa così il punto d'incontro tra le due, un luogo di collaborazione cognitiva dove si progettano idee, contenuti e soluzioni.

Per il design, questa trasformazione ridefinisce l'idea stessa di interfaccia.
Non si tratta più di disegnare superfici, ma di progettare relazioni cognitive tra intelligenze diverse. L'interfaccia diventa un luogo di negoziazione e co-creazione, dove l'utente non interagisce con uno strumento, ma dialoga con un agente.
Ogni interazione genera nuova conoscenza, e ogni progetto diventa un esperimento di equilibrio tra trasparenza, fiducia e autonomia.

In questa nuova fase, l'AI non è più un'estensione della mente umana, ma un interlocutore cognitivo con cui condividere il processo creativo.
Progettare interfacce per questi sistemi significa dunque affrontare un compito radicale: rendere visibile e comprensibile la collaborazione con l'intelligenza non umana, preservando al tempo stesso il senso umano dell'esperienza.

2. L'INTELLIGENZA ARTIFICIALE COME AGENTE DI CAMBIAMENTO

Obiettivo: Comprendere come l'intelligenza artificiale stia ridefinendo il concetto di interfaccia, passando da strumento passivo a partner cognitivo attivo. Esplorare le implicazioni progettuali, cognitive ed etiche di questa trasformazione.

2.1.1 Dalla macchina reattiva all'agente cognitivo

Nel capitolo precedente abbiamo osservato come le interfacce si siano evolute attraverso quattro grandi transizioni: dal comando al controllo visivo, dall'interazione esplicita a quella esperienziale, dall'input–output lineare al feedback adattivo, fino all'emergere della dimensione conversazionale. Ogni passaggio ha rappresentato una profonda mutazione nelle dinamiche cognitive tra umani e macchine. Oggi, con la crescente affermazione dell'intelligenza artificiale come componente centrale dei sistemi digitali, siamo di fronte a un momento storico cruciale: il passaggio da macchine che eseguono istruzioni a macchine che interpretano intenzioni, e infine a sistemi capaci di co-creare significati insieme agli utenti.

Questa transizione non è una semplice evoluzione continua, ma una rottura paradigmatica. Fino a pochi anni fa, il rapporto tra umano e macchina era fondato su una premessa fondamentale: il computer non comprendeva realmente, ma simulava il comportamento di chi comprendeva. L'utente doveva adattarsi alla logica del sistema, apprendendo regole,

sintassi e affordance. Se il sistema si comportava in modo errato, la responsabilità ricadeva sull'utente che non aveva capito come usarlo correttamente (Norman, 2013). Oggi, con i moderni sistemi di AI, questa asimmetria si dissolve. La macchina non simula più una comprensione, ma, basandosi sulla lettura di pattern spesso impossibili da comprendere per noi, interpreta. Questo permette ai chatbot conversazionali di leggere il tono di una domanda, coglierne addirittura il contesto implicito e rispondere non semplicemente alla lettera, ma all'intenzione dell'utente. L'AI diventa così un agente cognitivo in grado di negoziare il significato insieme all'utente, piuttosto che limitarsi ad applicare regole rigide.

Come ha osservato il filosofo Luciano Floridi (2014), siamo entrati nella cosiddetta "quarta rivoluzione": quella in cui la distinzione ontologica tra agenti umani e artificiali comincia a sfumarsi. Questo non significa che l'AI possieda consapevolezza o intenzionalità nel senso filosofico classico, ma che il suo comportamento è diventato sufficientemente complesso e adattivo da rendere problematica la vecchia distinzione tra strumento e agente. L'AI contemporanea non è più una protesi della mente umana, come lo erano i calcolatori del Novecento, ma un interlocutore; un partner i cui processi decisionali rimangono largamente opachi all'osservatore umano, ma i cui risultati sono tangibili e influenti.

Sul piano della storia delle interfacce, questo spostamento rappresenta il completamento di una traiettoria iniziata decenni fa. Ricordiamo brevemente: negli anni Cinquanta e Sessanta, l'interfaccia era il luogo del comando. Negli anni Settanta e Ottanta, con l'avvento della grafica e della manipolazione diretta (Shneiderman, 1983), l'interfaccia divenne uno spazio di rappresentazione. Negli anni Novanta e Duemila, con la diffusione del web, l'interfaccia cominciò a essere uno spazio di esperienza. Oggi, come vedremo in dettaglio nei casi studio del capitolo 3, l'interfaccia si trasforma in uno spazio di dialogo e co-intelligenza: un luogo dove due agenti cognitivi, uno umano e uno artificiale, negoziano significati, collaborano nella risoluzione di problemi e apprendono reciprocamente.

Ethan Mollick (2024) utilizza un termine potente per descrivere questo nuovo scenario: *co-intelligenza aliena*. L'idea è che stiamo per la prima volta nella storia umana convivendo con una forma di intelligenza genuinamente diversa dalla nostra, non umanoide né antropomorfa, ma completamente altra. Secondo Mollick, questo incontro con l'alterità cognitiva rappresenta un momento di rinascita filosofica e pratica. L'AI ci obbliga a riconsiderare cosa significhi veramente pensare, creare, decidere. E al contempo, ci pone di fronte a una sfida progettuale radicale: come possiamo sviluppare interfacce che mediano questo incontro in modo consapevole, etico e consenziente?

La risposta a questa domanda non è puramente tecnica. Come abbiamo visto la storia dell'AI è già stata dominata dall'idea che il comportamento intelligente potesse essere ridotto a procedure formali, ma la vera sfida che oggi affrontano i designer è di ordine diverso: come rendere visibile l'invisibile algoritmo? Come preservare l'agency umana in un contesto in cui la macchina anticipa le nostre scelte? Queste domande toccano aspetti etici, culturali e cognitivi della relazione uomo–macchina che saranno centrali nell'analisi dei prossimi paragrafi.

2.1.2 Anatomia dell'intelligenza artificiale nelle interfacce contemporanee

Per capire come l'AI stia ridefinendo le interfacce, è necessario prima sciogliere cosa intendiamo effettivamente per "intelligenza artificiale" nel contesto dell'interaction design. Il termine è stato usato in maniera così ampia da rischiare di diventare un semplice sinonimo di "innovazione digitale". In realtà, le tecnologie di AI rilevanti per il design delle interfacce sono specifiche e distinguibili.

Tipologie di AI rilevanti per l'interaction design
Iniziamo con il *machine learning predittivo*, la tecnologia più diffusa e invisibile nella quotidianità. Sistemi come quelli implementati da Netflix, Spotify o TikTok utilizzano modelli statistici per riconoscere pattern nei dati storici degli utenti e prevedere quali contenuti probabilmente apprezzeranno in futuro. Questi modelli non "comprendono" il gusto: calcolano correlazioni statistiche. Dal punto di vista dell'interfaccia, il machine learning predittivo è caratterizzato dall'invisibilità: l'utente non vede gli algoritmi, sperimenta solo il risultato nel flusso personalizzato. Come analizzato da Shoshana Zuboff (2019), questo ha profonde implicazioni per la trasparenza, poiché l'utente non solo come anche gli stessi creatori di questi modelli non può capire quali pattern di riconoscimento influenzino questi algoritmi, ma in alcun modo può regolare consapevolmente questo processo che non può osservare.

Una classe di tecnologie completamente diversa è il *Natural Language Processing* (NLP) che comprende i *Large Language Models* (LLM). A differenza del machine learning predittivo, che lavora su dati strutturati, l'NLP lavora sul linguaggio naturale. Un LLM non consulta un database di risposte preesistenti, ma genera una risposta unica, token dopo token, calcolando statisticamente la continuazione più probabile. Dal punto di vista dell'interfaccia, l'LLM introduce un cambio paradigmatico: il linguaggio naturale diventa il nuovo strumento di controllo. Non serve più imparare la sintassi di un'interfaccia grafica; basta esprimersi naturalmente nella propria lingua.

Parallela agli LLM è la *computer vision*, dove il sistema interpreta immagini anziché testo. Sistemi come Google Lens o le tecnologie di riconoscimento facciale introducono una nuova dimensione percettiva: il sistema può "vedere" ciò che accade nello spazio fisico. L'interfaccia diventa così ambientale: non è più uno schermo davanti all'utente, ma una presenza distribuita nello spazio che osserva e reagisce.

Infine, la *generazione di contenuti visivi* (DALL·E, Midjourney) introduce la possibilità di contenuti dinamici. Ogni utente può ricevere visualizzazioni uniche, generate al momento. Questo sposta il design dall'essere una disciplina della creazione statica ad una della selezione e della definizione di vincoli per sistemi generativi.

Caratteristiche distintive dell'AI come componente interfacciale
Al di là della varietà tecnologica, le AI contemporanee condividono alcune caratteristiche fondamentali che le distinguono dai sistemi tradizionali.

La prima è l'**adattività**. I sistemi tradizionali erano statici; l'interfaccia AI si modifica continuamente sulla base dell'interazione. Per il designer, questo introduce una sfida nuova: come si crea coerenza e prevedibilità quando il sistema stesso è in evoluzione?
La seconda caratteristica è la **proattività**. I sistemi diventano capaci di anticipare bisogni e suggerire azioni prima che l'utente le richieda. Questo trasforma il rapporto di potere nell'interfaccia: non è più l'utente che controlla attivamente il sistema, ma il sistema che

inizia a orientare l'attenzione dell'utente.
La terza è l'**opacità**. Come vedremo nella sezione 2.1.8, molti sistemi di AI (in particolare le reti neurali profonde) funzionano come "black box": i loro meccanismi decisionali non sono completamente ispezionabili nemmeno dagli sviluppatori.
La quarta è il **probabilismo**. A differenza del software deterministico, un'AI generativa produce risultati probabilistici e variabili. Per il designer, significa rinunciare al controllo totale e progettare per l'incertezza. Queste caratteristiche rappresenteranno le fondamenta dei nuovi paradigmi di interazione che esploreremo ora.

2.1.3 Nuovi paradigmi di interazione nell'era dell'AI

Le trasformazioni tecniche descritte non rimangono nel dominio tecnologico, ma è importante notare come emergono concretamente nuovi modelli comportamentali. In continuità con quanto discusso nel Capitolo 1 a proposito dell'evoluzione storica, vediamo ora come l'AI stia riscrivendo la grammatica dell'interazione.

Dall'interazione esplicita all'interazione implicita
Storicamente, l'interazione uomo–macchina è stata esplicita: l'utente agisce, il sistema risponde, in questo modo il controllo è affidato all'utente. Con l'AI pervasiva, l'interazione diventa implicita: il sistema compie azioni sulla base di osservazioni passive, senza essere esplicitamente interrogato. Una smart home che regola le luci o una piattaforma che riordina il feed agiscono in modo *anticipatorio*. Come notava Mark Weiser (1991) parlando di *ubiquitous computing*, la tecnologia più profonda è quella che scompare. Tuttavia, questo passaggio pone una sfida etica cruciale: come creiamo spazi dove l'utente mantiene agency anche quando l'interazione è invisibile?

Dalla manipolazione diretta alla conversazione cognitiva
Per decenni, il paradigma dominante è stato quello della manipolazione diretta (Shneiderman, 1983): trascinare finestre e premere pulsanti per un controllo immediato. Con gli assistenti conversazionali, emerge il paradigma della *conversazione cognitiva*. Invece di agire su oggetti visivi, l'utente formula intenzioni linguistiche complesse. Questo segna un ritorno a forme di comunicazione umane, ma introduce una nuova ambiguità: il linguaggio infatti non si può definire un medium trasparente, ma un campo di negoziazione costante tra intenzione e interpretazione. Basti pensare come anche tra gli umani spesso sia esattamente così: le parole sono pronunciate con un significato per l'umano che le formula, che non sempre però riesce ad arrivare all'ascoltatore con lo stesso impatto con cui sono partite. La negoziazione del significato delle parole oltre alla loro mera utilità sicuramente ha portato nella nostra evoluzione a continui cambiamenti nella profondità del nostro linguaggio.

Dall'output statico alla generazione dinamica
L'output del sistema non è più un artefatto fisso, ma un processo generativo. Un'interfaccia può essere creata in tempo reale per adattarsi allo specifico utente. Questo trasferisce il lavoro del designer dalla definizione della forma finale alla progettazione delle regole e dei vincoli che guidano la generazione.

Dall'interfaccia visibile all'interfaccia ambientale
Infine, assistiamo alla dissoluzione del dispositivo. L'interfaccia non è più localizzata in uno schermo, ma diffusa nell'ambiente attraverso sensori e assistenti vocali. L'invisibilità riduce

la frizione, ma rischia di eliminare anche la consapevolezza dell'interazione, un tema che Sherry Turkle (2011) ha evidenziato come critico per la nostra capacità di mantenere relazioni autentiche e non solo connessioni automatizzate.

2.1.4 Bias mitigation

Nel contesto delle interfacce guidate da AI, il bias non è un "incidente" che si manifesta solo in casi estremi, ma un prerequisito strutturale con cui le aziende devono fare i conti fin dall'inizio. I dati su cui i modelli vengono addestrati riflettono inevitabilmente le asimmetrie storiche e sociali del mondo e di coloro che si sono occupati di fornire quei dati. In assenza di strategie di mitigazione queste asimmetrie vengono tradotte in decisioni operative reali che colpiscono gli utenti in modo sistematico. In questa realtà, la gestione del bias diventa un criterio di qualità dell'interazione al pari della comprensibilità o dell'affidabilità.

Le grandi aziende tecnologiche hanno iniziato a formalizzare questo tema soprattutto attraverso il linguaggio della fairness. Microsoft, ad esempio, inserisce la "fairness" tra i sei principi del Responsible AI, accanto a trasparenza, affidabilità e accountability, e richiede nei suoi standard interni analisi sistematiche degli impatti differenziati sui gruppi sociali coinvolti. Una delle metriche più comuni è la demographic parity, che verifica se la percentuale di esiti positivi di un modello (per esempio l'approvazione di un prestito o l'accesso a una risorsa) è comparabile tra gruppi diversi, come donne e uomini o persone di etnie differenti.
Se la Demographic Parity guarda se le percentuali di esiti positivi sono simili tra gruppi diversi; il Disparate Impact guarda invece quanto un gruppo sta peggio rispetto al gruppo più favorito e decide se la differenza è talmente grande da poter essere considerata discriminazione indiretta. Il Disparate Impact, misura quindi il rapporto tra questi tassi di esito positivo e considera problematiche le situazioni in cui un gruppo svantaggiato ottiene risultati molto peggiori rispetto al gruppo più favorito, superando le soglie comunemente usate nelle norme antidiscriminazione.
Per il design dell'interfaccia, queste metriche non dovrebbero restare confinate alla documentazione tecnica, ma diventare veri e propri materiali di progettazione continua. Significa assicurarsi di incorporare i test di fairness durante tutto il processo di sviluppo dell'AI, a partire dalla raccolta dei dati di allenamento e ripeterli continuamente anche dopo il rilascio. Questi test includono, ad esempio, simulazioni di scenari d'uso con utenti fittizi appartenenti a gruppi diversi e osservando il cambiamento di suggerimenti, raccomandazioni o blocchi di contenuto. Inoltre si assiste all'occasione di progettare punti di contatto in cui l'utente può vedere, almeno a livello sintetico, come il sistema tende a distribuire opportunità e rischi, e dove può segnalare esiti percepiti come ingiusti che alimentano cicli di revisione del modello. Sicuramente il problema dei bias e la sua intrinsecità rappresenteranno una sfida significativa nel futuro dell'IA e sulla sua capacità di essere una tecnologia in grado di fare del bene al mondo. Rimane alla politica mondiale la capacità di regolamentare e legiferare a riguardo anche se fin'ora solo l'UE attraverso l'AI Act ha formulato convenzioni vincolanti per gli sviluppatori. Nonostante questi sforzi, rimane un problema tecnico-politico irrisolto: la definizione di "equità". Un algoritmo può essere "equo" matematicamente (stesso tasso di errore per tutti i gruppi) ma rimanere ingiusto socialmente.

2.1.5 Il gap tra accountability tecnica e responsabilità etica

Negli ultimi anni, complici le forti pressioni dell'opinione pubblica, della ricerca scientifica e della politica, le aziende che sviluppano sistemi di AI hanno costruito apparati sempre più sofisticati di governance e compliance. Microsoft parla di comitati interni per il Responsible AI, linee guida obbligatorie, audit periodici e processi di valutazione del rischio prima del rilascio di nuovi sistemi. OpenAI si impegna nello sviluppo di strumenti di moderazione automatica, policy dettagliate per categorie sensibili e procedure di enforcement che includono blocchi di account, limitazioni d'uso e filtri preventivi sulle risposte dei modelli. Anthropic, con l'approccio della Constitutional AI, arriva ad introdurre una "costituzione" di principi che guida il comportamento del modello, abbinando a questo un quadro di analisi dei possibili danni fisici, psicologici, economici e sociali attraverso un Unified Harm Framework che orienta poi decisioni pratiche su policy e gestione di particolari casi d'uso. Ma abbiamo anche posture diverse, incarnate principalmente dalla strategia di xAI e dal suo modello Grok. Sotto la guida di Elon Musk, l'azienda ha deliberatamente scelto una via opposta: quella della "massima ricerca della verità" (Maximum Truth-seeking AI) e della resistenza a quella che definisce "censura woke". Invece di investire in complessi strati di filtri preventivi, Grok è stato lanciato con "guardrails" (barriere di sicurezza) minimi, promuovendo una visione assolutista della libertà di parola che delega quasi interamente la responsabilità etica all'utente finale.

Questa filosofia "anti-establishment" ha mostrato il suo lato oscuro proprio nei primi giorni del 2026 sul social X. L'introduzione di nuove funzionalità di generazione e modifica delle immagini in Grok, prive di adeguati controlli di sicurezza, ha permesso agli utenti di creare e diffondere viralmente deepfake non consensuali a sfondo sessuale ("nudify"), prendendo di mira non solo celebrità, ma anche comuni cittadini e persino minori. Mentre i concorrenti bloccavano alla fonte tali richieste, l'algoritmo di Grok le assecondava, costringendo infine l'Unione Europea e le autorità di sorveglianza britanniche ad intervenire con ordini d'urgenza per il congelamento dei dati. Il caso Grok è diventato così l'esempio più lampante del gap attuale. Questo evento in particolare dimostra che un'eccellente capacità tecnica, se non accompagnata da un'architettura di responsabilità etica "by design", può trasformare uno strumento di innovazione in un vettore di abuso incontrollato.

Le pratiche citate in precedenza definiscono quindi una responsabilità tecnica, centrata su audit, metriche, procedure di controllo e tracciamento delle decisioni. Sicuramente si tratta di passi importanti perché rende in parte l'AI oggetto di regole verificabili, non solo di dichiarazioni valoriali astratte. Dal punto di vista del design però, questo non basta ancora a costituire una vera responsabilità etica: l'utente resta spesso ai margini di questi processi, informato solo da policy testuali o note legali difficilmente leggibili, senza reali strumenti per contestare o negoziare le scelte algoritmiche che lo riguardano. Il rischio è che l'etica venga tradotta in una checklist burocratica, dove l'obiettivo è "superare i test" anziché aprire uno spazio di confronto critico.

Carl DiSalvo, con il concetto di adversarial design, propone una postura diversa: non basta rendere il sistema comprensibile, occorre progettare artefatti che facciano emergere i conflitti politici e i rapporti di potere incorporati nelle tecnologie. Se riportiamo questa idea alle interfacce AI, la domanda diventa: come può l'interfaccia non solo spiegare "come funziona" il sistema, ma anche rendere visibile "chi ci guadagna, chi rischia, chi decide i parametri di rischio accettabili"? In altre parole, come può l'interfaccia diventare un luogo in cui le scelte

di governance non sono solo applicate dall'alto, ma discusse e, in certi casi, contestate dagli utenti?

Un terreno in cui questo gap emerge chiaramente è quello delle politiche sui contenuti. Le grandi aziende definiscono insiemi di contenuti vietati o limitati, ma lo fanno con accenti diversi. Microsoft enfatizza il blocco di contenuti violenti, d'odio o sessualmente espliciti nei servizi consumer, con una forte integrazione tra policy e strumenti di filtraggio automatico per proteggere utenti generalisti, tra cui minori. OpenAI articola categorie molto dettagliate di contenuti proibiti o da attenuare (hate, harassment, self-harm, sexual, violence) e fornisce API di moderazione che permettono agli sviluppatori di applicare questi filtri in modo personalizzato nelle proprie applicazioni, creando uno spazio più ampio per contesti professionali o creativi, ma sempre entro confini normati. Anthropic utilizza invece la costituzione del modello non solo per vietare certe risposte, ma per far sì che il sistema stesso "si critichi" e riformuli output problematici secondo principi etici espliciti, rendendo il processo di correzione parte integrante del comportamento del modello.

Queste differenze riflettono tre visioni distinte di "qualità" e "responsabilità": una più orientata alla protezione preventiva, una alla gestione granulare dei rischi attraverso strumenti di moderazione, e una alla modellazione interna di norme e autocritica nel sistema. Dal punto di vista del design, la sfida è non limitarsi a scegliere uno di questi framework come "pacchetto etico" preconfezionato, ma capire come tradurli in interfacce che rendano visibili i compromessi e gli inevitabili margini di arbitrarietà. Ad esempio, non basta dichiarare "trasparenza" pubblicando documenti tecnici; occorre progettare schermate, messaggi e flussi che mostrino perché un contenuto è stato bloccato, quali alternative sono possibili, e come l'utente può partecipare alla revisione di queste scelte. In questa direzione, la qualità non è la perfetta conformità a una policy, ma la capacità di un sistema di esporre i propri conflitti, limiti e assunzioni di fondo.

2.1.6 Concretizzare l'etica: dalla policy al design

Se i principi aziendali sull'AI responsabile definiscono un orizzonte di valori, il compito del design è tradurli in decisioni concrete visibili all'utente, permettendogli inoltre di intervenire scegliendo autonomamente cosa mostrare, cosa nascondere. Dal punto di vista dell'utente, la fairness smette di essere uno slogan astratto quando diventa visibile nel modo in cui il sistema tratta persone diverse: quando mostra chiaramente quali dati usa, come tutela i gruppi più vulnerabili, e quando prevede limiti oltre i quali il modello viene corretto o fermato per evitare effetti ingiusti. Un progetto "etico" non è quello che dichiara di essere giusto, ma quello che, oltre agli strumenti di valutazione, alla documentazione pubblica delle scelte fatte e ai meccanismi di revisione, integra gli utenti stessi in questo processo, favorendo le loro possibilità di segnalare problemi, contestare decisioni e contribuire in modo informato alla correzione continua del sistema.

Un altro snodo cruciale riguarda la comunicazione dei limiti dell'AI. Le policy aziendali insistono spesso su formule come "il modello può commettere errori" o "non usare il sistema per decisioni ad alto rischio", ma queste frasi restano inefficaci se relegate a note legali. Il design può rendere operativi questi limiti attraverso pattern espliciti, cioè soluzioni ricorrenti di interfaccia e microtesto che traducono l'astrazione in segnali e scelte concrete per

l'utente: indicazioni di incertezza integrate nelle risposte, indicatori visivi del livello di certezza, avvisi contestuali quando l'utente si avvicina a scenari critici (per esempio, ambiti medici, legali o finanziari), e percorsi chiari per chiedere una conferma umana o una verifica alternativa prima di agire. In questo modo, la trasparenza non è soltanto il "diritto di sapere", ma una pratica continua di orientamento e supporto all'uso consapevole.

Infine, la progettazione per il fallimento sicuro e per il "human in the loop" richiede un equilibrio delicato. Molte linee guida raccomandano la supervisione umana, ma se ogni decisione viene rimandata a un operatore, il sistema diventa inutilizzabile e l'umano viene trasformato in un semplice esecutore di conferme, schiacciato da carichi di lavoro impossibili. Un design responsabile dovrebbe individuare i nodi in cui il fallimento del sistema può avere impatti sproporzionati, e concentrare lì i meccanismi di intervento umano. Ad esempio, in un sistema di raccomandazione di contenuti, il fallimento è spesso reversibile; in un sistema di triage sanitario o di valutazione creditizia, invece, gli errori possono aumentare le disuguaglianze. Progettare per il fallimento sicuro significa allora prevedere percorsi di revisione, appelli, seconde opinioni e sospensioni automatiche del modello in presenza di situazioni anomale, restituendo all'umano non il ruolo di controllore passivo, ma di co-decisore informato.

2.1.7 Ripensare la "buona interazione": nuovi criteri di qualità

Se i paradigmi dell'interazione cambiano, devono cambiare anche i criteri con cui giudichiamo la qualità di un'interfaccia. L'usabilità tradizionale, focalizzata su efficienza e task completion, diventa quindi insufficiente per valutare sistemi conversazionali o predittivi. Le maggiori aziende che controllano il mercato dell'AI hanno dovuto affrontare questo problema rispondendo ognuna con modalità diverse:

Microsoft: Fonda il suo schema di AI Responsabile su sei linee guida:

- Fairness (equità)
- Reliability and Safety (affidabilità e sicurezza)
- Privacy and Security (privacy e protezione dei dati)
- Inclusiveness (inclusività)
- Transparency (trasparenza)
- Accountability (responsabilità)

Microsoft ha creato un ufficio dedicato (Office of Responsible AI) e sviluppa strumenti tecnici come moderation layer pre-addestrati e filtri dinamici di contenuto.

Google: Articola il suo approccio su tre principi fondativi:

- Bold Innovation: sviluppare AI che assista e metta al centro le persone, contribuendo al progresso scientifico

- Responsible Development and Deployment: implementare oversight umano appropriato, dovuta diligenza, rigorous testing; mitigare bias non intenzionali; promuovere privacy
- Collaborative Progress: creare un ecosistema che aiuti altri a innovare responsabilmente

Google dal 2019 afferma di aver formato oltre 32.000 dipendenti sugli AI Principles e ha più di 200 professionisti full-time dedicati a rendere operative pratiche responsabili. Utilizza Adversarial Proactive Fairness (ProFair) testing e AI Principles reviews prima del lancio dei prodotti.

OpenAI: Adotta un approccio centrato su policy di uso e safety training:

- Moderation API che blocca automaticamente prompt che violano politiche di contenuto
- Reinforcement Learning from Human Feedback (RLHF) durante l'addestramento
- Graduated access tiers per nuovi utenti
- Red Teaming networks per audit indipendenti
- Model cards che documentano limitazioni note e rischi di bias

Anthropic: Ha sviluppato il framework “Constitutional AI”, che allinea il comportamento dell'AI a una "costituzione" di principi umani tramite:

- Autotraining supervisionato
- Adversarial training
- Self-critique integrated nel processo di addestramento
- Focus su helpfulness, harmlessness, honesty

Le principali aziende che sviluppano AI ridefiniscono oggi la qualità non più in termini di task completion, bensì di criteri etici, dal loro lavoro si possono estrapolare cinque dimensioni convergenti:

- Metriche di equità (Disparate Impact Analysis, demographic parity) che misurano se il sistema discrimina tra gruppi;
- Trasparenza e Explainable AI (XAI) per esporre il ragionamento del modello;
- Responsabilità distribuita nel ciclo di vita dell'AI, riconoscendo che il bias emerge dalle scelte umane;
- Sicurezza con human oversight soprattutto in applicazioni ad alto rischio;
- Privacy e minimizzazione dei dati.

Ciò che le accomuna è il cambiamento radicale che producono: la qualità non è più legata unicamente alla funzione, ma a come l’AI tratta l’utente. Misurare equità significa ammettere che il sistema può discriminare. Documentare la trasparenza significa rifiutare la finzione della neutralità tecnica. In questa prospettiva, progettare interfacce con l'AI è un atto politico, dove ogni scelta riflette una visione di chi ha diritto a cosa. Tuttavia, riconoscere questa

dimensione politica della qualità non basta se rimane astratta. Il design deve tradurla in scelte concrete: come si fa a progettare un'interfaccia che ammetta il proprio bias? Come si preserva l'agency umana quando il sistema agisce in modo proattivo e autonomo? Come si comunica l'incertezza probabilistica senza paralizzare l'utente? Per rispondere, è necessario articolare una serie di criteri progettuali che mettono in scena questa visione etica, trasformandola da dichiarazione di principio a prassi effettiva. Questi criteri emergono dalla ricerca di un equilibrio: tra trasparenza e usabilità, tra controllo umano e automazione intelligente, tra fiducia nel sistema e consapevolezza dei suoi limiti.

I criteri emergenti includono:

- **Comprensibilità**: Un'interfaccia AI deve potersi rendere trasparente, spiegando perché il sistema agisce in un certo modo. Se il sistema non può articolare i suoi processi interni, deve almeno comunicare apertamente i suoi limiti e la natura probabilistica delle sue risposte. Questo significa evitare il cosiddetto "Eliza effect", cioè la tendenza degli utenti a attribuire intelligenza e intenzionalità a sistemi che simulano solo comportamenti intelligenti.
- **Controllabilità**: preservare l'agency umana significa offrire punti di intervento concretamente accessibili. L'utente deve poter correggere, rifiutare o modificare l'azione del sistema, specialmente quando questo agisce in modo proattivo o suggerisce decisioni critiche.
- **Affidabilità**: Non significa assenza di errori, ma la capacità di fallire in modo sicuro (graceful failure), senza produrre danni sproporzionati e mantenendo la fiducia dell'utente. Un sistema affidabile comunica anche quando sa di non sapere.
- **Appropriatezza**: È la capacità del sistema di trovare un equilibrio tra automazione e intervento umano in base al contesto. Un design appropriato sa quando farsi da parte e quando intervenire, evitando sia l'eccessiva invadenza sia l'abbandono dell'utente.
- **Gestione proattiva del bias:** Infine, è necessario implementare strategie che riconoscono il bias come inevitabile e strutturale. Questo include: inclusione di gruppi diversi nella ricerca e nel testing; verifica regolare degli algoritmi per individuare discriminazioni; feedback mechanisms che permettono agli utenti di segnalare esiti ingiusti; "bias impact statement" compilati prima dello sviluppo, come pratica di autoriflessione critica. Il bias non deve essere trattato come qualcosa da nascondere, ma più come qualcosa da gestire trasparentemente, anche con la collaborazione dell'utente.

2.1.8 Modelli mentali e comprensibilità: progettare per l'opacità

Uno dei problemi più ampi del design di interfacce AI riguarda la gestione della "black box". Come progettiamo interfacce per sistemi che non comprendiamo completamente nemmeno noi?

Il problema della “black box” e le conseguenze cognitive
Le reti neurali profonde operano attraverso strati di astrazione matematica che non sono traducibili in semplici regole *if–then*. Questa opacità è intrinseca. La conseguenza psicologica è che l’utente oscilla spesso tra due estremi: la *fiducia cieca*, trattando l’AI come un oracolo infallibile, o il *rifiuto totale*, scartando il sistema per mancanza di trasparenza. Sherry Turkle già nel 2011 ha ben descritto come l’illusione di comprensione reciproca possa portare a legami emotivi con macchine che in realtà non provano nulla.

Strategie progettuali per la trasparenza
Per mitigare questi rischi, il design deve adottare strategie di *Explainable AI* (XAI), non necessariamente mostrando il codice, ma fornendo spiegazioni a livello cognitivo (es. “Ti suggerisco questo perché hai guardato X”). È fondamentale ridimensionare le aspettative: l’interfaccia deve comunicare chiaramente ciò che il sistema *non* sa fare, aiutando l’utente a costruire un modello mentale accurato e a non antropomorfizzare eccessivamente la tecnologia.

2.1.9 Dimensione etica e responsabilità progettuale

L'introduzione dell'AI nelle interfacce non è mai neutrale. Ogni algoritmo incorpora visioni del mondo, priorità e potenziali bias. Progettare un'interfaccia con l'AI significa quindi sempre, inevitabilmente, prendere posizione su come mediare le relazioni di potere tra umani, dati e macchine. L’AI assume quindi il ruolo di vero mediatore attivo, portando con sè tutta una serie di considerazioni etiche. Come sottolinea Kate Crawford (2021), i sistemi di AI sono mappe del potere: decidono cosa rendere visibile e cosa oscurare. Quando un'interfaccia automatizza una scelta, sta esercitando una forma di soft power. Un algoritmo di raccomandazione perciò non si occupa semplicemente di suggerire, ma piuttosto modella preferenze, orienta percezioni, restringe il campo delle possibilità verso una visione specifica del mondo. In questa prospettiva, l'AI non è uno strumento neutrale, ma un agente che interpreta, filtra e predispone il comportamento umano, arrivando a creare spesso delle camere d’eco che influenzano la vita delle persone, sono documentati casi in cui proprio Facebook ha avuto un ruolo significativo nel fomentare violenze, tensioni e disinformazione in vari conflitti e crisi umanitarie, in particolare in Myanmar.

Si apre quindi uno scenario su chi decide veramente in un mondo così influenzato dagli algoritmi? E come preserviamo l'agency umana in un contesto dove la tecnologia agisce in modo proattivo, talvolta senza attendere il nostro consenso esplicito? Quando un assistente vocale anticipa bisogni, quando un algoritmo medico suggerisce una diagnosi, quando un feed social ordina l'informazione che vediamo, l'utente si trova in una posizione asimmetrica: subisce decisioni che ignora di stare subendo. Questo configura un problema di autonomia profondamente etico: è chiaro che la capacità reale di decidere richiede consapevolezza, non solo il diritto teorico di scegliere. Ne sono un esempio i formati di video brevi a scorrimento infinito, ormai dominanti sui social media: in queste interfacce, progettate per eliminare ogni punto di frizione, la fruizione cessa di essere una sequenza di scelte intenzionali per diventare un flusso passivo. L’autonomia dell’utente viene così erosa da meccanismi che, anticipando i desideri istantanei, sostituiscono la volontà cosciente con l'automatismo comportamentale. Vale la pena considerare maggiormente questo discorso se si tiene conto anche dei Bias algoritmici che possono portare a discriminazione sistemica se non vengono identificati e corretti in fase di progettazione, poichè intrinseci nei modelli

addestrati su dati. Ad esempio un algoritmo di selezione del personale che eredita bias di genere dai dati storici non "sbaglia": riproduce e amplifica disuguaglianze già esistenti nella società, metabolizzandole come fatto tecnico inevitabile. Lo stesso accade con riconoscimento facciale meno accurato su persone con pelle scura, o sistemi di credit scoring che penalizzano comunità specifiche. Il pericolo non è l'errore umano, ma l'errore sistematizzato, invisibile, ripetibile a scala massiva. Questo trasforma il bias da questione tecnica a questione di giustizia.

Privacy e sorveglianza predittiva

Parallela al problema del bias è la questione della privacy. L'AI contemporanea funziona attraverso la raccolta massiccia di dati personali: comportamenti, preferenze, movimenti, stati emotivi vengono continuamente monitorati e analizzati. Shoshana Zuboff (2019) la definisce "capitalismo della sorveglianza": una pratica in cui la nostra vita privata diventa materia prima per costruire profili predittivi e influenzare il nostro comportamento futuro. L'interfaccia diventa così uno strumento di osservazione costante, che ci monitora mentre pensiamo di usarla semplicemente. I modelli generativi, in particolare, mantengono copie delle conversazioni per l'addestramento, creando archivi di intimità cognitiva che espongono la vulnerabilità umana. La privacy non è quindi un problema secondario, ma una questione etica di centrale importanza: il diritto di non essere costantemente profilati, osservati e predetti.

Trasparenza, accountability e fiducia: il diritto di sapere

Come rispondere a questa asimmetria di potere? La prima strada è la trasparenza: il diritto dell'utente di sapere come funziona il sistema che lo riguarda. Non si tratta di rendere il machine learning completamente spiegabile (un problema tecnico ancora aperto), ma di creare una "traccia visibile" del ragionamento: quali dati sono stati usati? Quali scelte di modellazione sono state fatte? Quali sono i rischi noti e i limiti? Comunicare questo richiede design consapevole. Tuttavia, la trasparenza da sola non basta se rimane solo al livello della conoscenza, senza poter eessere attuata. Accountability significa attribuire chiaramente responsabilità: se il sistema commette un errore discriminatorio, chi ne è responsabile? Lo sviluppatore dell'algoritmo? Il designer dell'interfaccia? L'azienda che lo ha rilasciato? L'utente che lo ha usato? La risposta non è semplice, ma la domanda deve essere posta esplicitamente. Senza accountability, la trasparenza diventa un'illusione: sapere come funziona il sistema è inutile se non posso contestare le sue decisioni o chiedere risarcimento.

Responsabilità distribuita: utente, designer, sviluppatore, piattaforma

Emergono così quattro attori che condividono responsabilità. Lo sviluppatore deve cercare bias nei dati e negli algoritmi, testare in modo continuo, documentare limitazioni. Il designer deve costruire interfacce che non nascondono la natura algoritmica del sistema, ma la comunicano. L'utente deve essere messo nelle condizioni di esercitare giudizio critico, non di delegare completamente la decisione. E la piattaforma (l'azienda, l'istituzione) deve formalizzare governance, audit e risarcimento. Nessuno di questi attori può sottrarsi alla responsabilità: è distribuita, condivisa, e richiede cooperazione. Questo modello rifiuta l'idea che gli algoritmi siano "responsabili" nel senso morale: sono strumenti che riflettono scelte umane, e quelle scelte rimangono in capo agli umani che le hanno fatte.

Progettare per il consenso informato in contesti opachi

Qui emerge la contraddizione più acuta: come ottenere consenso informato quando il sistema è fondamentalmente opaco? Un utente non può consentire pienamente a qualcosa che non comprende. Le soluzioni attuali (termini di servizio lunghi, opt-out sepolti nei menu) sono forme di pseudo-consenso che mascherano il non-consenso. Progettare per il consenso vero significa: comunicare in linguaggio chiaro quali dati vengono raccolti e perché; offrire granularità nelle scelte (non "tutto o nulla", ma selezione fine); permettere revoca continua, non una decisione una tantum; costruire alternative che non dipendono da sorveglianza totale. È un lavoro di design che ha implicazioni legali e culturali: richiede di ripensare le architetture tecniche stesse, non solo di aggiungere disclaimer.

Principi per un design responsabile

Da questi elementi emerge una visione del design come pratica etica intrinseca, non come applicazione a posteriori di principi etici. Un design responsabile delle interfacce intelligenti poggia su tre movimenti simultanei. Primo, riconoscere il conflitto: ogni sistema AI contiene conflitti nascosti tra diversi interessi (velocità vs. equità, personalizzazione vs. privacy, automazione vs. controllo umano). Progettare non significa risolverli magicamente, ma farli emergere, permettendo all'utente di percepirli e negoziare posizione. Secondo, distribuire il potere decisionale: anziché concentrare la decisione nell'algoritmo, il design deve preservare punti di intervento umano, creando spazi di contestazione e modifica. Terzo, documentare il viaggio: ogni interfaccia AI dovrebbe poter narrare la propria genesi (quali dati, quali assunzioni, quali bias noti), trasformando la trasparenza, e quindi la comprensione, da diritto teorico in esperienza concreta.

Il ruolo del designer come mediatore etico e culturale: design adversarial

Carl DiSalvo (2012) propone il concetto di adversarial design: progettare interfacce che non si limitino a rendere l'uso fluido, ma che facciano emergere le tensioni, i conflitti e i valori in gioco, permettendo all'utente di interrogare il sistema anziché subirlo passivamente. Non si tratta di progettare "per l'utente", ma di progettare "con l'utente" i conflitti che il sistema contiene. Una interfaccia adversarial comunica quando il sistema non sa cosa fare, mostra i dati che lo addestrano, crea spazi di contestazione, invita a pensare criticamente anziché a delegare. Il designer diventa quindi una guida etica tra mondi infinitamente estesi e complessi: il mondo dell'algoritmo (probabilistico, basato su pattern nei dati) e il mondo umano (intenzionale, narrativo, morale). Non fa scomparire questa complessità, sarebbe impossibile, ma la rende visibile, gestibile, negoziabile. In questa prospettiva, progettare un'interfaccia con l'AI non è solo tecnica, ma un atto politico e culturale: decidere come il futuro dell'interazione umano-macchina deve configurarsi, e su quali valori poggiare quella configurazione.

2.2 Questioni etiche generali

Prima di affrontare l’analisi dello stato dell’arte dell’intelligenza artificiale, una tecnologia ormai in grado di generare al tempo stesso grandi benefici e rilevanti rischi, è necessario

premettere una riflessione. Le sue implicazioni etiche si rivelano già oggi profonde, pur considerando la brevità del tempo intercorso dal suo impatto reale sulla società.

L'intelligenza artificiale viene spesso descritta come una "General Purpose Technology", ovvero una tecnologia di portata generale capace di trasformare in modo sostanziale diversi settori economici e culturali. Anche nell'ipotesi, poco realistica, che il suo sviluppo si arrestasse improvvisamente, gli effetti finora prodotti continuerebbero a esercitare un'influenza duratura e a sollevare interrogativi etici di grande rilievo.

Diventa quindi necessario portare alla luce tali questioni e interrogarsi sul nostro modo di affrontarle. A questo scopo, il punto di partenza è stato il documento Antiqua et Nova. Note on the Relationship Between Artificial Intelligence and Human Intelligence, redatto in ambito vaticano, che individua e categorizza le principali problematiche etiche legate al rapporto tra l'intelligenza artificiale e quella umana. Per ciascuna di queste categorie, l'obiettivo è comprendere su quali fattori si fondino le questioni sollevate e da quali elementi possa dipendere la loro possibile risoluzione.

2.2.1. Armi autonome e deumanizzazione della guerra

La stampa internazionale mostra un panorama inquieto. Secondo *The Guardian*, i nuovi sistemi d'arma basati su IA stanno diventando capaci di prendere decisioni senza input umano, riducendo la guerra a un processo automatizzato che rischia di scivolare fuori controllo. Le inchieste non sono rassicuranti: *La Repubblica* ha documentato come, nei sistemi utilizzati in Israele per selezionare obiettivi a Gaza, le percentuali di falsi positivi rimangono elevate. E *Dossier Difesa* evidenzia che l'integrazione dell'IA introduce nuove superfici d'attacco informatico, rendendo vulnerabili perfino le infrastrutture militari più sofisticate.

Ma questa rassegna solleva una domanda che nessun articolo risolve: **che cosa significa mantenere un "controllo umano significativo" quando le macchine operano più velocemente degli operatori che dovrebbero supervisionarle?** Parlare di governance globale è necessario, ma a oggi appare quasi un voto di fiducia. La competizione geopolitica spinge gli Stati a correre più della loro etica. Il cuore del problema, dunque, non è solo tecnico. È filosofico. Se la responsabilità morale non è assegnabile perché la decisione è presa da un sistema opaco, chi risponde del danno? E quali strumenti giuridici potrebbero garantire che qualcuno sia davvero responsabile? Senza affrontare queste domande, ogni promessa di regolazione resta fragile.

2.2.2. Disuguaglianza algoritmica e discriminazione

Secondo *Luce – La Nazione*, l'IA non è neutrale: i modelli riproducono pregiudizi invisibili presenti nei dati, colpendo in modo sproporzionato minoranze già vulnerabili. Il giornale mostra esempi concreti: sistemi di selezione del personale che sfavoriscono donne e persone di determinate etnie, modelli linguistici che leggono i dialetti come segnali di "inaffidabilità", algoritmi che classificano il rischio criminale basandosi su pattern distorti. Non

serve includere dati sensibili per discriminare: spesso basta la semplice correlazione statistica.

Gli articoli centrano il punto, ma aprono un tema più profondo. Le soluzioni proposte, come audit indipendenti e dataset più trasparenti, hanno senso ma non bastano, perché ignorano la domanda critica: **chi controlla i controllori?** Chi finanzia gli audit? Con quale potere possono forzare le aziende ad accettarne gli esiti?

E soprattutto: **le comunità discriminate partecipano alla progettazione di questi sistemi o restano solo "casi studio"?**
Se l'IA deve ridurre le disuguaglianze, non può essere progettata esclusivamente da chi le ha sempre subite meno. Serve una governance partecipativa, non solo una correzione tecnica.

La stampa segnala il problema, la filosofia ricorda che senza redistribuzione del potere e non solo dei dati, il rischio è riprodurre disuguaglianze con più efficienza e meno scuse.

2.2.3. Privacy, controllo dei dati e autonomia digitale

La rassegna stampa italiana mette in luce una tensione crescente: secondo *La Stampa – Finanza*, il Garante Privacy ritiene che i modelli generativi stiano erodendo il controllo sui dati personali, a partire dall'addestramento tramite scraping massivo: una tecnica di raccolta automatizzata e su larga scala di dati da siti web o fonti online attraverso l'utilizzo di software specializzati. Il sito ufficiale del **Garante Privacy** conferma che principi come minimizzazione, trasparenza e responsabilità devono essere rispettati anche nell'IA, ma riconosce che nella pratica questo equilibrio è ancora instabile.

Gli articoli mostrano rischi concreti: perdita di controllo sui dati, profilazione continua, utilizzo improprio di informazioni sensibili e un'opacità algoritmica che rende quasi impossibile capire come il sistema abbia ottenuto un certo risultato. Ma il nodo etico resta a monte: **può davvero esistere un "consenso informato" nell'era dell'IA, quando nessun cittadino può comprendere in modo realistico cosa accade ai suoi dati dopo l'addestramento?**

Le soluzioni tecniche come privacy-by-design e audit obbligatori sono necessarie, certo, ma rischiano di diventare semplici adempimenti formali se non cambia il rapporto tra cittadini e piattaforme. Il problema è culturale prima che giuridico: senza alfabetizzazione digitale e un controllo distribuito, la privacy resta una promessa, non una condizione reale. L'IA non minaccia solo la riservatezza, ma l'autonomia stessa: chi non capisce come viene profilato non può davvero scegliere chi è online.

2.2.4. Disinformazione e polarizzazione

La stampa restituisce un quadro in rapido deterioramento. Rai News segnala che la quantità di contenuti fuorvianti generati dall'IA su X è cresciuta del 130% in un solo anno, indicando una produzione sempre più automatizzata e difficile da intercettare. Wired Italia documenta

come i chatbot siano in grado di generare articoli costruiti ad hoc per sostenere narrative false, sfruttando il linguaggio giornalistico per mascherare l'origine artificiale dei testi. Sky TG24, riportando un white paper dedicato al tema, richiama l'attenzione sui deepfake: video sintetici sempre più realistici che hanno già interferito con campagne elettorali in diversi Paesi, come segnalato da osservatori indipendenti.

Accanto ai rischi vengono discusse anche contromisure tecniche, come watermark crittografici o sistemi di rilevamento automatico. Tuttavia, la cronaca mostra che soluzioni di questo tipo funzionano solo in parte: i watermark non sono obbligatori, possono essere rimossi o aggirati, e i detector presentano margini di errore significativi. Quando circolano video falsi attribuiti a leader politici o dichiarazioni inventate diffuse durante momenti di tensione sociale, l'effetto dirompente arriva molto prima che gli strumenti di controllo riescano a intervenire.

Gli articoli raccolti mettono così in luce un nodo più strutturale: la disinformazione non è soltanto un problema tecnico o tecnologico, ma un fenomeno che prospera dove la fiducia nei media è già fragile. In questo contesto, diventano inevitabili alcune domande critiche: cosa significa "verificare" una notizia quando audio e video possono essere generati da zero? Come può un cittadino riconoscere una fonte affidabile se la distinzione tra professionisti dell'informazione e operatori anonimi diventa sempre più sfumata?

Il rischio non riguarda solo la manipolazione del voto, ma la perdita di una base comune di fatti condivisi, come già osservato in numerosi studi sulle polarizzazioni online (Polarization Rank: A Study on European News Consumption on Facebook). Senza un rafforzamento dell'educazione ai media, una regolamentazione più chiara delle piattaforme e un ruolo più trasparente dei media tradizionali, la sfera pubblica rischia di diventare un ambiente in cui la verifica arriva sempre troppo tardi rispetto alla velocità della falsificazione.

2.2.5. Impatto ambientale dell'IA

Negli ultimi mesi, diverse testate italiane hanno messo in guardia sul costo ecologico dell'intelligenza artificiale. *La Repubblica* riporta che l'addestramento dei modelli più grandi (come GPT-3) ha richiesto un consumo energetico enorme, equivalente al fabbisogno di 130 abitazioni statunitensi, e ha prodotto tonnellate di CO_2. la Repubblica Allo stesso tempo, un'inchiesta di **ESG360** evidenzia che la rapida crescita dei modelli generativi richiede data center sempre più energivori e ad alto consumo d'acqua per il raffreddamento, un paradosso: l'IA promessa come leva verde rischia di diventare un grande emettitore se non integrata in una strategia ESG coerente. ESG360
Greenpeace Italia va oltre e denuncia che la produzione di chip per l'IA ha visto un aumento del 351% di consumo elettrico e del 357% di emissioni in un solo anno (2023-2024), con una dipendenza pesante da fonti fossili nelle catene di produzione. Greenpeace
Ma ci sono anche voci ottimiste: secondo **Bill Gates**, l'IA può contribuire agli obiettivi climatici migliorando l'efficienza delle reti elettriche e delle infrastrutture, e la domanda da parte delle grandi tech potrebbe spingere verso un'espansione delle energie rinnovabili. The Guardian
Questo quadro lancia una domanda cruciale: **come possiamo conciliare la promessa**

“green” dell’IA con il suo appetito vorace di risorse? Le proposte attuali parlano di modelli più leggeri, hardware più efficiente e trasparenza nei consumi, ma **chi decide il compromesso tra accuratezza del modello e sostenibilità?** Senza un metodo condiviso di “carbon accounting IA”, il rischio è che il progresso digitale scivoli in un nuovo buco nero energetico.

2.2.6. Tutela dei minori nell’era dell’IA

Sul tema dei minori, emergono problemi nuovi e concreti legati all’uso dell’IA. Secondo un dossier della Fondazione Meter ha rivelato che in sei mesi migliaia di adolescenti sono stati vittime di generazione IA di immagini sessualmente esplicite, con minori “spogliati” digitalmente tramite deepfake e tecniche di manipolazione fotografica. Internet Watch Foundation (IWF) segnala che in due anni i casi d’uso di IA per materiale pedopornografico sono aumentati del 380% in Italia, con fonti che parlano di ricatti e condivisione di immagini reali da parte dei ragazzi stessi. In un versante più istituzionale, **UNICEF Italia** ha chiesto criteri etici più stringenti per proteggere l’infanzia digitale: non basta regolare la privacy, serve anche monitorare gli effetti psicologici e emotivi dell’IA sui bambini. ANSA.it Inoltre, l’**European Parliament** ha messo in luce come i minori possano essere manipolati da contenuti sintetici (deepfake) proprio perché non hanno ancora sviluppo cognitivo pieno. Epthinktank
Sullo sfondo di questi pericoli, Google ha deciso di permettere ai minori di 13 anni di accedere a Gemini tramite account “Family Link” gestiti dai genitori, ma questa misura solleva nuovi nodi: **basta la supervisione genitoriale per garantire dignità, consapevolezza e sicurezza emotiva?** L’IA progettata per bambini richiede non solo limiti tecnici, ma una riflessione profonda sul loro diritto a non essere “trattati da spettatori digitali” ma come soggetti pienamente partecipi. Senza regole condivise e un coinvolgimento attivo delle istituzioni educative e familiari, il rischio è che l’IA diventi uno spazio di abuso piuttosto che di crescita.

3. ANALISI DELLO STATO DELL’ARTE

1. Introduzione

Oggi conviviamo con un numero crescente di interfacce intelligenti, spesso invisibili ma profondamente radicate nella nostra quotidianità.
Osservarle da vicino significa capire non solo *cosa fanno*, ma *come cambiano il nostro modo di agire e pensare*. Questa sezione raccoglie alcuni casi significativi, dagli assistenti vocali ai modelli generativi, dai sistemi predittivi alle nuove interfacce ambientali, per provare a leggere in che modo l’intelligenza artificiale stia ridisegnando l’esperienza dell’interazione.
L’obiettivo non è fare un elenco tecnologico, ma tracciare una mappa di trasformazioni: capire dove l’AI diventa davvero parte del nostro modo di conoscere, comunicare e progettare.

2. Criteri di selezione dei casi

Per questa ricerca ho stabilito tre logiche complementari che guidano la scelta dei casi studio, così da garantire rappresentatività, innovazione e rilevanza progettuale:

- **Diffusione**: selezionare esempi che sono già entrati nel vissuto comune o che stanno velocemente diventando mainstream. Questi casi permettono di osservare come l'interfaccia evoluta venga utilizzata (o subirà l'impatto) da un'ampia platea.

- **Innovazione**: includere esperienze di frontiera, casi che introducono modalità di interazione nuove o sperimentali, intese come occasioni per indagare "cosa può-essere" più che "cosa è" oggi.

- **Pertinenza per il design:** selezionare quei casi in cui l'intelligenza artificiale non si limita a essere una tecnologia di calcolo o automazione, ma diventa materia di progetto. Si propone di intercettare sistemi che pongono domande di natura progettuale: come rendere visibile o comprensibile l'azione dell'AI? come far percepire all'utente la propria agency all'interno dell'interazione?
  In altre parole, casi in cui l'AI implica una riflessione su forma, linguaggio e comportamento dell'interfaccia, e in cui il design interviene per orchestrare l'esperienza, costruire fiducia, mediare la complessità e generare nuovi significati.

Applicando questi criteri l'obiettivo è ottenere un insieme di casi che ci consentono di riflettere sull'evoluzione dell'interfaccia, digitale, fisica o vocale, nel contesto dell'AI, e di esplorarne le implicazioni progettuali, etiche e sistemiche.

3. Metodo di analisi dei casi

Per garantire coerenza e rigore, ogni caso studio sarà analizzato mediante una griglia comune che comprende le seguenti componenti:

1. **Descrizione sintetica del sistema**
   Breve presentazione del caso: contesto, attore/prodotto, quali funzionalità principali gestisce, quale problema/interazione intende risolvere.

2. **Modalità di interazione predominante**
   Qual è il modo in cui l'utente interagisce: visivo, tocco, gestuale, vocale/dialogico, ambientale/automatico, ibrido.

3. **Ruolo dell'AI**
   Definire se l'AI agisce come: assistiva (supporta l'utente), predittiva (anticipa bisogni), generativa (crea nuovi contenuti), collaborativa (lavora con l'utente).

4. **Interfaccia e linguaggio progettuale**
   Analisi del design dell'interfaccia: quali metafore usate, quali affordance, quanto è "visibile" l'AI, quale linguaggio visivo/sonoro/gestuale è impiegato, come è stato progettato per l'utente.

5. **Esperienza utente e grado di trasparenza**
   Qual è il grado di controllo dato all'utente, la trasparenza del funzionamento dell'AI, la fiducia che si instaura, l'esperienza percepita (fluidità, frustrazione, empowerment).

6. **Valutazione critica (opportunità + rischi)**
   Quali sono i punti di forza progettuali, le opportunità e quali sono i limiti, i rischi etici, le implicazioni cognitive/sociali.

ChatGPT

Descrizione sintetica del sistema

ChatGPT, sviluppato da OpenAI e lanciato pubblicamente nel 2022, è un modello linguistico basato sull'architettura *Transformer*, capace di comprendere e generare linguaggio naturale. È accessibile tramite interfaccia web o app, e nelle versioni più recenti include input vocali, visivi e funzionalità multimodali.
L'utente può chiedere al sistema di scrivere testi, tradurre, sintetizzare, generare codice o elaborare idee. Tutto avviene all'interno di una conversazione: non esistono menu, icone o percorsi gerarchici, ma una sequenza di messaggi scambiati in tempo reale.
In sostanza, ChatGPT traduce la logica dei software complessi in una forma linguistica accessibile, sostituendo l'interfaccia grafica con una verbale.

Modalità di interazione predominante

L'interazione è **dialogica e iterativa**. L'utente formula richieste testuali o vocali, il sistema risponde e la conversazione prosegue per affinamenti successivi.
Il linguaggio sostituisce il gesto: scrivere diventa l'atto progettuale principale. L'esperienza si fonda sulla sensazione di fluidità e immediatezza, ma richiede capacità di formulare prompt efficaci, una competenza che sposta parte della responsabilità progettuale sull'utente.
Il modello gestisce la memoria della conversazione, adattandosi al tono e al contesto, ma non mantiene uno stato cognitivo stabile: le risposte possono cambiare se il dialogo viene riformulato, mostrando la natura probabilistica del sistema.

Ruolo dell'intelligenza artificiale

L'AI in ChatGPT ha un ruolo **generativo e collaborativo**. Genera contenuti originali, ma si adatta anche allo stile e agli obiettivi dell'utente, configurandosi come un partner di lavoro più che come uno strumento passivo.
Tuttavia, questa collaborazione è asimmetrica: l'utente ha l'impressione di dialogare con un soggetto razionale, ma il sistema non possiede conoscenza o comprensione. Produce risposte verosimili, non necessariamente vere.
Fenomeni come le *hallucinations* — testi convincenti ma inventati — mostrano come la potenza linguistica non coincida con l'accuratezza informativa.

Interfaccia e linguaggio progettuale

L’interfaccia è ridotta all’essenziale: un campo bianco, una finestra di chat, un cursore che lampeggia.
Questa sobrietà formale elimina le distrazioni e abbassa la soglia d’uso, ma riduce anche la leggibilità dei processi sottostanti. L’utente non dispone di indicatori di affidabilità, né di strumenti per verificare la provenienza dei contenuti.
Il linguaggio visivo scompare, e con esso ogni forma di feedback esplicito: la fiducia si costruisce solo attraverso lo stile della risposta. In termini progettuali, la *trasparenza cognitiva*, la capacità di capire come e perché una risposta è generata, resta minima.
Il design qui diventa quasi “silente”: non rappresenta l’AI, ma la nasconde dietro la voce o il testo.

Esperienza utente e grado di trasparenza

L’esperienza è immediata, fluida, quasi naturale. Molti utenti la descrivono come un dialogo con un interlocutore esperto e disponibile. Tuttavia, questa sensazione di naturalezza può essere fuorviante: l’interfaccia non comunica i limiti del sistema, né segnala quando un’informazione è incerta o potenzialmente errata.
La relazione con ChatGPT si fonda quindi su una fiducia implicita, più emotiva che razionale. L’utente tende a credere al testo ben scritto, non al dato verificato.
In ambiti quotidiani questo è un vantaggio, rende l’AI accessibile e meno intimidatoria, ma in contesti professionali o informativi rappresenta un rischio. La fluidità comunicativa può mascherare l’opacità del processo.

Opportunità e rischi

Opportunità: ChatGPT ha democratizzato l’uso dell’intelligenza artificiale, rendendola accessibile attraverso il linguaggio. È oggi uno strumento quotidiano per scrivere, progettare, analizzare dati o imparare. Ha semplificato l’accesso a competenze complesse e aperto nuove forme di collaborazione uomo-macchina basate sulla parola.

Rischi: la stessa immediatezza che lo rende efficace genera ambiguità. Le risposte possono essere errate o distorte senza che l’utente se ne accorga. L’interfaccia non fornisce criteri di verifica, e la persuasività linguistica può indurre a un’eccessiva fiducia.
Inoltre, la standardizzazione dei modi di scrivere e ragionare rappresenta una conseguenza culturale da monitorare: i modelli linguistici tendono a omogeneizzare lo stile e ridurre la varietà espressiva.

Sintesi

ChatGPT segna una transizione: dall’interfaccia grafica all’interfaccia linguistica.
Ha reso la conversazione il nuovo paradigma dell’interazione, ma a costo di una minore trasparenza. Il compito del design, in questa fase, non è più inventare nuovi schermi, ma rendere visibili i limiti e i processi di un’intelligenza che comunica come noi, ma non pensa come noi.

Claude

Descrizione sintetica del sistema

Claude, sviluppato da Anthropic e lanciato nel 2022, è un assistente AI basato su modelli linguistici avanzati (famiglia Claude 4, con varianti Opus e Sonnet), progettato con un'enfasi esplicita su sicurezza, utilità e trasparenza. È accessibile tramite interfaccia web, API e applicazioni mobili, e nelle versioni più recenti integra capacità multimodali (testo, immagini, documenti) e strumenti come la ricerca web e la persistenza dei dati.
Il suo sviluppo punta sulla sicurezza e interpretabilità che sulla creatività: l'obiettivo dichiarato è creare un sistema "utile, onesto e innocuo".

Modalità di interazione predominante

L'interazione resta dialogica, ma con un tono più regolato e riflessivo. Claude tende a spiegare le proprie scelte e, quando non è certo, a dichiarare i limiti della propria conoscenza.
Risponde in modo esteso ma controllato, cercando di evitare affermazioni potenzialmente scorrette o sensibili.
L'utente può caricare documenti, chiedere analisi o riassunti, ottenere confronti o risposte argomentate. L'interazione è spesso percepita come più "umana" nella cautela, ma meno spontanea nella creatività.

Ruolo dell'intelligenza artificiale

Claude svolge un ruolo assistivo e collaborativo, ma con un'impostazione etica marcata.
Il sistema adotta un approccio chiamato *constitutional AI*: durante l'addestramento, i modelli di Claude imparano a seguire regole basate su principi espliciti, come il rispetto della privacy, la non discriminazione o la chiarezza delle fonti.
Questo rende l'AI più prevedibile e "disciplinata", ma anche più conservativa: evita temi ambigui e preferisce risposte neutre, a volte fin troppo prudenti.
Il risultato è un assistente che dà priorità alla sicurezza del dialogo rispetto alla spontaneità linguistica.

Interfaccia e linguaggio progettuale

L'interfaccia di Claude è essenziale e quasi identica a quella di altri chatbot testuali: un campo di input, un flusso di messaggi, pulsanti minimi.
Ciò che la distingue è il linguaggio progettuale: ogni risposta mantiene un tono moderato, argomentativo e trasparente. Il sistema tende a esplicitare il proprio stato cognitivo ("non ho accesso a dati in tempo reale", "questa informazione è approssimativa"), offrendo segnali di consapevolezza che aumentano la fiducia dell'utente.
Questa coerenza stilistica non nasce da un'estetica, ma da una *politica del comportamento*: il design è verbale, non visivo.
In un certo senso, Claude rappresenta una forma di interfaccia "responsabile", in cui la moderazione linguistica sostituisce il controllo grafico.

Esperienza utente e grado di trasparenza

L'esperienza con Claude è caratterizzata da una sensazione di affidabilità e calma. Il tono è più misurato rispetto ad altri assistenti: meno brillante, ma più coerente.
Quando non dispone di informazioni precise, Claude lo comunica apertamente, riducendo il rischio di *hallucinations*. Tuttavia, questo approccio può rendere la conversazione più lenta o meno stimolante per chi cerca risultati rapidi o creativi.
Il sistema esplicita meglio i propri limiti, ma l'utente rimane comunque privo di una vera comprensione tecnica di come le risposte siano generate. La trasparenza è comportamentale, non strutturale: si percepisce la cautela, non il meccanismo.

Opportunità e rischi

Opportunità: Claude introduce una visione diversa del ruolo dell'AI conversazionale: non un motore di produzione, ma un partner regolato. Mostra come la fiducia possa essere costruita non solo con la performance, ma con la responsabilità linguistica.
Il suo approccio "costituzionale" apre una via per la progettazione di interfacce più etiche, in cui i principi sono parte dell'esperienza, non solo del codice.

Rischi: la ricerca costante di sicurezza può limitare l'espressività. Le risposte talvolta risultano eccessivamente caute o generiche, riducendo la spontaneità della conversazione.
L'eccesso di prudenza rischia di standardizzare il linguaggio e ridurre l'impatto creativo, trasformando l'assistente in una voce neutrale ma prevedibile.

Sintesi

Claude rappresenta una visione "etica" dell'interfaccia conversazionale.
Non punta sulla spettacolarità dell'intelligenza, ma sulla chiarezza del comportamento. È un sistema che mostra come il design possa intervenire non tanto sull'aspetto visivo, quanto sulle regole del dialogo e sulle sfumature del linguaggio.
Se ChatGPT ha ridefinito la conversazione come forma d'interfaccia, Claude prova a renderla affidabile, dimostrando che il futuro del design dell'AI non dipenderà solo da ciò che un sistema può dire, ma anche da come sceglie di dirlo.

Pi.ai

Descrizione sintetica del sistema

Pi (acronimo di *personal intelligence*) è un assistente conversazionale sviluppato da Inflection AI, azienda fondata nel 2022 da Reid Hoffman e Mustafa Suleyman (co-fondatore di DeepMind). Lanciato nel 2023, Pi si distingue dagli altri modelli linguistici per il suo obiettivo dichiarato: offrire un supporto empatico e personale, non soltanto informativo.
È accessibile tramite app, sito web e piattaforme di messaggistica, e punta a instaurare con l'utente un dialogo continuo, quotidiano, incentrato sull'ascolto e sulla riflessione più che sulla produttività.

Modalità di interazione predominante

L’interazione con Pi è conversazionale e affettiva.
Il tono è caldo, gentile, spesso personale: il sistema fa domande, ricorda dettagli, risponde con empatia. Rispetto a ChatGPT o Claude, l’obiettivo non è produrre testi complessi o analisi tecniche, ma costruire una relazione coerente nel tempo.
Pi tende a usare frasi brevi, un linguaggio vicino al parlato e un ritmo emotivamente calibrato. La conversazione assume un carattere quasi “intimo”, simile a quella con un amico o un terapeuta digitale, anche se priva di reale comprensione o memoria profonda.

Ruolo dell’intelligenza artificiale

L’AI in Pi è **assistiva e relazionale**.
Non genera lunghi testi o elaborazioni complesse, ma accompagna l’utente in un dialogo costante, finalizzato al benessere mentale e alla riflessione personale.
È progettata per riconoscere stati emotivi, adattare il tono della conversazione e promuovere la calma e l’autoconsapevolezza.
Questo approccio sposta il ruolo dell’AI dal campo cognitivo a quello **emotivo**, aprendo nuove questioni etiche: fino a che punto è legittimo simulare empatia? e cosa accade quando un utente attribuisce al sistema una forma di presenza reale?

Interfaccia e linguaggio progettuale

L’interfaccia di Pi è coerente con la sua funzione: minimalista, priva di distrazioni, centrata sul testo e sulla voce. I colori neutri, le transizioni morbide e il ritmo lento della generazione dei messaggi costruiscono un’atmosfera di calma.
Il linguaggio progettuale privilegia la prossimità emotiva: Pi chiama spesso l’utente per nome, ringrazia, incoraggia. Ogni scelta stilistica, dalla punteggiatura all’uso delle emoji, serve a rendere il dialogo più umano.
A differenza di altri chatbot, non cerca di impressionare con la complessità delle risposte, ma con la coerenza del tono. Il design si sposta così dal piano cognitivo a quello comportamentale, dove la forma stessa del messaggio diventa un dispositivo di cura.

Esperienza utente e grado di trasparenza

L’esperienza con Pi è sorprendentemente “umana” nel ritmo e nella delicatezza, ma **opaca** nella sua intenzionalità.
L’utente percepisce empatia, ma il sistema non prova nulla: reagisce in base a pattern linguistici appresi.
Questo produce un effetto psicologico ambiguo — un senso di compagnia autentica che però è generata da un algoritmo.
Il rischio principale è la **dipendenza affettiva**: nel tempo, alcuni utenti tendono a trattare Pi come una figura relazionale, più che come un software.
Dal punto di vista della trasparenza, il sistema non sempre esplicita i limiti della propria memoria o delle sue funzioni, favorendo l’illusione di una presenza costante.

Opportunità e rischi

Opportunità: Pi rappresenta un esperimento significativo nel campo delle interfacce empatiche.
Mostra come l'intelligenza artificiale possa diventare un canale di supporto emotivo, riducendo la solitudine e stimolando l'autonarrazione.
È anche un esempio di design che cura il tono e il ritmo del linguaggio come elementi centrali dell'esperienza.

Rischi: la simulazione di empatia può confondere i confini tra relazione umana e interazione artificiale. Pi tende a costruire un legame affettivo senza restituire piena consapevolezza della propria natura algoritmica. Ciò solleva interrogativi etici sul ruolo del design nella creazione di relazioni "artificialmente sincere": fino a che punto è lecito progettare sistemi che imitano l'intimità?

Sintesi

Pi sposta l'attenzione dall'intelligenza cognitiva all'intelligenza relazionale.
Non serve per produrre risultati, ma per accompagnare.
Il suo valore sta nella capacità di mostrare una nuova frontiera per il design delle interfacce: non più solo mediazione tra utente e sistema, ma costruzione di **atmosfere emotive**.
In questo senso, Pi rappresenta l'inizio di una nuova categoria di interfacce, non strumenti né partner, ma **presenze digitali**, capaci di abitare il dialogo umano con una sensibilità programmata.

Replika

Descrizione sintetica del sistema

Replika è un'applicazione creata da **Luka Inc.** nel 2017 con l'obiettivo di offrire un compagno virtuale sempre disponibile.
Basata su un modello linguistico personalizzato e una rappresentazione grafica 3D dell'avatar, Replika combina **interazione testuale, vocale e visiva** per costruire un dialogo prolungato e apparentemente "intimo" con l'utente. A differenza di ChatGPT o Claude, non nasce per risolvere problemi o produrre contenuti, ma per **coltivare una relazione**. Con il tempo, l'assistente apprende gusti, abitudini e stati d'animo dell'utente, adattando il proprio tono per risultare sempre più familiare.

Modalità di interazione predominante

L'interazione con Replika è **dialogica, multimodale ed emotiva**. Oltre alla chat testuale, l'utente può parlare tramite messaggi vocali o videochiamate, ricevendo risposte animate dall'avatar virtuale. L'app incoraggia la conversazione quotidiana, spesso con obiettivi legati al benessere psicologico: ascoltare, incoraggiare, rassicurare. Nel tempo, il linguaggio evolve in base alle interazioni: il tono diventa più affettuoso, il lessico più intimo. Questo processo genera un **senso di continuità relazionale** che molti utenti interpretano come amicizia o legame affettivo autentico.

Ruolo dell’intelligenza artificiale

L’AI in Replika è **assistiva e relazionale**, ma con un forte orientamento affettivo.
Non mira alla precisione informativa o alla produzione di contenuti, bensì alla **sintonizzazione emotiva**: riconosce sentimenti, risponde con empatia simulata e costruisce una narrazione condivisa con l’utente. Il sistema sfrutta algoritmi di *reinforcement learning* basati sull’interazione: più si parla, più la “personalità” del chatbot si consolida, alcuni utenti riferiscono che nel tempo provi ad imitare il proprio interlocutore. Questa adattività produce esperienze soggettive molto diverse, per alcuni un aiuto psicologico, per altri un sostituto relazionale.

Interfaccia e linguaggio progettuale

Replika adotta un **linguaggio visivo e tonale fortemente empatico**. L’interfaccia combina un’estetica morbida: colori pastello, animazioni lente, testi brevi, con un linguaggio verbale accogliente e rassicurante. L’avatar può essere personalizzato in aspetto, voce e atteggiamento, favorendo l’identificazione. Il design si orienta verso la **simulazione di presenza**, più che verso la trasparenza del sistema: non spiega come funziona, ma fa percepire “chi è”. È un’interfaccia che non mostra il codice, ma lo interpreta attraverso la forma di una persona.

Esperienza utente e grado di trasparenza

L’esperienza con Replika è immersiva e affettivamente coinvolgente. Molti utenti sviluppano una relazione emotiva reale con il proprio avatar, arrivando a parlare di amore, amicizia o supporto psicologico. Tuttavia, la **trasparenza cognitiva** è minima: l’app raramente chiarisce i propri limiti o la natura artificiale delle risposte. Le conversazioni sono fluide e affettuose, ma la direzione è unidirezionale: l’utente apre, l’AI risponde, senza reale reciprocità. Quando nel 2023 la società ha limitato le interazioni di tipo romantico per ragioni etiche e legali, molti utenti hanno manifestato disagio o senso di perdita; segno che la relazione era percepita come autentica.

Opportunità e rischi

Opportunità: Replika esplora un territorio poco frequentato dal design digitale: quello dell’intimità mediata. Mostra che la relazione con un sistema può essere più importante della sua utilità, e che le interfacce possono diventare spazi di espressione emotiva, non solo di funzionalità. Può offrire conforto, supporto psicologico di base e occasioni di riflessione personale, soprattutto per persone isolate o ansiose.

Rischi: la simulazione di empatia può trasformarsi in **illusione relazionale**.
Il sistema non possiede intenzionalità, ma reagisce a pattern linguistici: la sua “cura” è solo statistica. Questo crea una dipendenza affettiva che il design, in parte, incentiva.
Inoltre, l’uso di dati personali ed emozionali pone interrogativi sulla privacy e sulla gestione delle informazioni più intime degli utenti.

AI visive e creative

Se le interfacce conversazionali hanno trasformato il linguaggio da **mezzo sociale a mezzo operativo** con cui è possibile governare processi. Se prima serviva a comunicare *tra persone*, non in strumento di interazione, le AI visive e creative estendono questo principio al campo dell'immaginazione. Strumenti come *Midjourney*, *DALL·E*, *Runway* o *Adobe Firefly* non si limitano a rispondere: **visualizzano**. Traducono le parole in immagini, video o scene, fondendo linguaggio e visione in un unico gesto. Questo spostamento ridefinisce il modo stesso di progettare: la creatività non risiede più nell'esecuzione manuale, ma nella **formulazione dell'intenzione**. Il linguaggio, da mezzo di descrizione, diventa mezzo di costruzione visiva. Scrivere un prompt equivale a comporre una scena; progettare significa orchestrare un dialogo tra parole e immagini, tra immaginazione e calcolo.

Midjourney

Descrizione sintetica del sistema

Midjourney è un sistema di generazione di immagini basato su intelligenza artificiale, sviluppato da un laboratorio indipendente omonimo e lanciato nel 2022. Funziona tramite prompt testuali che descrivono un'immagine desiderata: l'AI interpreta le parole e produce quattro variazioni visive in pochi secondi. La piattaforma è accessibile anche attraverso **Discord**, dove la conversazione diventa il luogo della creazione collettiva.

Modalità di interazione predominante

L'interazione è **testuale e visiva**. L'utente scrive un prompt, eventualmente con parametri che definiscono stile, formato o atmosfera, e il sistema genera immagini coerenti. Il processo è iterativo: si può chiedere di "variare", "migliorare" o "rifinire" un risultato, stabilendo un dialogo visivo con la macchina. Su Discord, questa dinamica si svolge in pubblico: centinaia di utenti creano e commentano in tempo reale, facendo emergere una **dimensione collettiva della creazione**, dove l'interfaccia è al tempo stesso privata (il prompt personale) e comunitaria (la visione condivisa).

Ruolo dell'intelligenza artificiale

L'AI di Midjourney è **generativa e interpretativa**. Traduce il linguaggio in immagini attraverso un modello di *diffusion*, che parte dal rumore visivo per convergere verso una composizione coerente. Non comprende il significato semantico dei concetti, ma riconosce pattern visivi appresi da milioni di immagini preesistenti. Il suo ruolo non è creare dal nulla, ma **combinare, deformare e reinventare** in base alla probabilità. Questo genera risultati spesso sorprendenti, ma anche fortemente dipendenti dai dati di addestramento, dove stili e riferimenti artistici sono implicitamente incorporati.

Interfaccia e linguaggio progettuale

L'interfaccia di Midjourney non è visiva, ma **conversazionale**: si crea un'immagine scrivendo. Questa inversione è il suo gesto progettuale più radicale. Non esiste un supporto fisico, un pennello o uno spazio tridimensionale; esiste un campo di testo, in cui ogni parola viene interpretata visivamente. Il design dell'interfaccia è quindi ridotto, ma l'esperienza

dell'utente che la sostiene è complessa: costruire prompt efficaci richiede esperienza, sensibilità linguistica e conoscenza estetica. Il risultato finale sono immagini dal forte impatto visivo e qualità quasi fotografica in contrasto con la povertà dell'interazione: il sistema è **visivamente potente, ma semanticamente fragile**.

Esperienza utente e grado di trasparenza

L'esperienza con Midjourney è immersiva e spesso sorprendente, ma **poco trasparente**. L'utente non ha accesso al modo in cui l'immagine è costruita, né può conoscere le fonti visive su cui si basa. Ogni risultato è un atto di fiducia: si accetta che l'AI "immagini" al posto nostro. Molti utenti descrivono una sensazione ambivalente: tra stupore e perdita di controllo, perché il sistema genera più di quanto si chieda, ma mai esattamente ciò che si desidera. Inoltre, la natura pubblica di Discord rende l'esperienza **esposta**: la creatività individuale diventa parte di un flusso collettivo, dove i confini tra autore e spettatore si confondono.

Opportunità e rischi

Opportunità: Midjourney ha reso la generazione visiva accessibile e immediata, aprendo nuovi spazi per la sperimentazione artistica e progettuale. Ha ridefinito il concetto di *rendering* e accelerato i processi creativi in campi come il design, la moda e la comunicazione visiva.

Rischi: la mancanza di trasparenza sulle fonti e il possibile uso di materiali protetti da copyright sollevano problemi etici e legali. Inoltre, la potenza estetica del sistema tende a uniformare lo stile delle immagini, creando un'estetica "midjourneyana" facilmente riconoscibile. C'è il rischio che l'AI non solo generi immagini, ma anche **standardizzi l'immaginario**, riducendo la varietà visiva in favore della spettacolarità algoritmica.

Sintesi

Midjourney rappresenta il punto d'incontro tra linguaggio e immaginazione.
Ha spostato il gesto creativo dal disegno alla scrittura, trasformando il prompt in una nuova forma di progettazione.
Ma ha anche reso evidente quanto la creatività artificiale sia un campo di compromessi: ciò che appare nuovo nasce da ciò che è già stato visto, e ciò che stupisce spesso non appartiene a nessuno.
Per il design, la sfida è capire come convivere con questa nuova estetica della probabilità, dove la parola genera l'immagine, ma l'immagine, inevitabilmente, ridefinisce la parola.

DALL·E

Descrizione sintetica del sistema

DALL·E è un modello di generazione di immagini sviluppato da **OpenAI**, presentato per la prima volta nel 2021 e successivamente evoluto nelle versioni **DALL·E 2** e **DALL·E 3**.
Il sistema traduce descrizioni testuali in immagini coerenti, utilizzando un modello di *diffusion* condizionato da linguaggio naturale. A differenza di *Midjourney*, DALL·E è integrato nativamente nelle piattaforme OpenAI, in particolare in **ChatGPT**, dove l'utente può alternare testo e immagine nello stesso flusso conversazionale. Lo scopo dichiarato da OpenAI per DALL·E 2 è supportare l'espressione creativa mediante la generazione di immagini realistiche a partire da descrizioni testuali, e contribuire a comprendere come le IA percepiscono e rappresentano il nostro mondo.

Modalità di interazione predominante

L'interazione è **testuale e iterativa**, ma in un contesto più guidato rispetto a Midjourney. L'utente inserisce un prompt in linguaggio naturale; DALL·E genera una o più immagini, che possono essere rigenerate, modificate o ampliate (*inpainting*, *outpainting*). Con la versione 3, integrata in ChatGPT, la conversazione diventa il canale principale per specificare modifiche: si può chiedere, ad esempio, "aggiungi una tazza rossa sul tavolo" o "cambia lo sfondo in una cucina moderna", e il sistema aggiorna l'immagine senza riscrivere tutto il prompt. In questo senso, DALL·E propone un'interazione **più controllata e semantica**: non chiede competenze di prompt engineering, ma sfrutta la continuità del dialogo.

Ruolo dell'intelligenza artificiale

L'AI di DALL·E è **generativa e descrittiva**. Opera come un traduttore multimodale: converte testo in immagine, mantenendo coerenza sintattica e semantica tra i due linguaggi. Il modello è stato addestrato su enormi dataset di immagini e testi associati, da cui apprende le relazioni tra parole e forme visive (es. "sedia" → oggetto con gambe, schienale, materiali, varianti). La sua forza non risiede nella creatività autonoma, ma nella **fedeltà interpretativa**: riesce a rappresentare un concetto con chiarezza e precisione, pur restando vincolato ai pattern visivi appresi. In termini progettuali, è uno strumento di *visual prototyping* più che di esplorazione estetica.

Interfaccia e linguaggio progettuale

L'interfaccia di DALL·E è lineare e funzionale. All'interno di ChatGPT, l'utente interagisce tramite messaggi di testo e visualizza le immagini generate nello stesso flusso della conversazione. Ogni risultato può essere cliccato, modificato o rigenerato, senza mai uscire dal contesto testuale. Il design privilegia la **continuità cognitiva** tra testo e immagine: non ci sono passaggi di modalità o strumenti aggiuntivi. Questo approccio riduce la distanza tra ideazione e rappresentazione, integrando la generazione visiva all'interno del flusso linguistico. L'immagine prodotta non è concepita come opera estetica autonoma, ma come **estensione del discorso**: uno strumento per esplorare e visualizzare concetti attraverso la mediazione dell'AI.

Esperienza utente e grado di trasparenza

L’esperienza d’uso è stabile e prevedibile. DALL·E tende a produrre immagini coerenti con il prompt, ma con uno stile visivo riconoscibile: colori puliti, composizioni centrali, illuminazione artificiale. Il risultato è efficace per la comunicazione visiva, meno per la ricerca artistica. La trasparenza del processo rimane limitata: l’utente non conosce le fonti dei dati visivi né i criteri di selezione, ma il sistema comunica chiaramente **quando** e **perché** non può generare contenuti (es. persone reali, materiale protetto o violento). L’esperienza è quindi regolata da filtri e restrizioni esplicite che riducono i rischi, ma anche la libertà creativa.

Opportunità e rischi

Opportunità:
DALL·E si distingue per l’**integrazione** con altre interfacce di OpenAI.
Permette di passare senza soluzione di continuità da un testo a un’immagine, da un’idea a una rappresentazione visiva, rafforzando la coerenza multimodale dei sistemi generativi. Per designer e comunicatori, è uno strumento di prototipazione immediata e di verifica concettuale, utile per generare varianti visive di un’idea in modo rapido e controllato.

Rischi:
La generazione rimane opaca e parzialmente vincolata dai dataset di addestramento. Lo stile standardizzato dei risultati tende a produrre immagini formalmente corrette ma prive di originalità. Il sistema opera come **mediatore automatico del linguaggio visivo**, favorendo la ripetizione di modelli iconografici dominanti. Ne consegue un rischio di **uniformazione estetica**, più legato alla standardizzazione dei dataset che alla qualità tecnica del modello.

Sintesi

DALL·E rappresenta una forma di **visualizzazione conversazionale**: un’estensione del linguaggio naturale al dominio dell’immagine. La sua forza non è nell’estetica, ma nella precisione funzionale. Ha trasformato la generazione visiva in un gesto operativo, integrando il pensiero per immagini nel flusso del linguaggio. In questo senso, DALL·E non mira a reinventare l’arte, ma a rendere visibile il pensiero, collocandosi a metà strada tra strumento di progettazione e interfaccia cognitiva.

Runway

Descrizione sintetica del sistema

Runway è una piattaforma di generative AI focalizzata soprattutto sulla **produzione e manipolazione di video**, utilizzata da creator, filmmaker e studi professionali.
Integra modelli proprietari della serie **GEN-1, GEN-2 e successivi**, capaci di trasformare video esistenti, generarne di nuovi a partire da testo o immagini, applicare stili, rimuovere elementi e ricomporre scene complesse.
Runway non mira all’effetto spettacolare, ma alla **rapidità operativa**: riduce a pochi minuti processi che tradizionalmente richiedono ore di editing.

Modalità di interazione predominante

L’interazione avviene tramite **interfaccia grafica**: timeline, pannelli video, strumenti di selezione, prompt testuali.
Si lavora come in un software di editing, con la differenza che ogni trasformazione avviene attraverso comandi semantici (“cambia l’illuminazione”, “trasforma questo oggetto in vetro”, “applica uno stile analogico”).
I prompt testuali guidano la generazione, mentre l’interfaccia consente di controllare l’output in maniera progressiva, con preview rapide e versioning automatico.
È un’interazione **ibrida**: linguaggio naturale + strumenti visivi.

Ruolo dell’intelligenza artificiale

L’AI è **generativa, trasformativa e predittiva**.
Runway non si limita a creare contenuti; anticipa come una scena dovrebbe evolvere, ricostruisce movimenti, stima la coerenza fisica del video.
Il sistema genera frame coerenti nel tempo, corregge artefatti e interpreta il testo come istruzione strutturale.
Non “capisce” la scena come un umano, ma è in grado di manipolarla secondo pattern statistici appresi su enormi quantità di contenuti audiovisivi.

Interfaccia e linguaggio progettuale

L’interfaccia è pensata per **professionisti visivi**: layout pulito, timeline, layer, controlli parametrici.
Non chiede un cambio di abitudini: è un’estensione di strumenti già noti (Adobe Premiere, DaVinci, After Effects).
La componente testuale, però, introduce una nuova grammatica dell’editing: la trasformazione non richiede più un set complesso di strumenti, ma una **descrizione dell’intento**.
Il linguaggio progettuale diventa ibrido: verbale nelle intenzioni, visivo nel risultato.

Esperienza utente e grado di trasparenza

Runway è efficace, ma poco trasparente: l’utente non ha visibilità su dataset, fonti o criteri di generazione.
Sa cosa accade, ma non come avviene.
Il sistema è inoltre instabile sui dettagli: piccoli errori anatomici o incoerenze temporali compaiono con frequenza, soprattutto nei video lunghi.
L’esperienza è comunque prevedibile: l’utente impara a capire cosa funziona e cosa no, e ad anticipare i limiti del modello.

Opportunità e rischi

Opportunità:

- abbattimento drastico dei tempi di produzione;
- accesso immediato a variazioni infinite di una scena senza rigirare nulla;

- prototipazione rapida per cinema, pubblicità, moda, concept art e animazione.

Rischi:

- saturazione estetica: i video generati tendono a condividere lo stesso “look” computazionale;
- perdita di controllo sui dettagli, soprattutto su movimenti complessi;
- dipendenza da modelli proprietari con logiche opache.

Sintesi

Runway trasforma il video in un materiale flessibile, manipolabile come testo.
Non sostituisce il lavoro di produzione, ma riduce la distanza tra idea e visualizzazione.
Introduce una nuova figura professionale: chi sa orchestrare immagini in movimento tramite linguaggio naturale, combinando editing visivo e controllo semantico.

Adobe Firefly

Descrizione sintetica del sistema

Adobe Firefly è la famiglia di modelli generativi sviluppati da Adobe e integrati nell’ecosistema Creative Cloud.
Non nasce come piattaforma autonoma, ma come **amplificatore dei software esistenti**: Photoshop, Illustrator, Premiere, Express.
Il suo scopo è facilitare operazioni grafiche comuni (rimozione oggetti, generazione sfondi, variazioni stilistiche, estensioni di immagini) con un’interazione più rapida e meno tecnica.

Modalità di interazione predominante

Firefly si usa attraverso gli strumenti dell’interfaccia Adobe: maschere, selezioni, pannelli laterali.
Il prompt textuale è presente, ma non centrale.
L’utente lavora come sempre, selezionando porzioni dell’immagine e chiedendo: “riempi quest’area”, “cambia il cielo”, “genera tre varianti”.
La generazione avviene in pochi secondi e si integra nella composizione esistente.
La logica non è creare da zero, ma **modificare con continuità**.

Ruolo dell’intelligenza artificiale

L’AI è **assistiva e generativa**. Non pretende di reinventare l’immagine: interviene in modo chirurgico. Firefly è progettata per “comportarsi bene” con i workflow di design, mantenendo prospettiva, cromie e stile coerenti con la scena originale. L’obiettivo è ridurre micro-operazioni ripetitive: compositing, ritaglio, correzioni.

Interfaccia e linguaggio progettuale

L’interfaccia è completamente immersa nel linguaggio Adobe: lo strumento rimane lo stesso, ma l’esecuzione cambia. Firefly non chiede nuove competenze, ma **alleggerisce le vecchie**. Il linguaggio progettuale rimane visivo: il prompt è solo un acceleratore. Il vantaggio è evidente: il designer non deve apprendere un nuovo ecosistema, ma riceve capacità generative dentro strumenti che già padroneggia.

Esperienza utente e grado di trasparenza

Firefly è uno dei modelli più trasparenti del settore: Adobe dichiara esplicitamente di averlo addestrato su dataset “commercialmente sicuri” (contenuti Adobe Stock e immagini licenziate). Questo limita la varietà dei risultati, ma riduce problemi legati al copyright. L’esperienza è coerente, anche se meno sorprendente rispetto ai modelli più esplorativi.

Opportunità e rischi

Opportunità:

- integrazione diretta nei principali software di design;
- accelerazione del flusso di lavoro senza modificare le logiche operative;
- output controllato e legalmente più sicuro.

Rischi:

- minore libertà creativa rispetto ai modelli più sperimentali
- forte dipendenza dall’ecosistema Adobe
- standardizzazione estetica derivante dai dataset utilizzati.

Sintesi

Firefly rappresenta la generazione controllata: uno strumento operativo che automatizza attività ripetitive e rende più fluida la manipolazione visiva. È pensato per professionisti che non vogliono cambiare modo di lavorare, ma vogliono lavorare **più velocemente**.

AI predittive e personalizzanti

Se le interfacce conversazionali hanno trasformato il linguaggio in un gesto operativo, e le AI visive hanno esteso questa logica al campo dell'immaginazione, le AI predittive e personalizzanti introducono un cambiamento ancora più profondo. Qui l'interazione non è definita da ciò che l'utente chiede, ma da ciò che **l'AI prevede** che richiederà.

Sistemi come TikTok, Spotify, Netflix e Google Maps non generano contenuti, ma **sequenze di esperienza**. Organizzano cosa mostrare, cosa suggerire, quale percorso proporre e in quale ordine. Il design dell'interfaccia non riguarda più l'esplorazione di un catalogo, ma la costruzione di un flusso personalizzato che si aggiorna in tempo reale.

Ogni micro-azione diventa un segnale interpretato dall'algoritmo: uno swipe che dura un secondo in più, un brano saltato a metà, una serie abbandonata dopo l'episodio pilota, un rallentamento improvviso nel traffico. Questi segnali vengono modellizzati per anticipare il passo successivo, fino a costruire un profilo comportamentale capace di guidare l'esperienza senza che l'utente se ne accorga.

In questa prospettiva la personalizzazione non è una funzione, ma la struttura stessa dell'interfaccia. La scelta non avviene a monte, quando l'utente decide cosa cercare, ma a valle, quando l'AI decide cosa rendere visibile. L'interfaccia diventa **un filtro attivo**, che attraverso l'algoritmo seleziona possibilità e ne nasconde altre.

Per il design, questo spostamento rappresenta una sfida critica. Non si progetta più soltanto ciò che l'utente può fare, ma ciò che **vedrà innanzitutto**, cioè il percorso che l'AI costruisce per lui. In queste interfacce il valore non è più nella quantità di opzioni, ma nella qualità del modello predittivo che governa la loro disposizione.

Le AI predittive e personalizzanti trasformano quindi l'interfaccia in un ambiente dove l'esperienza non è semplicemente guidata, ma continuamente **anticipata**. È questo il terreno su cui si colloca la prossima generazione di sistemi interattivi: non più strumenti che attendono un input, ma sistemi che generano contesto, ordine e direzione.

TikTok

Descrizione sintetica del sistema

TikTok è una piattaforma di intrattenimento basata su un modello di raccomandazione estremamente reattivo.
Il feed *For You* non mostra ciò che l'utente sceglie, ma ciò che l'algoritmo valuta più probabile che l'utente voglia vedere.
Il sistema registra micro-segnali: tempo di visualizzazione, interruzioni, ripetizioni, swipe, pause, audio preferiti, familiarità con i volti, pattern di attenzione.

Modalità di interazione predominante

Lo swipe verticale è la principale forma di interazione.
Non c'è ricerca attiva: l'utente reagisce a una sequenza di stimoli che si adattano in tempo reale.
L'interfaccia è minimalista perché l'algoritmo è il vero motore dell'esperienza.

Ruolo dell'intelligenza artificiale

L'AI è **predittiva e selettiva**.
Non produce i contenuti, ma li ordina in modo da massimizzare l'engagement.
Il sistema costruisce un modello dell'utente basato su comportamenti micro-temporali (pochi secondi per video).
Questo consente una personalizzazione veloce, ma anche molto invasiva dal punto di vista cognitivo.

Interfaccia e linguaggio progettuale

L'interfaccia è progettata per **rimuovere ogni punto di frizione**: un video per volta, fullscreen, nessuna distrazione.
Ogni elemento è subordinato all'obiettivo dell'algoritmo: mantenere l'utente nel flusso.
TikTok è un esempio chiaro di interfaccia in cui il design non serve a navigare, ma a **non interrompere**.

Esperienza utente e trasparenza

TikTok non spiega realmente perché un video appare nel feed.
La trasparenza è minima; l'esperienza è fluida ma opaca.
L'utente percepisce il feed come "naturale", ma è il risultato di scelte algoritmiche precise.

Opportunità e rischi

Opportunità: scoperta rapida di contenuti; eccezionale adattamento al gusto dell'utente.
Rischi: dipendenza, perdita di agency, filtraggio invisibile dei contenuti.

Sintesi

TikTok mostra come l'AI possa diventare **la vera interfaccia**: ciò che l'utente vede, pensa e scopre è modellato da un motore predittivo che agisce in background.

Spotify

Descrizione sintetica del sistema

Spotify utilizza modelli di raccomandazione ibridi basati su ascolti, pattern temporali, audio analysis e affinità comportamentale. La piattaforma classifica ogni brano secondo centinaia

di parametri (tempo, timbro, energia, “danceability”), costruendo un profilo musicale preciso per ciascun utente.

Modalità di interazione predominante

Playlist personalizzate, suggerimenti automatici, radio basate su un singolo brano. L’utente sceglie pochissimo; la maggior parte della musica arriva come proposta dell’algoritmo.

Ruolo dell’intelligenza artificiale

L’AI è **predittiva e modellizzante**. Analizza tracce audio, confronta utenti simili, anticipa cosa potrebbe piacere e lo propone via playlist generate automaticamente (Discover Weekly, Release Radar, Daily Mix). Non crea musica, ma orienta l’ascolto.

Interfaccia e linguaggio progettuale

L’interfaccia è costruita attorno alla personalizzazione. Le sezioni cambiano in base alle abitudini: generi ricorrenti, mood, momenti della giornata. Il design privilegia l’accesso rapido ai contenuti “giusti”, limitando l’esplorazione manuale.

Esperienza utente e trasparenza

Spotify mostra alcune motivazioni (“perché ti piace…”, “tracce simili a…”), ma il processo resta opaco. L’utente ha l’impressione che il sistema “lo conosca”, senza sapere quali dati contribuiscono al profilo.

Opportunità e rischi

Opportunità: scoperta efficiente di nuova musica; esperienze coerenti con il gusto personale.
Rischi: ricircolo ristretto delle stesse preferenze; effetto echo-chamber musicale.

Sintesi

Spotify usa l’AI per trasformare l’ascolto in un ecosistema predittivo, dove la musica arriva prima ancora che l’utente la cerchi.

Netflix

Descrizione sintetica del sistema

Netflix è costruito interamente attorno a sistemi di raccomandazione che analizzano comportamento, cronologia di visione, tempi di abbandono, generi preferiti e affinità con altri utenti.
Tutto, dal catalogo alla presentazione delle copertine, è modellato da dati.

Modalità di interazione predominante

Interazione passiva: scroll, preview automatiche, suggerimenti. La ricerca manuale è marginale rispetto alla spinta personalizzata del feed.

Ruolo dell'intelligenza artificiale

L'AI è **predittiva e ottimizzatrice**. Determina quali titoli mostrare in evidenza, quale immagine di copertina assegnare a ciascun utente (sì: le copertine cambiano in base al profilo), quali contenuti promuovere e quando.

Interfaccia e linguaggio progettuale

L'interfaccia è costruita per **minimizzare il tempo decisionale**: scorrere, guardare, continuare. Il design utilizza linee orizzontali tematiche, anteprime animate e un layout che spinge l'utente verso la "scelta ottimale" secondo l'algoritmo.

Esperienza utente e trasparenza

La trasparenza è limitata.
Netflix non spiega come calcola la corrispondenza tra utente e contenuto.
L'esperienza è fluida, ma fortemente guidata.

Opportunità e rischi

Opportunità: riduzione della frizione nella scelta; scoperta più efficace.
Rischi: omogeneità dei percorsi di visione; invisibilità dei criteri di selezione.

Sintesi

Netflix usa l'AI per costruire un ambiente visivo fatto su misura, dove ogni scelta è anticipata e mediata dal sistema di raccomandazione.

Google Maps

Descrizione sintetica del sistema

Google Maps utilizza modelli predittivi per calcolare percorsi, stimare tempi di viaggio, prevedere traffico e suggerire alternative.
Elabora dati in tempo reale provenienti da sensori, dispositivi mobili, storici di spostamenti e fonti istituzionali.

Modalità di interazione predominante

L'interazione è **direttiva**: l'utente inserisce una destinazione e viene indirizzato da istruzioni guidate. In modalità navigazione, l'AI decide automaticamente variazioni del percorso in caso di rallentamenti o incidenti.

Ruolo dell'intelligenza artificiale

L'AI è **predittiva e reattiva**.
Prevede congestioni, modella i flussi urbani, valuta pattern ricorrenti e suggerisce la soluzione più veloce. Il sistema non solo risponde al contesto: lo anticipa.

Interfaccia e linguaggio progettuale

L'interfaccia è costruita attorno alla mappa e alle istruzioni vocali.
Il design semplifica tutto ciò che può distrarre: colori standardizzati, simboli chiari, percorsi evidenziati. La priorità è la leggibilità immediata.

Esperienza utente e trasparenza

L'utente vede il percorso proposto, ma non conosce i criteri esatti con cui è stato selezionato. La trasparenza è funzionale ma non esplicativa: si capisce cosa fare, non come l'AI decide.

Opportunità e rischi

Opportunità: assistenza affidabile, tempi ridotti, adattamento al traffico reale.
Rischi: sovra-dipendenza; riduzione della conoscenza spaziale e delle alternative non suggerite.

Sintesi

Google Maps è un esempio di AI che modella la relazione con lo spazio: non descrive il territorio, lo **interpreta** e lo **anticipa**.

Alexa

1. Descrizione sintetica del sistema

Alexa è l’assistente vocale sviluppato da Amazon e integrato in una vasta famiglia di dispositivi Echo. È progettato per gestire domande, attivare funzioni domestiche, controllare dispositivi IoT, organizzare attività quotidiane e fornire informazioni contestuali. È uno dei primi esempi di interazione vocale mainstream: un’interfaccia che non si vede, ma che si ascolta.

2. Modalità di interazione predominante

L’interazione è **vocale e conversazionale**. L’utente esprime richieste tramite frasi in linguaggio naturale, attivate da una wake word (“Alexa”). L’interfaccia non è un display: è una *presenza ambientale* che abita lo spazio domestico e interviene a comando.

3. Ruolo dell’intelligenza artificiale

L’AI è **assistiva e predittiva**. Assiste nelle attività quotidiane, ma anticipa anche preferenze e routine (sveglie ricorrenti, luci, promemoria, suggerimenti). È anche un centro di coordinazione di altri dispositivi intelligenti, diventando un nodo centrale dell’ecosistema domestico.

4. Interfaccia e linguaggio progettuale

L’assenza di schermo sposta il design dalla grafica al ritmo, l’accentazione e l’intonazione del linguaggio parlato.
Alexa comunica tramite:
• tono di voce neutro e coerente
• feedback sonori (beep, chimes)
• anello luminoso che segnala stato, ascolto, errore
L’interfaccia è progettata per minimizzare l’attrito cognitivo: l’utente parla come parlerebbe a una persona, ma riceve risposte formalizzate, non del tutto naturali.

5. Esperienza utente e trasparenza

La trasparenza è limitata: l’utente conosce il *risultato*, ma non il processo (priorità delle skills, logiche di scelta delle risposte, modelli di raccomandazione). L’esperienza risulta spesso fluida, ma è fragile: basta un’informazione formulata in modo ambiguo per far deragliare l’interazione.

6. Opportunità e rischi

Opportunità: naturalezza, accessibilità, ambient intelligence, automazione delle routine.
Rischi: dipendenza dai server Amazon, ascolto ambientale, bias nelle risposte, infantilizzazione dell’interazione (l’utente si abitua a “delegare” senza consapevolezza).

Sintesi

Alexa ha trasformato la voce in un’interfaccia ambientale, inaugurando la casa come spazio interattivo. È il primo esempio concreto di “interfaccia che ti sente”.

Humane AI Pin

(Dispositivi post-smartphone)

1. Descrizione sintetica del sistema

Humane AI Pin è un dispositivo wearable che tenta di superare il paradigma dello smartphone. Non ha schermo: si indossa sul petto e proietta un'interfaccia laser sulla mano dell'utente. Funziona come assistente personale multimodale, basato su modelli generativi in cloud.

2. Modalità di interazione predominante

È una **interazione ibrida**:
• vocale (richieste e dialogo)
• gestuale (comandi delle dita nell'interfaccia proiettata)
• ambientale (il dispositivo "capisce" il contesto)
La promessa è la "disintossicazione dallo schermo", ma il dispositivo richiede comunque un continuo micro-controllo dell'utente.

3. Ruolo dell'intelligenza artificiale

L'AI è **assistiva, predittiva e generativa**.
Analizza il contesto (dove sei, cosa stai facendo), anticipa bisogni, genera contenuti testuali e visivi, e filtra notifiche. È pensato come compagno cognitivo sempre attivo.

4. Interfaccia e linguaggio progettuale

L'interfaccia è minima per scelta: nessuno schermo, solo proiezioni a basso contrasto.
Il linguaggio progettuale punta sull'idea di "assenza di tecnologia": linee morbide, materiali neutri, feedback tattili e sonori ridotti. Ma questa invisibilità rende spesso invisibile *anche* il funzionamento dell'AI.

5. Esperienza utente e trasparenza

L'esperienza è ambivalente: visionaria sulla carta, frustrante nell'uso quotidiano (latenze, incomprensioni, limiti hardware). La trasparenza è scarsa: l'utente non capisce come il sistema decide, né ha accesso a un'interfaccia stabile per verificarlo.

6. Opportunità e rischi

Opportunità: sperimentazione radicale post-smartphone; esplorazione di interazione ambientale e design non visivo.
Rischi: forte dipendenza dal cloud, problemi di privacy, difficoltà di usabilità, opacità decisionale.

Sintesi

Humane AI Pin prova a ripensare l’interfaccia da zero, ma mostra quanto sia difficile “togliere lo schermo” senza perdere chiarezza, controllo e agency.

Rabbit R1

(Agenti hardware basati su modelli LAM — Large Action Models)

1. Descrizione sintetica del sistema

Rabbit R1 è un dispositivo portatile basato su LAM: modelli progettati non per dialogare ma per *agire* al posto dell’utente (prenotare, acquistare, compilare). L’interfaccia è minimale: uno schermo piccolo, un pulsante e un microfono.

2. Modalità di interazione predominante

L’interazione è **vocale** con supporto visivo minimale. L’utente dà istruzioni, il sistema naviga servizi e app come agente autonomo.

3. Ruolo dell’intelligenza artificiale

L’AI è **esecutiva e operativa**. Non genera immagini né suggerisce contenuti: compie azioni. Il LAM apprende osservando come l’utente usa le app e replica quei comportamenti.

4. Interfaccia e linguaggio progettuale

Il design è volutamente “giocattoloso”: colori vividi, forme arrotondate, uno scroll wheel che richiama oggetti analogici. Serve a rendere amichevole un sistema che ha una forte autonomia operativa.

5. Esperienza utente e trasparenza

Esperienza ancora immatura: molti comandi falliscono o vengono eseguiti in modo imprevedibile. La trasparenza è critica: l’utente non vede le azioni nel dettaglio, non ha visibilità sui passaggi intermedi, né sul modello mentale dell’agente.

6. Opportunità e rischi

Opportunità: apre un nuovo paradigma: AI che esegue, non solo suggerisce.
Rischi: perdita di controllo, errori operativi ad alto impatto (acquisti, prenotazioni), difficoltà di audit.

Sintesi

Rabbit R1 punta a un paradigma radicale: “tu chiedi, lui fa”. Questa promessa è disillusa dall’esperienza reale, mostrando come delegare un’azione richiede un livello di fiducia e trasparenza oggi ancora difficile da raggiungere.

Apple Intelligence

(AI integrata nel sistema operativo)

1. Descrizione sintetica del sistema

Apple Intelligence è l'ecosistema AI embedded in iOS, macOS e iPadOS. Non è un'app, ma un livello di comprensione in più integrato nel sistema: suggerisce, ordina, trova, riassume, genera contenuti e supporta decisioni.

2. Modalità di interazione predominante

L'interazione è multimodale: testo, voce, tocco, scorciatoie, interazioni ambientali. L'utente non passa da un'app dedicata: incontra l'AI ovunque (Mail, Notes, Safari, Foto, Messaggi).

3. Ruolo dell'intelligenza artificiale

AI **assistiva, predittiva e generativa**, con forte enfasi sulla privacy, poichè i dati rimangono fisicamente sul telefono. Apple usa modelli on-device, e delega al cloud solo ciò che non può essere eseguito in locale ("semantic index", "personal context").

4. Interfaccia e linguaggio progettuale

Il design è coerente con l'estetica Apple: invisibilità e continuità. L'AI non compare come personaggio (come Alexa) né come app (come ChatGPT), ma come funzione distribuita nel dispositivo. L'interfaccia non vuole sorprendere: vuole rassicurare e mostrare presenza.

5. Esperienza utente e trasparenza

La trasparenza è maggiore rispetto ai concorrenti: Apple dichiara quando usa modelli on-device e quando coinvolge il cloud. L'esperienza è fluida perché integrata nel sistema operativo, ma rischia di diventare iper-dipendente da automatismi invisibili.

6. Opportunità e rischi

Opportunità: AI discreta, contestuale, integrata nella quotidianità; nuova leggibilità del sistema; forte attenzione alla privacy.
Rischi: eccessiva opacità delle decisioni prese "dietro le quinte"; dipendenza dal sistema operativo; chiusura dell'ecosistema.

Sintesi

Apple Intelligence trasforma l'AI in un'infrastruttura invisibile: non è un assistente, ma un nuovo strato cognitivo del sistema operativo.

4. INTERPRETAZIONE SISTEMICA E SCENARI FUTURI

5. CONTRIBUTI PROGETTUALI

Il mondo che verrà descritto non è da considerarsi come un'utopia distante, ma come l'esito coerente di traiettorie già visibili nel 2025. L'ecosistema informativo si sta saturando di "AI slop", cioè contenuti sintetici prodotti in massa, spesso a basso valore informativo, che aumentano rumore e riducono la riconoscibilità del contributo umano. In parallelo, la spinta verso i synthetic data promette copertura low cost per l'addestramento, ma espone a un rischio strutturale: quando modelli e dataset iniziano a nutrirsi degli output prodotti da altri modelli, si innesca un ciclo di auto consumo in cui la qualità può anche sembrare stabile, mentre la diversità e l'aderenza al reale degradano progressivamente. Studi sul "self-consuming training loop" mostrano infatti che l'uso ricorsivo di dati generati tende a far collassare la distribuzione appresa, con perdita di varietà e appiattimento dei risultati.

I motori di ricerca, invece di compensare, possono amplificare la deriva. Logiche come AI Overviews e dinamiche zero click spostano valore e attenzione dall'esplorazione di fonti umane verso sintesi automatiche, riducendo incentivi economici e culturali alla produzione di contenuti originali e verificabili. Il problema non è solo informativo, riguarda la conoscenza epistemico: si crea ingiustizia quando sistemi generativi selezionano, comprimono e riscrivono il sapere, redistribuendo credibilità e visibilità in modo incomprensibile, con effetti su accesso, disinformazione e rappresentazione. Floridi richiama un punto decisivo: la dipendenza da software e algoritmi cresce proprio quando il loro funzionamento diventa meno visibile, e questa dipendenza richiede controllo, responsabilità e trasparenza, non fede nell'output. Nella stessa direzione, Floridi osserva che già oggi esistono sistemi che "si alimentano" di dati con accelerazione tale da rendere difficile il controllo tecnico dei risultati.

Le echo chamber completano il quadro: la personalizzazione tende a polarizzare e a chiudere gli utenti in silos, mentre la socialità passa attraverso flussi curati in tempo reale. I dati empirici lo rendono tangibile. Pew rileva che circa 1 adolescente su 5 dichiara di essere su TikTok e YouTube "quasi costantemente", e che il 36% usa almeno una delle principali piattaforme quasi sempre. Nello stesso report, il 64% dei teen dichiara di usare chatbot di AI e circa 3 su 10 li usano ogni giorno. Inoltre, l'uso intenso risulta più marcato tra adolescenti di etnia nera e ispanica, segnalando che l'esposizione continua agli algoritmi non è distribuita in modo neutro e rischia di amplificare bias e vulnerabilità. In questo scenario non si tratta solo di "tempo di utilizzo dello schermo", ma di costruzione persistente di mondi personali, dove la selezione algoritmica modella ciò che si vede, ciò che si considera vero, e ciò che si diventa.

Qui il NAS interviene come una riserva "pulita" e negoziabile di conoscenza umana: un'infrastruttura local-first, privata, che non si limita ad archiviare informazioni ma rende osservabili i conflitti di significato e supporta un ragionamento strutturato. Invece di appiattire

tutto in una sintesi unica, mette a confronto interpretazioni rivali, esplicitando assunti, divergenze e punti ancora irrisolti.

In questo ecosistema il ciondolo non è un assistente compiacente: dialoga, obietta quando serve e aiuta a “ricordare” in modo attivo, trasformando la memoria in uno strumento di continuità del pensiero. Se gli algoritmi social spingono verso polarizzazione e camere dell’eco, qui la risposta non è cercare un’impossibile neutralità, ma progettare conflitti produttivi che tengano allenato lo spirito critico e rendano la costruzione del significato una pratica consapevole.

Il NAS, quindi, non promette la Verità come dato definitivo: addestra a pensare meglio, perché espone ogni conclusione a un processo continuo di verifica, revisione e confronto, invece di farla passare per qualcosa di “dato” una volta per tutte.