

L'interfaccia liquida,
Spatial Computing, Generative UI,
il design post — schermo

L'interfaccia liquida,
Spatial Computing, Generative UI,
il design post — schermo

[Abstract]

L'evoluzione dell'interfaccia uomo-macchina (HCI) sta attraversando la sua più radicale trasformazione dall'introduzione dell'interfaccia grafica a finestre (GUI) negli anni Ottanta. Dopo decenni dominati dalla metafora dello schermo bidimensionale, un confine rettangolare che separa il mondo fisico da quello digitale, l'avvento dello Spatial Computing, della realtà estesa (XR) e delle interfacce cervello-computer (BCI) sta progressivamente erodendo questa barriera, spingendo il digitale a fondersi con l'ambiente fisico e biologico dell'utente.

La presente tesi si propone di indagare le implicazioni progettuali di questo passaggio generazionale. Se lo schermo scompare, quali diventano le regole della composizione? Come si progetta l'invisibile? Attraverso un percorso che intreccia analisi storica, ergonomia cognitiva e teoria del design speculativo, il lavoro è strutturato in quattro macro-sezioni.

Il primo capitolo traccia l'evoluzione storica dell'interazione, evidenziando la tendenza alla progressiva smaterializzazione dell'hardware e alla riduzione dell'attrito cognitivo. Il secondo capitolo analizza l'anatomia della nuova interazione spaziale, definendo una grammatica basata sulla prossemica digitale, sull'input multimodale (occhio, mano, voce) e sulla consapevolezza del contesto. Il terzo capitolo esplora le frontiere dell'intelligenza artificiale applicata al design, teorizzando l'avvento di interfacce "liquide" e generative, capaci di auto-costruirsi in tempo reale in risposta all'intento dell'utente.

A chiusura dell'indagine, il quarto capitolo valida il framework teorico proposto adottando la metodologia del Design Fiction. Attraverso la simulazione di tre scenari applicativi (domestico, urbano e professionale), vengono messe alla prova le linee guida formulate, dimostrando come esse risolvano concretamente le frizioni tra tecnologia ubiqua, privacy e contesto sociale.

L'obiettivo finale della ricerca non è solo descrittivo, ma propositivo: la tesi formula un framework euristico composto da sei linee guida per il design post-schermo, offrendo agli operatori del settore una bussola metodologica per navigare la transizione da un design di artefatti visivi a un design di comportamenti adattivi e neuro-etici.

[Indice]

[1] L'evoluzione dell'interfaccia:
dalla macchina di calcolo alla simbiosi

8

[1.1] L'Archeologia dell'interfaccia: dal Batch Processing alla linea di comando (1945 – 1975)	10
[1.2] La rivoluzione visiva e il paradigma WIMP (1973 – 1995)	15
[1.3] L'Interfaccia ipertestuale e la navigazione (1990 – 2005)	20
[1.4] L'Interfaccia somatica: Il mobile e la NUI (2007 – 2015)	24
[1.5] La dissoluzione dell'interfaccia: Zero-UI, Spatial Computing e AI (2015 – Oggi)	29

[2] Architetture dell'interazione
post-schermo: spazio, corpo e contesto

36

[2.1] La prossemica digitale: organizzare l'Informazione nello spazio	38
[2.2] Embodied interaction: il corpo come cursore	41
[2.3] Multimodalità: la convergenza dei sensi	45
[2.4] Interazione situata e context-awareness: l'ambiente come interfaccia	49
[2.5] L'architettura del suono: audio spaziale e sonificazione	51
[2.6] Inclusive design: accessibilità nello Spatial Computing	52

[3] L'interfaccia liquida:
Generative UI e adattamento cognitivo

54

[3.1] La crisi del layout statico: oltre la metafora del contenitore	56
[3.2] Generative UI: l'intelligenza artificiale come co-designer in tempo reale	58
[3.3] Anticipatory Design: la riduzione dell'attrito decisionale e il rischio della bolla	60
[3.4] Proposta di guidelines per l'interfaccia post-schermo: un framework euristico	64

[4] Validazione sperimentale:
scenari di design fiction

70

[4.1] Metodologia: il prototipo diegetico come strumento di verifica	72
[4.2] Scenario A: l'ambiente domestico e la “Calm Technology”	72
[4.3] Scenario B: mobilità urbana e la bolla di privacy	74
[4.4] Scenario C: ambito professionale e Agency collaborativa	76
[4.5] Sintesi dei risultati	78

[1] L'evoluzione dell'interfaccia: dalla macchina di calcolo alla simbiosi

Per comprendere la direzione futura dell'interazione uomo-macchina, verso un paradigma spaziale liquido e guidato dall'intelligenza artificiale, è indispensabile analizzare la traiettoria storica che ci ha condotto fin qui. La storia dell'Human-Computer Interaction (HCI) non deve essere letta come una mera cronologia di avanzamenti tecnici o di potenza di calcolo, ma come un processo continuo e inesorabile di riduzione dell'attrito.

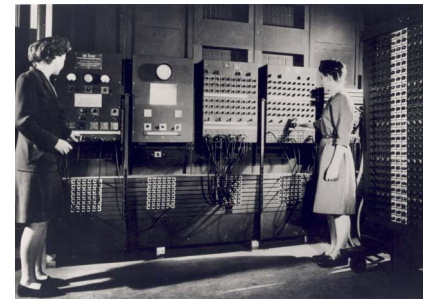
Dalle prime macchine di calcolo, che richiedevano all'uomo di adattare il proprio pensiero alla logica binaria della macchina, stiamo assistendo a un'inversione di tendenza in cui è la macchina a doversi adattare ai modelli mentali, fisici e biologici dell'essere umano. Questo capitolo ripercorre le tappe fondamentali di questa evoluzione, identificando come ogni cambio di paradigma, dalla Command Line all'interfaccia grafica, fino al touch e alle interfacce neurali, abbia avuto come obiettivo primario la rimozione degli intermediari fisici e cognitivi. Analizzando il passaggio dai sistemi rigidi e astratti del passato verso sistemi sempre più naturali e "invisibili", tratteremo la linea evolutiva che porta dritto alla visione profetica di J.C.R. Licklider: la simbiosi totale tra uomo e computer.

[1.1] L’archeologia dell’interfaccia: dal Batch Processing alla Command Line (1945 – 1975)

[1.1.1] L’assenza di contatto: Il Batch Processing e il “sacerdozio” informatico

Il termine “interfaccia” è oggi molto diffuso e si usa per descrivere qualsiasi superficie di contatto tra un utente e un sistema digitale. Tuttavia, la sua adozione nel lessico informatico, è relativamente recente. Etimologicamente, il termine deriva dal latino inter-facies (“tra due facce”) e fu utilizzato per la prima volta nella fisica del XIX secolo per descrivere la superficie di separazione tra due stati della materia (ad esempio, tra un liquido e un gas). Solo negli anni Sessanta il termine migrò nel campo della computazione per indicare, inizialmente, la connessione fisica tra due dispositivi hardware e, successivamente, lo spazio di traduzione cognitiva tra l’essere umano e il codice binario. Per comprendere l’evoluzione delle interfacce contemporanee, è fondamentale analizzare quella che possiamo definire l’archeologia dell’interazione: il periodo storico in cui l’interfaccia non era progettata per facilitare l’utente, ma per proteggere la macchina, una risorsa all’epoca rara e costosissima.

Nei primi due decenni dell’informatica commerciale e accademica (anni Quaranta e Cinquanta del Novecento), il concetto di User Interface (UI) come lo intendiamo oggi non esisteva. I grandi calcolatori mainframe, come l’ENIAC o l’UNIVAC, e successivamente i sistemi IBM della serie 700, operavano secondo il paradigma del Batch Processing (elaborazione a lotti). In questo contesto, l’interazione uomo-macchina era caratterizzata da una totale asincronia in quanto l’utente, spesso uno scienziato o matematico, non aveva accesso diretto all’elaboratore. L’interfaccia fisica era costituita dalla scheda perforata (punched card), un cartoncino rigido su cui le istruzioni venivano codificate tramite la presenza o l’assenza di fori in posizioni specifiche. La programmazione era un’attività “offline”. L’utente scriveva il codice su carta, lo traduceva mentalmente in fori, e utilizzava una macchina perforatrice per creare il pacchetto fisico di schede.



Programmatrici al lavoro sull’ENIAC (1946). In questa fase primordiale, l’interfaccia coincideva con l’hardware stesso: la riprogrammazione richiedeva la riconfigurazione manuale di cavi e interruttori.

[1.1.2] Le radici teoriche: cibernetica e Memex

Una volta preparato il “batch”, esso veniva consegnato fisicamente a un operatore specializzato. Come affermano Campbell-Kelly e Aspray nella loro storia del computer (1996), si creava una vera e propria gerarchia sociale, spesso definita il “sacerdozio” (priesthood) dei mainframe.

I computer erano isolati in stanze climatizzate (“glass rooms”), accessibili solo ai tecnici in camice bianco. L’utente restava fuori, separato da un vetro, in attesa dei risultati.

L’aspetto critico di questa fase archeologica è l’assenza del feedback loop. In un sistema interattivo moderno, l’errore viene segnalato istantaneamente come ad esempio una sottolineatura rossa mentre si digita. Nel Batch Processing invece il feedback loop poteva durare giorni. Se l’utente commetteva un singolo errore di sintassi o perforava male una scheda, l’intero processo si bloccava. Il risultato, restituito ore o giorni dopo sotto forma di tabulati cartacei, riportava semplicemente l’errore, costringendo l’utente a ricominciare l’intero ciclo da capo. L’interfaccia, dunque, non era un luogo di dialogo, ma un muro di separazione. La macchina era un oracolo silenzioso e l’interazione era un atto di sottomissione ai tempi e ai rigidi protocolli dell’hardware (Ceruzzi, 2003).

Mentre la prassi tecnica era dominata dalla lentezza e complessità del Batch Processing, sul piano teorico si stavano gettando le basi concettuali per l’interazione in tempo reale. Per parlarne dobbiamo introdurre due figure successivamente riconosciute come gli architetti del pensiero moderno sull’interfaccia: Vannevar Bush e J.C.R. Licklider. Ancora prima dell’avvento dei computer digitali, nel 1945, Vannevar Bush pubblicò sull’Atlantic Monthly il saggio visionario “As We May Think”. Bush, preoccupato dalla crescente difficoltà nel gestire l’enorme mole di conoscenza umana, teorizzò il Memex (Memory Extender). Sebbene immaginato come un dispositivo

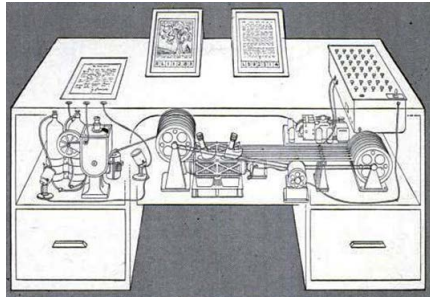


Illustrazione del “Memex” di Vannevar Bush (Life Magazine, 1945). Una scrivania elettromeccanica teorica concepita per archiviare e collegare informazioni per associazione, anticipando il concetto di ipertesto.

elettromeccanico basato su micro-film, il Memex introduceva concetti che sarebbero diventati centrali quarant’anni successivi:

L’archiviazione associativa: Bush criticava l’indicizzazione alfabetica o numerica, tipica delle biblioteche e dei primi computer, considerata “artificiale”. La mente umana opera per associazione: un pensiero ne evoca un altro istantaneamente. L’interfaccia ideale doveva quindi permettere di creare “trails” di connessioni tra documenti diversi. Questo concetto sarà precursore all’introduzione dell’ipertesto (Bush, 1945).

Il Memex poneva l’utente al centro del processo di recupero dell’informazione, non come mero calcolatore, ma come esploratore di conoscenze.

Nello stesso periodo il concetto di cibernetica teorizzato da Norbert Wiener (1948) fornì il substrato matematico per l’interazione. Wiener definì la comunicazione come un processo circolare di azione e retroazione (feedback). Perché una macchina potesse essere considerata “intelligente” o utile in senso adattivo, doveva essere in grado di comunicare il proprio stato all’operatore e modificare il proprio comportamento in base ai nuovi input (Wiener, 1948). Questa teoria smantellava concettualmente la linearità del Batch Processing a favore di un sistema dinamico.

La formulazione di queste idee trova terreno fertile nel lavoro di J.C.R. Licklider, psicologo sperimentale e informatico presso l’ARPA. Nel suo fondamentale paper del 1960, “Man-Computer Symbiosis”, Licklider propose una nuova visione radicale: il computer non doveva essere un sostituto dell’uomo, ma un partner cognitivo. Licklider identificò una divisione del lavoro precisa:

- La macchina eccelle nelle operazioni algoritmiche, ripetitive e deterministiche.
- L’essere umano eccelle nel pensiero euristico, nella formulazione di ipotesi e nel riconoscimento di pattern vaghi.

Per realizzare questa simbiosi, Licklider sosteneva che fosse

[1.1.3] La rivoluzione del Time-Sharing e l’emergere del terminale



Un terminale Teletype (TTY). Con il Time-Sharing nasce l’interazione in tempo reale, ma il dialogo uomo-macchina avviene ancora tramite output cartaceo impresso meccanicamente.

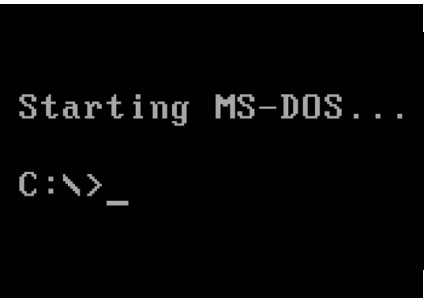
necessario abbattere la barriera temporale del Batch Processing: “Se l’uomo deve cooperare con la macchina, i due devono essere in grado di scambiarsi informazioni a ritmi compatibili” (Licklider, 1960).

La transizione dalle tesi teoriche sopra enunciate all’applicazione pratica avvenne a metà degli anni Sessanta con lo sviluppo dei sistemi di Time-Sharing, come il CTSS del MIT e successivamente il sistema Multics.

Il Time-Sharing sfruttava la velocità di elaborazione del computer per simulare la simultaneità: la CPU passava da un utente all’altro in frazioni di secondo, dando a ciascuno l’illusione di avere la macchina interamente a propria disposizione. Questo cambiamento di architettura della macchina trasformò radicalmente l’interfaccia. Non era più necessario consegnare schede perforate a un operatore, ci si poteva sedere direttamente di fronte a un terminale ed agire direttamente. Inizialmente, questi terminali erano telescriventi elettromeccaniche (Teletype o TTY), derivate dalla tecnologia telegrafica: ovvero delle macchine da scrivere collegate via cavo al computer che stampavano le risposte su rotoli di carta continua. Nonostante l’output fosse ancora cartaceo, l’interazione si era trasformata in una conversazione uomo-macchina. L’utente digitava una riga, premeva invio, e la macchina rispondeva stampando la riga successiva.

Il passo successivo fu la sostituzione della carta con il fosforo grazie all’avvento dei Video Display Terminals (VDT) o dei “Glass Teletypes”. Lo schermo a tubo catodico (CRT) introdusse due nuove dimensioni: la volatilità e la velocità. Il testo poteva apparire, scorrere e scomparire istantaneamente. Come osserva Lev Manovich nel “Il linguaggio dei nuovi media” (2001), lo schermo divenne non solo uno spazio di visualizzazione, ma uno spazio di controllo. La luce divenne la materia prima dell’interfaccia.

[1.1.4] L’egemonia della Command Line Interface (CLI)



L’interfaccia a riga di comando (CLI). L’utente si confronta con uno schermo vuoto e un prompt in attesa, un sistema che richiede un alto carico mnemonico (Recall) per conoscere la sintassi corretta.

Con la diffusione dei terminali video negli anni Settanta e l’arrivo dei primi microcomputer, si consolidò il paradigma dominante della Command Line Interface (CLI), sfruttato da sistemi come Unix e in un secondo momento, MS-DOS. La CLI rappresenta un’interfaccia linguistica pura. L’utente si trova di fronte a uno schermo vuoto, spesso nero o verde, abitato solo da un cursore lampeggiante (il prompt). Questo cursore è un segnale di attesa: la macchina è pronta, ma muta. Dal punto di vista della psicologia cognitiva e dell’ergonomia, la CLI presenta sfide enormi, analizzabili attraverso la distinzione fatta da Nielsen (1993) tra Recall e Recognition.

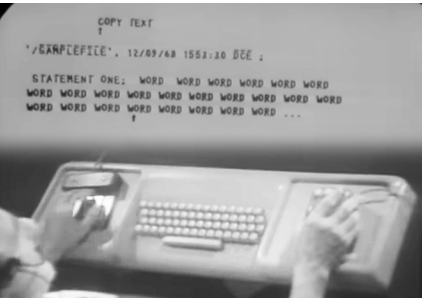
Nell’interfaccia a riconoscimento (es. GUI/Menu) l’utente vede le opzioni disponibili (es. “Salva”, “Stampa”) e deve solo riconoscerle e selezionarle per questo motivo il carico cognitivo richiesto all’utente è basso. Nell’ interfaccia a richiamo (CLI) invece lo schermo non mostra nulla, l’utente deve ricordare a memoria il comando esatto (es. grep, chmod, dir), la sua sintassi e i parametri opzionali. L’interfaccia a riga di comando non ammette errori, un singolo carattere sbagliato provoca il fallimento dell’operazione, spesso con messaggi di errore difficili da decifrare (“Syntax Error”). Neal Stephenson, nel suo celebre saggio “In the Beginning... Was the Command Line” (1999), paragona l’interazione via CLI a un dialogo con un interlocutore onnisciente ma letterale fino all’ossessione, privo di qualsiasi capacità di intuire l’intento dell’utente (Stephenson, 1999). Nonostante la sua ripida curva di apprendimento, la CLI non deve essere vista solo come un limite tecnologico. Essa offriva, e offre tuttora agli utenti esperti, una potenza espressiva e una velocità operativa ineguagliabili. Attraverso meccanismi come il piping (in cui l’output di un comando diventa l’input del successivo), la CLI permetteva di concatenare operazioni complesse creando veri e propri

[1.2] La rivoluzione visiva e il paradigma WIMP (1973 – 1995)

“programmi al volo”. Tuttavia, finché l’interazione rimaneva vincolata alla memorizzazione di sintassi testuali astratte, il computer sarebbe rimasto uno strumento per un’élite tecnica. Per realizzare la visione di Licklider di una vera simbiosi accessibile a tutti, era necessario un nuovo paradigma in cui non era più necessario comunicare alla macchina cosa fare digitandolo, ma mostrarglielo agendo.

Se l’interfaccia a riga di comando aveva trasformato il computer in un interlocutore linguistico, richiedendo all’utente di apprendere una sintassi rigorosa, la fase successiva dell’evoluzione delle interfacce mirava a trasformarlo in un ambiente spaziale navigabile. Questa transizione, che portò alla nascita e alla diffusione della GUI (Graphical User Interface), non rappresentò un semplice miglioramento estetico, ma un radicale cambio di paradigma epistemologico. Si passò dalla descrizione verbale di un’azione alla sua esecuzione tramite manipolazione diretta.

[1.2.1] Dalla sintassi allo spazio: l’eredità di Engelbart e il dilemma dell’input



Douglas Engelbart durante la “Mother of All Demos” (1968). La sua visione prevedeva un input bimanuale ad alta efficienza: mouse per puntare, tastiera a accordi (keyset) per digitare.

Sebbene l’esplosione commerciale della GUI sia associata agli anni Ottanta, le sue radici concettuali e tecnologiche risalgono alla visione pionieristica di Douglas Engelbart presso lo Stanford Research Institute (SRI). Engelbart, diversamente dai suoi contemporanei, non era interessato all’automazione d’ufficio fine a se stessa, ma perseguiva l’obiettivo dell’ “Augmented Human Intellect”: l’uso della tecnologia per estendere le capacità cognitive umane grazie all’azione in tempo reale. Nella storica dimostrazione del 9 dicembre 1968, nota come “The Mother of All Demos”, Engelbart presentò il sistema NLS (oN-Line System). Per la prima volta lo schermo non mostrava solo un flusso di testo, ma era diviso in “finestre” (viewports) che permettevano la visualizzazione simultanea di più documenti, integrando il testo e la grafica vettoriale (Engelbart, 1968). Tuttavia, l’aspetto più rivoluzionario riguardava la filosofia che si celava dietro l’input. Engelbart

[1.2.2] Xerox PARC e la crociata contro i “modi”

comprendeva che la tastiera era insufficiente per navigare in uno spazio bidimensionale. Tale pensiero lo portò a sviluppare il “mouse” (all’epoca un semplice blocco di legno con due ruote metalliche perpendicolari) pensato per controllare un puntatore sullo schermo. È fondamentale sottolineare, al fine di comprendere successiva evoluzione, che la visione originale di Engelbart era basata su un’interazione che faceva uso di entrambe le mani. Mentre la mano destra guidava il mouse per la navigazione spaziale, la mano sinistra operava su un dispositivo chiamato Chord Keyset: una tastiera a cinque tasti simile ai tasti di un pianoforte. L’idea era permettere all’utente esperto di inserire comandi o digitare testo (usando combinazioni di accordi binari) senza mai staccare la mano dal mouse e lo sguardo dallo schermo. Il futuro di questi due dispositivi spiega bene la prima lezione darwiniana nell’Interaction Design: la bassa barriera d’ingresso vince sull’efficienza a lungo termine. Da un lato il mouse divenne uno strumento universale per la sua immediatezza cognitiva grazie al mapping diretto del movimento fisico della mano nello spazio reale in quello virtuale. D’altra parte il Keyset scomparve in quanto richiedeva un carico mnemonico e un addestramento motorio troppo elevati. Questo dimostrò come l’interfaccia di successo è quella che minimizza lo sforzo di apprendimento iniziale, anche a costo di una minore velocità operativa per l’utente esperto (Bardini, 2000).

Le intuizioni di Engelbart trovarono la loro piena applicazione allo Xerox PARC (Palo Alto Research Center) negli anni Settanta. Qui, un team di ricercatori guidati da Alan Kay, Butler Lampson e Larry Tesler sviluppò lo Xerox Alto (1973), quello che possiamo considerare come primo personal computer il quale introdusse per la prima volta il Bit-mapped Display. L’idea di trattare ogni pixel dello schermo come un bit di memoria manipolabile rese possibile il concetto di WYSIWYG (What You See Is What You Get),



L’interfaccia dello Xerox Star 8010 (1981). Sviluppata al PARC di Palo Alto, introdusse per la prima volta icone, finestre sovrapponibili e il concetto di “WYSIWYG” (What You See Is What You Get) in un prodotto commerciale.

[1.2.3] La psicologia del paradigma WIMP

eliminando la discrepanza tra ciò che si vedeva a video e il risultato stampato. Oltre all’hardware in sé, il contributo più significativo portò dal PARC fu la filosofia dell’editing modeless (senza modi). Fino ad allora, gli editori di testo (come Vi o Bravo) operavano per stati modali distinti: l’utente doveva entrare in “Insert Mode” per scrivere e in “Command Mode” per cancellare o muoversi. Questa modalità generava frequenti errori cognitivi: se l’utente dimenticava in quale stato si trovava il sistema, la digitazione di una parola poteva inavvertitamente attivare comandi distruttivi. Larry Tesler, ricercatore al PARC, considerava i modi una barriera cognitiva inaccettabile e fonte di ansia (“mode errors”). Tesler sviluppò i comandi universali di Cut, Copy e Paste (taglia, copia, incolla), una metafora derivata dall’editing fisico dei giornali con forbici e colla. Questa innovazione rese l’editing fluido e privo di stati nascosti permettendo all’utente di selezionare e agire in qualsiasi momento. Questo principio di trasparenza dello stato del sistema divenne poi un pilastro fondamentale delle moderne linee guida di usabilità (Tesler, 1981).

All’interno del PARC venne codificato il paradigma che ancora oggi domina l’informatica desktop: il WIMP (Windows, Icons, Menus, Pointer). Ogni componente di questo sistema fu progettato per risolvere specifici limiti cognitivi umani, basandosi sugli studi nascenti di psicologia cognitiva e sul modello GOMS (Goals, Operators, Methods, and Selection rules) elaborato da Card, Moran e Newell nel 1983.

- Windows (finestre): introducono il multitasking visivo. Essendo la memoria di lavoro umana limitata, le finestre fungono da memoria esterna, permettendo all’utente di mantenere il contesto di più attività (“context switching”) senza perdere informazioni, mimando la disposizione di fogli su una scrivania reale.
- Icons (icone): rappresentano l’applicazione della semiotica all’informatica. Le icone

[1.2.4] La democratizzazione: Apple Macintosh e l'umanizzazione della tecnologia

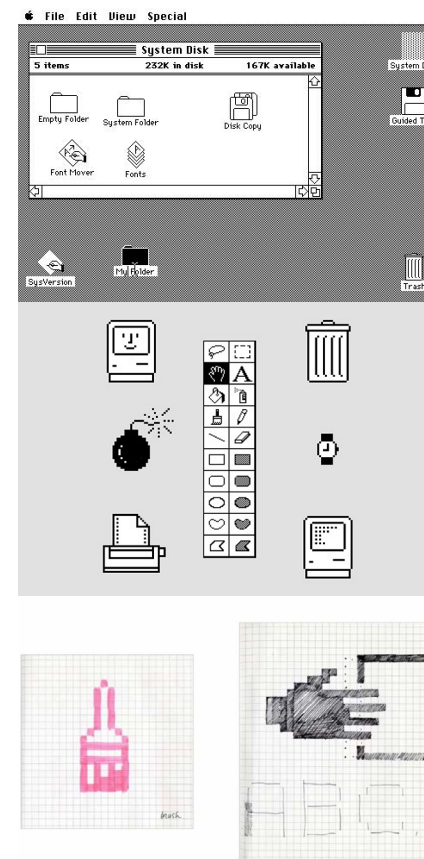
funzionano come ancore cognitive che concretizzano concetti astratti. In questo modo un file di codice binario diventa un documento leggibile permettendo in tal modo agli oggetti manipolabili di essere facilmente manipolabili.

- Menus (menu): risolvono definitivamente il problema del Recall vs Recognition. L'utente non è più costretto a ricordare a memoria i comandi come avveniva nella CLI, ma ha solo bisogno di riconoscerli all'interno di una lista visibile permettendo in tal modo di ridurre drasticamente il carico mnemonico richiesto all'utente
- Pointer (puntatore): estende la propriocezione dell'utente nello spazio digitale, creando un senso di “agency” fisica diretta sugli oggetti virtuali.

Il culmine di questo lavoro e dello sviluppo di tali teorie fu lo Xerox Star (1981), il primo sistema commerciale ad implementare rigorosamente la metafora della scrivania (Desktop Metaphor).

I progettisti scelsero lo skeuomorfismo concettuale caratterizzato dall'uso di cartelle, cestini e schedari al fine di creare un ponte cognitivo per gli impiegati d'ufficio. Sfruttare le loro conoscenze pregresse del mondo fisico permetteva di facilitare l'uso del dispositivo e la comprensione del mondo digitale (Smith et al., 1982).

Se lo Xerox Star definì le regole, colui che le rese accessibili e culturalmente rilevanti fu Apple. Dopo la celebre visita al PARC nel 1979, Steve Jobs e il suo team, comprendendo che la complessità di Star (costoso e lento) ne avrebbe impedito la diffusione, intrapresero un processo di radicale semplificazione. Un esempio emblematico di questo approccio riguarda il design del mouse. Il mouse Xerox possedeva tre pulsanti, il che richiedeva all'utente di scegliere l'azione corretta prima di cliccare. Apple, collaborando con la ditta di design IDEO, sviluppò un mouse a un solo pulsante. Questa scelta, spesso criticata dagli utenti esperti per la sua limitatezza, fu in realtà una decisione ergonomica precisa per eliminare



L'interfaccia del Macintosh System 1 (1984). La metafora della scrivania (Desktop), con cartelle e cestino, rese il computer accessibile ai non tecnici nascondendo la complessità del codice.

[1.2.5] Ergonomia e standardizzazione: la legge di Fitts

l'incertezza decisionale e ridurre il carico cognitivo per l'utente che approcciava il computer per la prima volta. (Gladwell, 2011). Con il lancio del Macintosh nel 1984, l'interfaccia subì anche una trasformazione estetica ed emotiva grazie al lavoro di Susan Kare. Fino ad allora, le icone erano disegnate da programmatori ed erano spesso rappresentazioni tecniche e fredde. Kare, designer grafica con un background in belle arti, portò una sensibilità umanistica nel design dei pixel. Lavorando su fogli a quadretti per simulare la griglia di pixel, Kare creò un linguaggio visivo amichevole e universale: l'orologio da polso per indicare l'attesa (al posto di una clessidra astratta), la bomba per l'errore di sistema (chiara ma ironica) e il “Happy Mac” che accoglieva l'utente all'avvio. Un caso studio significativo della semiotica di Kare è l'icona del tasto “Command”. Dovendo rappresentare un tasto di funzione astratto senza usare testo, Kare sfogliò un dizionario di simboli e trovò il nodo di Bowen, utilizzato in Svezia per indicare luoghi di interesse turistico nei campeggi. Quel simbolo, ancora presente sulle tastiere odierne, dimostra come l'interfaccia sia un costrutto culturale che attinge a significati esterni per rendere comprensibile la tecnologia (Kare, 2019).

L'approccio di Apple alla GUI introdusse anche il rigore scientifico nell'ergonomia del software, esemplificato dall'applicazione della Legge di Fitts. Questa legge, derivata dalla psicologia motoria, predice il tempo necessario per raggiungere un bersaglio in funzione della sua distanza e dimensione. Bruce Tognazzini, esperto di interfacce Apple, utilizzò questo principio per giustificare la posizione della Menu Bar del Macintosh, ancorata al bordo superiore dello schermo, in contrasto con la scelta di Microsoft Windows di posizionare i menu all'interno di ogni finestra. Poiché il cursore del mouse si arresta contro il bordo fisico dello schermo, quel bordo possiede un'altezza virtuale “infinita”. L'utente può lanciare il

mouse verso l’alto con un gesto rapido senza dover mirare con precisione, rendendo l’accesso ai menu su Mac significativamente più veloce e meno faticoso a livello cognitivo-motorio rispetto a Windows (Tognazzini, 1992). Con la diffusione globale di Windows 3.0 (1990) e Windows 95, il paradigma WIMP e la metafora della scrivania divennero lo standard dell’interazione uomo-macchina. Le Human Interface Guidelines di Apple e Microsoft imposero una coerenza che permise a milioni di persone di trasferire le proprie competenze da un software all’altro. Per quasi due decenni, l’interfaccia si cristallizzò in questa forma: un utente seduto, uno schermo verticale e una mano che manipola oggetti virtuali su una scrivania illusoria.

Mentre il paradigma WIMP consolidava la gestione del computer locale, trasformandolo in un’efficiente scrivania virtuale, nei laboratori di ricerca europei si stava verificando una rivoluzione che avrebbe ridefinito l’interfaccia non solo come spazio di lavoro, ma come un nuovo spazio di connessione. L’avvento del World Wide Web negli anni Novanta costrinse i progettisti a ripensare radicalmente le metafore dell’interazione. Se da un lato la GUI desktop riguardava la manipolazione di oggetti statici, la GUI del nascente Web riguardava il movimento attraverso l’informazione.

Nel 1989, Tim Berners-Lee, ricercatore presso il CERN di Ginevra, propose un sistema per la gestione delle informazioni distribuite. Il problema che Berners-Lee intendeva risolvere un problema di natura cognitiva e organizzativa che provocava la perdita di informazioni causata dal turnover dei ricercatori e l’incompatibilità tra diversi sistemi informatici. La sua soluzione era il World Wide Web nato dall’unione di due tecnologie preesistenti: l’interne (infrastruttura di rete) e l’ipertesto. Il concetto di ipertesto non era

nuovo infatti era stato già teorizzato da Vannevar Bush nel 1945 e coniato da Ted Nelson nel 1963, ma l’implementazione di Berners-Lee introdusse una caratteristica tecnica che cambiò la natura dell’interfaccia: l’unidirezionalità del link. Nel sistema Xanadu immaginato da Ted Nelson, i collegamenti dovevano essere bidirezionali e mantenere l’integrità referenziale ovvero se si spostava un documento, il link si sarebbe aggiornato. Berners-Lee d’altra parte, al fine di favorire la scalabilità globale, accettò il rischio del “link rotto” (Error 404). Questa decisione architetturelle ebbe un impatto profondo sull’esperienza utente: la navigazione web divenne un’attività intrinsecamente incerta, un’esplorazione euristica in cui l’utente deve costantemente valutare l’affidabilità del percorso (Berners-Lee, 1999).

Quando Berners-Lee presentò la sua proposta formale al suo supervisore, Mike Sendall, questi scrisse sulla copertina del documento una nota passata alla storia: “Vague but exciting” (“Vago ma eccitante”). Questa vaghezza era intrinseca al modello di interfaccia proposto, non c’era una gerarchia centrale, non c’era un indice maestro, l’interfaccia del Web era, per definizione, decentralizzata. L’utente non “apriva” un programma, ma “navigava” (browsing). Questo termine suggerisce un movimento orizzontale e serendipico, tipico della ricerca web, piuttosto che l’approfondimento verticale tipico del database (Berners-Lee, 1999).

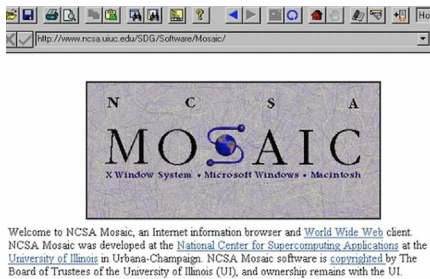
Il primo browser di Berners-Lee (WorldWideWeb) era anche un editor, mantenendo viva l’idea dell’utente come creatore attivo. Tuttavia, la vera esplosione del Web come fenomeno di massa avvenne nel 1993 con il rilascio di NCSA Mosaic da parte del team guidato da Marc Andreessen. Mosaic introdusse due cambiamenti cruciali nell’interfaccia:

– L’interfaccia grafica (GUI) per il web, aggiunse pulsanti visivi per “indietro”, “avanti” e “home”, fornendo orientamento spaziale in un ambiente privo di mappe.

[1.3] L’Interfaccia ipertestuale e la navigazione (1990 – 2005)

[1.3.1] La genesi del link: Berners-Lee e la rottura della linearità

[1.3.2] Mosaic e la nascita della “pagina” come metafora



NCSA Mosaic (1993). L'inclusione delle immagini nel flusso del testo ha definito la metafora della “pagina web”, imponendo lo scrolling verticale come interazione dominante.

[1.3.3] Architettura dell'informazione e il problema del “Lost in Hyperspace”

– Il tag . Prima di Mosaic, le immagini venivano aperte in finestre separate. Andreessen decise di permettere l'inclusione di immagini inline, ovvero all'interno del flusso del testo. La decisione di Andreessen di inserire le immagini nel flusso del testo fu controversa tra i puristi dell'informazione, ma definì la metafora dominante del Web: la pagina. A differenza della “scrivania” la quale è un ambiente 2D dove gli oggetti possono essere sovrapposti liberamente, la “pagina web” eredita la struttura dell'editoria cartacea, un layout verticale che richiede lo scrolling la quale divenne l'interazione primaria. Mentre nelle prime GUI lo scrolling era considerato un difetto poiché nascondeva le informazioni, nel web divenne la norma. Questo spostò il modello mentale dell'utente rendendo l'interfaccia non più un contenitore finito come una finestra di dialogo, ma una pergamena potenzialmente infinita. L'utente imparò a “scansionare” visivamente lunghe porzioni di contenuto alla ricerca di parole chiave o link blu, sviluppando quella che Jakob Nielsen avrebbe poi definito lettura a F-Pattern (Nielsen, 1999).

Con la crescita esponenziale delle pagine web emerse un nuovo problema di usabilità. L'utente poteva sentirsi disorientato. Il disorientamento è noto in letteratura come il fenomeno del “Lost in Hyperspace”. L'assenza di fisicità tipica di un libro e di confini dati da un file system locale gli utenti faticavano a creare una mappa mentale del sito in cui si trovavano. Al fine di risolvere questa crisi cognitiva, nacque una nuova disciplina conosciuta come “Architettura dell'Informazione” (IA). Louis Rosenfeld e Peter Morville, nel loro testo fondamentale “Polar Bear Book”, codificarono i sistemi di navigazione che oggi diamo per scontati: briciole di pane (breadcrumbs), barre di navigazione globali, sitemap e tassonomie gerarchiche (Rosenfeld & Morville, 1998). Questi sistemi di navigazione pongono le loro basi concetto del “profumo dell'informazione” (Information Scent). Esso

rappresenta assieme alla teoria del foraggiamento informativo (Information Foraging) uno dei concetti teorici più affascinanti sviluppati in questo periodo dai ricercatori dello Xerox PARC (Pirulli e Card). Gli utenti navigano nel web come animali in cerca di cibo. Quando un utente guarda un link, valuta il suo “profumo informativo” (Information Scent): una stima probabilistica e imperfetta di quanto quel link lo avvicinerà al suo obiettivo. Se il profumo è forte (il testo del link corrisponde esattamente all'intento dell'utente), il click è immediato, se è debole l'utente esita o abbandona. Questa teoria ha dettato le regole del micro-copy nelle interfacce web: i link generici come “Clicca qui” sono stati sostituiti da link descrittivi, proprio per massimizzare la traccia odorosa cognitiva e ridurre l'incertezza della navigazione (Pirulli & Card, 1999).

[1.3.4] La ribellione estetica: l'era di Flash e l'interfaccia cinematografica

Tra la fine degli anni Novanta e i primi anni Duemila, si creò una frattura ideologica nel design delle interfacce web. Da una parte c'erano i sostenitori dell'usabilità (Nielsen, Krug) che predicavano layout standard, link blu sottolineati e velocità di caricamento. Dall'altra, i designer visivi che trovavano l'HTML troppo rigido e limitante. Questa tensione esplose con la diffusione di Macromedia Flash, poi diventato Adobe Flash. Flash permetteva ai designer di ignorare le regole del browser, permetteva loro di usare font non standard, animazioni vettoriali complesse, suoni e video. L'interfaccia cessava di essere una “pagina” e diventava un “filmato interattivo”. Il simbolo di quest'era è il pulsante “Skip Intro”. Molti siti Flash accoglievano l'utente con lunghe animazioni cinematografiche, loading bars elaborate, loghi che ruotavano, prima di mostrare il menu di navigazione. Mentre queste interfacce offrivano un'esperienza immersiva ed emotiva superiore, violavano quasi tutte le euristiche di usabilità. Rompevano il pulsante “indietro” del browser intrappolando l'utente nell'applicazione, non erano indicizzabili dai

[1.3.5] Usabilità come ingegneria: “Don’t Make Me Think”

motori di ricerca e ignoravano le convenzioni di navigazione standard. La caduta di Flash, accelerata dal rifiuto di Apple di supportarlo su iPhone nel 2010, segnò la vittoria dell’approccio funzionalista. Secondo tale corrente di pensiero l’interfaccia web doveva essere prima di tutto utile, accessibile e conforme agli standard, sacrificando l’espressività cinematografica in favore della prevedibilità e della compatibilità universale (Manovich, 2001; Krug, 2000).

La sintesi di questo periodo storico si trova nel lavoro di Steve Krug, il cui libro “Don’t Make Me Think” (2000) divenne la bibbia dei web designer. Krug argomentava che la risorsa più scarsa sul web non è la larghezza di banda, ma l’attenzione dell’utente. Ogni volta che un’interfaccia costringe l’utente a fermarsi per decodificare la navigazione “È un bottone o solo un’immagine?”, “Dove sono finito?”, si consuma una “riserva di pazienza” limitata. L’evoluzione dell’interfaccia web, dal caos iniziale dei primi siti personali GeoCities fino ai layout standardizzati di Google o Amazon caratterizzati dal logo a sinistra, ricerca in alto, carrello a destra, è stata un processo di riduzione del carico cognitivo attraverso la standardizzazione delle convenzioni. Abbiamo imparato collettivamente a usare il web creando un linguaggio di pattern condivisi, rendendo l’interfaccia trasparente per permettere l’accesso immediato al contenuto.

[1.4] L’Interfaccia somatica: il mobile e la NUI (2007 – 2015)

Se il paradigma WIMP aveva definito l’interazione uomo-macchina per un quarto di secolo, interponendo un dispositivo meccanico come il mouse tra l’intenzione dell’utente e l’azione digitale, l’avvento dell’era mobile ha segnato il ritorno al corpo. Questa fase, che possiamo definire dell’interfaccia somatica (dal greco soma, corpo), è caratterizzata dalla rimozione degli intermediari. Lo schermo non è più solo una superficie di rappresentazione (output), ma diventa, attraverso la tecnologia multi-touch, la superficie stessa di

[1.4.1] L’antefatto: la tirannia del pennino e il fallimento dei Tablet PC

manipolazione (input). Si realizza così il sogno della manipolazione diretta in senso letterale: l’utente non sposta più un cursore che agisce su un oggetto, ma tocca l’oggetto stesso.

Per comprendere la portata rivoluzionaria dell’interfaccia tattile capacitiva, è necessario analizzare il contesto tecnologico dei primi anni Duemila. Esistevano già dispositivi touch (come il Palm Pilot o i Pocket PC di Microsoft), ma si basavano sulla tecnologia resistiva. Questi schermi richiedevano una pressione fisica decisa e, soprattutto, l’uso di uno stilo per colpire piccoli target precisi d’interfaccia, ripresi direttamente dal concetto di desktop (pulsanti “X” per chiudere le finestre grandi pochi pixel). Bill Gates nel 2001, presentò il Microsoft Tablet PC, affermando che sarebbe diventato il computer più popolare d’America. Nonostante tale dichiarazione fallì miseramente non tanto per l’hardware in sé, ma per la sua interfaccia. Microsoft aveva semplicemente trapiantato Windows XP, progettato per il mouse e la tastiera, su uno schermo touch senza ripensare le nuove implicazioni che l’interazione touch avrebbe comportato. L’utente si trovava a dover gestire menu a tendina complessi e barre di scorrimento sottili con un pennino di plastica, generando un attrito cognitivo ed ergonomico non sostenibile. Secondo la biografia redatta da Walter Isaacson (2011), la genesi dell’iPhone nacque proprio da una reazione viscerale di Steve Jobs a questo approccio. Un ingegnere di Microsoft, durante una festa di compleanno, si vantò con Jobs dei progressi del Tablet PC e del suo stilo. Jobs, infastidito, tornò in ufficio il giorno dopo e ordinò al suo team: “facciamogli vedere come si fa davvero. Dio ci ha dato dieci stilo. Non ne inventiamo un altro”(Isaacson, 2011). Questa decisione filosofica e il rifiuto della mediazione strumentale a favore del dito nudo costrinse il team Apple a ripensare l’interfaccia da zero: i target dovevano essere grandi (almeno 44x44 pixel, la dimensione media di un

[1.4.2] La rottura del 2007 e la fisica dell'illusione



Steve Jobs presenta l'iPhone (2007). Il rifiuto del pennino a favore del multitouch capacitivo segna il passaggio all'interazione diretta con il dito nudo.

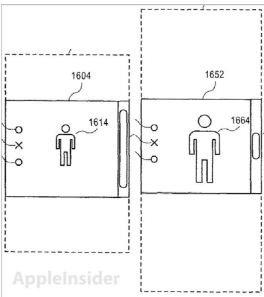
[1.4.3] Il vocabolario gestuale: pinch, swipe e la semantica del gesto

polpastrello), non potevano esistere cursori e il feedback doveva essere immediato.

Il lancio dell'iPhone nel gennaio 2007 introdusse il paradigma delle NUI (Natural User Interfaces). A differenza delle GUI, che richiedono l'apprendimento di convenzioni artificiali come il doppio clic e il tasto destro, le NUI sfruttano abilità motorie innate. Tuttavia, l'aspetto più innovativo non fu il semplice tocco, ma la fisica simulata. Nelle interfacce precedenti infatti gli oggetti digitali erano statici. Nell'interfaccia sviluppata per iOS d'altra parte gli oggetti possedevano inerzia, attrito e peso. Uno dei dettagli più sottili ma cruciali dell'interazione mobile è lo scorrimento inerziale con rimbalzo (Rubber Banding). Prima del 2007, quando l'utente scorreva una lista su un computer e arrivava alla fine lo scorrimento si arrestava bruscamente. Su iPhone, quando l'utente arriva a fine lista, il contenuto continua a muoversi leggermente staccandosi dal bordo, per poi "rimbalzare" indietro come se fosse attaccato a un elastico. Questo dettaglio, brevettato da Apple (e causa di lunghe guerre legali con Samsung), non è una decorazione estetica ma un feedback cognitivo cruciale. Esso comunica all'utente, senza usare parole o finestre di errore, che la lista è finita conferendo all'oggetto digitale una materialità fisica. Questo dettaglio inganna il cervello nel percepire i pixel come materia tangibile. Come osserva Bret Victor (2011), il design dell'interazione mobile di successo è quello che rispetta le leggi della fisica newtoniana, perché il nostro sistema sensomotorio si è evoluto per manipolare oggetti fisici, non astrazioni logiche (Victor, 2011).

L'introduzione del Multi-touch permise di creare un nuovo vocabolario gestuale. Il gesto non è più un semplice comando (come premere un tasto), ma un'azione semantica che trasforma lo spazio.

- Pinch-to-zoom: è il gesto iconico dell'era mobile. Pizzicare lo schermo per rimpicciolire o allargare le



Disegno da un brevetto Multi-touch. Il tocco multiplo ha permesso la creazione di un nuovo vocabolario gestuale (come il pinch-to-zoom) dove l'azione fisica mappa direttamente la trasformazione spaziale.

[1.4.4] L'estetica didattica: l'apogeo e il declino dello Skeuomorfismo



Confronto tra lo skeuomorfismo di iOS 6 (sinistra) e il Flat Design di iOS 7 (destra). Le texture realistiche sono servite come "rotelle" per l'apprendimento iniziale, poi rimosse una volta acquisita l'alfabetizzazione gestuale.

dita per zoomare, questo gesto è cognitivamente rilevante perché mappa un'azione fisica (comprimere/espandere) su un risultato visivo analogo eliminando l'astrazione, l'utente "afferra" lo spazio e lo manipola.

- Swipe (scorrimento): sostituisce la barra di scorrimento (scrollbar). Invece di muovere un piccolo rettangolo laterale per spostare il contenuto (azione indiretta), l'utente spinge il contenuto stesso. Bill Buxton, ricercatore di Microsoft e pioniere del multitouch, nel suo testo Multi-Touch Systems that I Have Known and Loved (2007), sottolinea come questa "fedeltà cinematografica" cambi la natura dell'errore. Nelle interfacce mouse, l'errore è spesso fatale o frustrante, nelle interfacce gestuali fluide invece, l'utente può correggere il tiro in tempo reale: se zumo troppo, restringo subito. Il ciclo di feedback è continuo (millisecondi), creando una sensazione di controllo totale ed istantaneo (Buxton, 2007).

L'era mobile ha visto anche un'evoluzione radicale dell'estetica, fondamentale per accompagnare l'apprendimento di queste nuove modalità di interazione. Nei primi anni (2007 - 2012), sotto la guida di Scott Forstall, l'interfaccia di iOS era dominata da uno skeuomorfismo estremo. L'applicazione "Note" sembrava un blocco di carta gialla strappata; il "Game Center" imitava il feltro verde di un tavolo da casinò; la bussola sembrava fatta di ottone e vetro. Molti critici di design in seguito hanno deriso questo stile come "kitsch". Tuttavia, come difeso da Donald Norman, lo skeuomorfismo aveva una funzione pedagogica fondamentale. Toccando per la prima volta uno schermo di vetro liscio, l'utente non aveva alcun feedback tattile. La texture visiva fungeva da affordance percepita. Se un pulsante sembrava lucido e tridimensionale, il cervello dell'utente suggeriva: "Questo sporge, quindi posso premerlo". Se l'agenda sembrava rilegata in pelle, suggeriva: "Questo è un oggetto personale, non un foglio di calcolo". Lo skeuomorfismo è stato le "rotelle della

[1.4.5] L’ergonomia del pollice e la “Thumb Zone”

bicicletta” per l’umanità digitale, ci ha insegnato a interagire con lo schermo touch. Una volta che l’alfabetizzazione gestuale è stata acquisita dalla massa, le rotelle sono state tolte. Nel 2013, con iOS 7, Apple, alla cui guida era salito Jony Ive, passò improvvisamente al Flat Design. Le textured e le ombre realistiche vennero rimosse, l’interfaccia divenne astratta, basata su tipografia, colore e livelli di traslucenza. Questo passaggio segnò la maturità dell’interfaccia somatica in quanto non avevamo più bisogno che il telefono fingesse di essere un oggetto analogico, avevamo accettato il dispositivo digitale come un oggetto con una propria estetica nativa e autentica (Manovich, 2013).

L’interfaccia mobile ha imposto nuovi vincoli ergonomici. A differenza del mouse, che permette di raggiungere qualsiasi pixel con lo stesso sforzo, il pollice umano ha un raggio d’azione limitato quando si impugna lo smartphone con una sola mano. Lo studio di Steven Hoober sulla “Thumb Zone” risalente al 2013 ha dimostrato che il 49% degli utenti usa il telefono con una sola mano e che certe aree dello schermo, specialmente gli angoli superiori, sono “zone ostili”, difficili da raggiungere (Hoober, 2013). Questo dato biometrico ha forzato la migrazione degli elementi di navigazione. Se nel desktop i menu sono situati in alto, nelle app mobili moderne le barre di navigazione (tab bar) sono quasi sempre in basso, all’interno della zona di comfort del pollice. L’ascesa dei “Phablet” ha accresciuto tale problema, portando alla nascita di soluzioni di interfaccia come il “Reachability” di Apple (che abbassa lo schermo a metà) o le interfacce “One UI” di Samsung, che spostano i titoli in alto, considerata area visiva, e i pulsanti in basso divenuta l’area interattiva, separando per la prima volta l’area di viewing dall’area di doing.

[1.5] La dissoluzione dell’interfaccia: Zero-UI, Spatial Computing e AI (2015 – Oggi)

[1.5.1] Zero-UI e la sfida delle interfacce vocali (VUI)



Amazon Echo (1a generazione, 2014). Un esempio radicale di design “Zero-UI”: la rimozione totale dello schermo costringe l’interazione a basarsi esclusivamente sulla voce, eliminando l’affordance visiva ma permettendo un uso ambientale.

Se la storia delle interfacce finora raccontata è stata una storia di schermi, dal fosforo verde dei terminali, al desktop luminoso, fino al vetro degli smartphone, l’ultima fase evolutiva segna un’inversione di tendenza: la dissoluzione dello schermo. Siamo entrati nell’era che il designer Golden Krishna ha provocatoriamente definito “The Best Interface is No Interface” (Krishna, 2015). Dopo aver raggiunto il punto di saturazione, in cui siamo circondati da rettangoli luminosi che richiedono costante attenzione visiva, la nuova frontiera del design mira a spostare l’interazione nel dominio dell’invisibile, dell’ambientale e del predittivo.

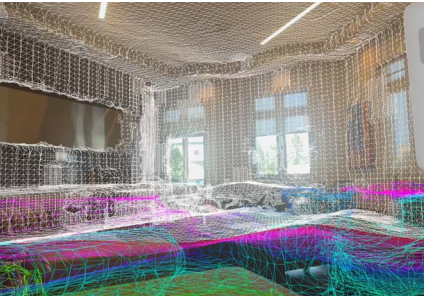
La prima manifestazione di questa dissoluzione è l’ascesa delle VUI (Voice User Interfaces), popolarizzate da assistenti intelligenti come Amazon Alexa, Apple Siri e Google Assistant. L’interfaccia vocale rappresenta il ritorno a una forma di comunicazione pre-tecnologica caratterizzata dal dialogo orale. Eliminando lo schermo si elimina la necessità di apprendere una GUI, non ci sono più menu da navigare, ma si introducono nuove, complesse sfide cognitive.

Quando Amazon lanciò l’Echo nel 2014, fu una scelta di design radicale: un cilindro nero senza schermo, senza tastiera, con solo un anello di luce. Il problema fondamentale delle VUI è la mancanza di Affordance Visiva. In una GUI, se vedo un pulsante, so che posso premerlo d’altra parte nella VUI l’utente si trova di fronte a un “golfo dell’esecuzione” (Norman, 1988) enorme: quali sono i limiti del sistema? Questo genera spesso frustrazione e la necessità di memorizzare comandi specifici, riportando ironicamente l’utente a problemi simili a quelli della Command Line Interface (CLI): l’interazione si basa sul Recall (memoria) e non sul Recognition (vista). Quello che ha reso e mantenuto il successo di questi dispositivi risiede nella loro natura di Ambient Computing: permettono l’interazione “eyes-free” e “hands-free”, mentre si cucina, mentre si guida, integrandosi nel flusso

[1.5.2] Spatial Computing: il mondo come schermo



L'interfaccia di Google Glass (2013). Un esempio di Heads-Up Display (HUD): l'informazione fluttua in un piano fisso rispetto all'occhio, creando distrazione e fallendo nell'integrazione contestuale con l'ambiente fisico.



Visualizzazione della mappatura spaziale (Spatial Mapping). Affinché l'interfaccia possa ancorarsi stabilmente al mondo fisico, il dispositivo deve ricostruire in tempo reale la geometria dell'ambiente (mesh), trasformando la realtà in dati computabili.

[1.5.3] Smart Glasses e Neural Interfaces

della vita quotidiana senza interromperlo visivamente.

Parallelamente alla voce, la realtà aumentata (AR) e la realtà mista (MR) stanno trasformando lo spazio fisico in interfaccia, dando vita al paradigma dello Spatial Computing. A differenza della realtà virtuale che isola l'utente in un mondo sintetico, lo Spatial Computing sovrappone le informazioni digitali al mondo reale. Qui, l'interfaccia non è più confinata in una cornice, il campo visivo dell'utente diventa il display. Per capire l'evoluzione dello Spatial Computing, è istruttivo confrontare due approcci. Nel 2013, Google Glass fallì non solo per problemi di privacy (il fenomeno dei "Glassholes"), ma per un errore di design della sua interfaccia. I Glass erano un Heads-Up Display (HUD) che mostravano notifiche fluttuanti in un angolo della vista distraendo l'utente dalla realtà senza realmente integrandosi con essa. Era come avere uno smartwatch incollato all'occhio. Al contrario, dispositivi moderni come Microsoft HoloLens o Apple Vision Pro utilizzano sensori di mappatura spaziale (LiDAR e SLAM - Simultaneous Localization and Mapping) per comprendere la geometria della stanza. Questo permette l'Ancoraggio Spaziale (Spatial Anchoring). L'interfaccia non "galleggia" davanti agli occhi, ma può essere "posata" su un tavolo reale o "appesa" a una parete reale diventando parte integrante della realtà e dell'ambiente stesso. Quando l'utente si muove, l'oggetto digitale rimane fisso nello spazio fisico. Questa stabilità è cruciale per il cervello perché trasforma l'elemento digitale in materia percepita. Come notano Scoble e Israel (2017), lo Spatial Computing realizza finalmente la visione di "fondere atomi e bit", rendendo l'interfaccia contestuale e situata (Scoble & Israel, 2017).

Mentre Apple con Vision Pro ha spinto l'interfaccia verso lo Spatial Computing ad alta fedeltà basato su schermi pass-through, ovvero real time video che riproducono la



Il braccialetto EMG (elettromiografia) di Meta. Il dispositivo rileva i segnali neurali che viaggiano dal cervello alla mano, permettendo un'interazione basata su micro-gesti impercettibili o sulla pura intenzione motoria.

realtà, Meta ha delineato una traiettoria alternativa che punta alla totale smaterializzazione del dispositivo attraverso due punti chiave: l'intelligenza multimodale senza schermo (Ray-Ban Meta) e l'AR olografica (Orion). L'interfaccia senza display: Meta Ray-Ban e l'AI Multimodale con l'evoluzione dei Ray-Ban Meta (2023-2024), assistiamo a un fenomeno interessante di "sottrazione". A differenza dei fallimentari Google Glass i quali cercavano di aggiungere uno schermo alla vista, i Ray-Ban rimuovono il display visivo affidandosi interamente all'audio e all'intelligenza artificiale multimodale. L'interazione si basa sul paradigma "Look and Ask": le telecamere non servono più solo a scattare foto, ma agiscono come occhi per l'AI. L'utente guarda un ingrediente e chiede: "Cosa posso cucinare con questo?". In questo scenario, l'interfaccia si trasforma in un agente contestuale sussurrato. Non vi è né GUI né menu. L'interfaccia è ridotta alla pura semantica della conversazione, situata nel contesto visivo dell'utente. Questo realizza la visione di Mark Weiser in modo inaspettato il quale sosteneva la tecnologia non fosse "calma" perché è ambientale, ma perché nascosta in un oggetto (l'occhiale da sole) che è già socialmente accettato da decenni, superando la barriera dell'imbarazzo sociale che aveva condannato i Glass e che sta penalizzando d'altro canto l'Apple Vision Pro nell'uso quotidiano (Zuckerberg, 2024). Il prototipo Orion rappresenta il culmine della ricerca sull'interfaccia ottica see-through e forse il vero punto di arrivo per Meta. A differenza dei visori VR/MR (come il Quest o il Vision Pro) che chiudono l'utente dietro a degli schermi opachi riproiettando il mondo via video, Orion utilizza proiettori Micro-LED e lenti in carburo di silicio (Silicon Carbide) per proiettare ologrammi direttamente nel campo visivo naturale dell'utente. Tuttavia, la vera rivoluzione di Orion porta non risiede nell'output (display), ma nell'input. Il problema irrisolto dell'AR è

sempre stato l'input: come cliccare su un'icona che fluttua nell'aria senza stancare le braccia (il cosiddetto "Gorilla Arm effect")? Meta ha introdotto con Orion un braccialetto basato sull'elettromiografia (EMG). Questa tecnologia non legge il movimento della mano, ma l'intenzione neurale, il braccialetto decodifica i segnali elettrici che i motoneuroni inviano dal cervello ai muscoli delle dita millisecondi prima che il movimento avvenga, o anche se il movimento è impercettibile (movimenti isometrici). L'interfaccia neurale EMG permette un'interazione "micro-gestuale": l'utente può fare "clic" semplicemente pensando di muovere il pollice, tenendo le mani in tasca o lungo i fianchi. Siamo di fronte a un cambio di paradigma ontologico, si sta passando dall'interfaccia che richiede un'azione fisica evidente come premere un tasto o toccare uno schermo, ad un'interfaccia che legge l'intenzione motoria latente. È il passo definitivo verso la fusione tra sistema nervoso e sistema computazionale, rendendo l'interazione non solo invisibile agli altri, ma quasi priva di sforzo per l'utente (Meta Reality Labs, 2024).

Questo sviluppo segna la fine della "dittatura dello schermo": l'interfaccia futura non sarà un rettangolo che teniamo in mano, ma un layer intelligente che percepisce la nostra volontà (EMG) e arricchisce la nostra visione (AR), dissolvendosi completamente nell'esperienza umana.

Se le tecnologie EMG di Meta intercettano l'intenzione motoria mentre viaggia lungo il sistema nervoso periferico, l'ultima frontiera dell'Interaction Design mira a eliminare del tutto la necessità del trasporto del segnale connettersi direttamente alla sorgente, la corteccia cerebrale. Entriamo qui nel campo delle BCI (Brain-Computer Interfaces), una tecnologia che, pur avendo radici mediche decennali (Nicoletti, 2001), ha trovato in Neuralink la sua prima incarnazione commerciale orientata al concetto di "simbiosi totale". La filosofia progettuale di Neuralink, esplicitata da Elon Musk



Il dispositivo N1 di Neuralink. I filamenti flessibili ("threads"), più sottili di un capello umano, contengono gli elettrodi per rilevare l'attività neuronale, puntando a una connessione ad alta larghezza di banda tra cervello e cloud.

nel lancio del 2017, parte da una critica radicale alle interfacce attuali (tastiere, mouse, touch, voce). Musk definisce questi metodi come "lossy" (a perdita di dati) e incredibilmente lenti. Mentre il cervello umano elabora pensieri a velocità elevatissime, l'output verso la macchina è limitato dalla velocità meccanica delle dita o delle corde vocali. Siamo, secondo questa visione, limitati dall'I/O (Input/Output). L'obiettivo dell'interfaccia neurale diretta è aumentare la "larghezza di banda" della comunicazione uomo-macchina, permettendo una fusione cognitiva che J.C.R. Licklider aveva iniziato a sognare nel 1960. Non si tratta più di usare un computer, ma di essere il computer (Musk & Neuralink, 2019). Il momento "Engelbart" delle BCI è avvenuto nel gennaio 2024 con l'impianto del chip "Telepathy" (N1) nel primo paziente umano, Noland Arbaugh, un uomo tetraplegico. L'architettura dell'interfaccia di Neuralink differisce dai precedenti sistemi medici (come il Utah Array) per la sua invisibilità e densità: 1024 elettrodi su 64 fili flessibili ("threads"), più sottili di un capello umano, inseriti da un robot chirurgico per evitare i vasi sanguigni. L'aspetto rivoluzionario dal punto di vista dell'Interaction Design è l'esperienza descritta da Arbaugh. Dopo un breve periodo di calibrazione (immaginare di muovere la mano paralizzata), il modello mentale è cambiato. Arbaugh ha descritto l'uso del cursore per giocare a scacchi o a Mario Kart come "usare la forza" di Star Wars: "Non stavo più cercando di muovere la mano. Guardavo semplicemente un punto sullo schermo e volevo che il cursore fosse lì, e succedeva." (Neuralink Livestream, 2024). Questo rappresenta l'eliminazione definitiva del golfo dell'esecuzione teorizzato da Norman. Non c'è traduzione fisica (muscolo), non c'è traduzione strumentale (mouse). L'intenzione diventa azione pura. L'inserimento di Neuralink nel panorama delle interfacce segna una biforcazione storica. Da un lato, la strada non invasiva (seguita da Meta con Orion/EMG o Apple

[1.5.4] L'ultima frontiera: Brain-Computer Interfaces (BCI) e Neuralink

[1.5.5] Oltre l'Input: l'AI agentive e la simbiosi cognitiva

con l'Eye Tracking), che mira a ridurre l'attrito mantenendo l'integrità corporea. Dall'altro, la strada Invasiva (Neuralink, Synchron), che richiede una modifica biologica dell'utente (craniotomia) per ottenere un segnale più "pulito" e veloce. Mentre oggi queste tecnologie sono confinate all'ambito clinico (ripristino delle funzioni per persone con disabilità), la roadmap di Neuralink punta esplicitamente all'utenza consumer. Ciò pone interrogativi inediti per il designer: come si progetta l'interazione per un sistema che non si può "spegnere" o rimuovere? Come si gestisce la privacy quando l'interfaccia ha accesso non alle azioni, ma ai pensieri grezzi? Con le BCI, l'interfaccia cessa di essere un "dispositivo di mediazione" esterno per diventare una componente interna dell'essere umano, realizzando e superando le visioni più estreme della cibernetica di Wiener.

Se le interfacce neurali (BCI) come Neuralink promettono di abbattere la barriera dell'input, permettendo all'intenzione umana di fluire direttamente nella macchina, resta aperto il problema della complessità dell'output. In un mondo iper-complesso, trasmettere il comando mentale "organizza una riunione" è inutile se la macchina non possiede l'autonomia per eseguirlo. Qui entra in gioco l'ultimo tassello della dissoluzione dell'interfaccia: il passaggio dal paradigma della manipolazione diretta a quello della delega. Per decenni, l'Interaction Design si è focalizzato sul potenziamento degli strumenti. Christopher Noessel, nel suo *Designing Agentive Technology* (2017), distingue nettamente tra:

- strumenti: estensioni che richiedono attenzione continua e manipolazione diretta (il mouse, Photoshop, il cursore controllato mentalmente da Noland Arbaugh).
- agenti: sistemi che estendono la volontà dell'utente lavorando in background.

In questo scenario, l'interfaccia neurale e l'AI Generativa convergono. L'hardware (Neuralink/Meta Orion) cattura l'intento grezzo ("Voglio andare a casa"), mentre il software

(l'Agente AI) gestisce l'esecuzione (calcolare il percorso, chiamare l'auto, impostare il termostato). L'interfaccia scompare non solo perché è invisibile, ma perché il bisogno di micro-interagire svanisce. L'obiettivo diventa l'Intent Prediction, il sistema anticipa il bisogno prima ancora che venga formulato consciamente. Tuttavia, questa delega totale solleva il problema della Black Box. Se l'interfaccia decide per me, come posso fidarmi? Il design delle interfacce future si sposterà quindi dalla progettazione di bottoni alla progettazione della fiducia e della spiegabilità (explainability). Sarà necessario garantire sempre un meccanismo di Human-in-the-loop, per evitare che la sensazione di "potere magico" offerta dalla simbiosi AI-cervello si trasformi in una perdita di Agency (controllo) umana.

Conclusioni

L'analisi storica condotta in queste pagine traccia una parabola chiara e inarrestabile. Siamo partiti dall'era del sacerdozio tecnico (1945-1975), dove l'uomo doveva apprendere il linguaggio oscuro della macchina (CLI). Abbiamo attraversato l'era della metafora visiva (1973-2007), in cui abbiamo costruito scrivanie virtuali per rendere la logica computazionale comprensibile (GUI). Siamo passati per l'era dell'interfaccia somatica (2007-2015), rompendo il vetro e toccando l'informazione con le nostre mani (NUI). Oggi, siamo sulla soglia della simbiosi. Con l'avvento dello Spatial Computing, delle interfacce neurali e dell'AI agentiva, l'interfaccia cessa di essere un dispositivo di mediazione esterno (un oggetto che si usa) per diventare una componente ambientale e biologica (un sistema in cui si vive). Progettare in questo nuovo paradigma non significa più disegnare schermate o flussi di navigazione, ma progettare comportamenti, etica e relazioni. L'interfaccia del futuro non sarà qualcosa che guardiamo, ma qualcosa attraverso cui guardiamo il mondo.

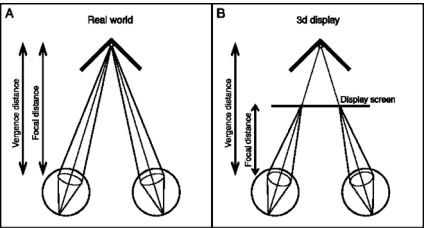
[2] Architetture dell'interazione post-schermo: spazio, corpo e contesto

Se il primo capitolo ha tracciato l'evoluzione storica dell'interfaccia come un processo di progressiva smaterializzazione dell'hardware, questo secondo capitolo si propone di analizzare la grammatica strutturale che emerge quando lo schermo cessa di essere il confine dell'esperienza digitale. L'avvento dello Spatial Computing e della realtà aumentata (AR) non rappresenta una semplice estensione della GUI (Graphical User Interface) in una terza dimensione, ma un cambio di paradigma ontologico. Come osserva Lev Manovich, ci stiamo spostando da una "poetica dello schermo", basata sulla composizione visiva all'interno di una cornice rettangolare, a una "poetica dello spazio aumentato", in cui l'informazione pervade l'ambiente fisico, sovrapponendosi all'architettura esistente.

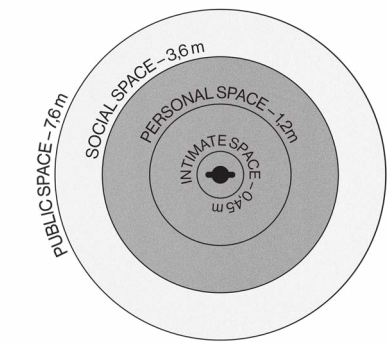
In questo nuovo contesto, il designer non è più un impaginatore di pixel, ma un architetto di esperienze. Le regole del Graphic Design, come la gestione del colore o della griglia, diventano insufficienti se non integrate con principi derivati dall'architettura, dall'ergonomia fisica e dalla prossemica. Questo capitolo esplorerà le tre dimensioni fondamentali dell'interazione post-schermo: lo spazio come contenitore, il corpo come strumento di input e il contesto come chiave interpretativa.

[2.1] La prossemica digitale: organizzare l'informazione nello spazio

[2.1.1] Le zone di interazione e il conflitto visivo



Il conflitto vergenza-accomodazione (VAC). Nei visori attuali, l'occhio deve mettere a fuoco a una distanza fissa (lo schermo), ma convergere su oggetti virtuali a distanze diverse. Questo disaccoppiamento è la causa primaria della fatica visiva in VR/AR.



Le zone prossemiche codificate da Edward T. Hall (1966). Nel design spaziale, queste distanze definiscono dove posizionare i contenuti: l'interazione diretta avviene nella zona personale (entro 1,2m), mentre i contenuti informativi risiedono nella zona sociale.

Il primo problema fondamentale che l'interfaccia spaziale deve risolvere è il posizionamento degli elementi. In un ambiente desktop tradizionale, la posizione di una finestra è determinata da coordinate cartesiane arbitrarie su un piano bidimensionale; in un ambiente spaziale, l'oggetto digitale occupa un volume reale e deve relazionarsi necessariamente con la fisiologia dell'utente e con la materialità degli oggetti circostanti. Per strutturare questa relazione, l'Interaction Design attinge alla prossemica, la disciplina fondata dall'antropologo Edward T. Hall nel 1966, che studia come l'essere umano utilizzi lo spazio come forma di comunicazione non verbale. (Hall, 1966)

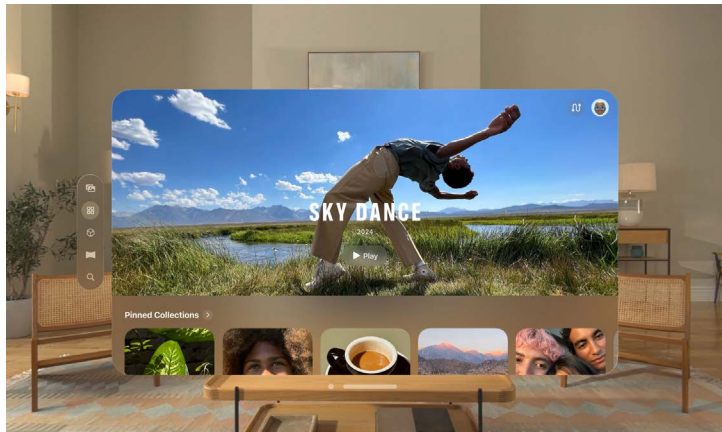
Applicando il modello di Hall alla Human-Computer Interaction (HCI), come teorizzato da ricercatori quali Saul Greenberg (2011), è possibile definire una tassonomia delle distanze digitali che garantisca leggibilità e comfort. La zona più critica è quella che Hall definisce "intima", compresa tra il corpo dell'utente e i 45 centimetri di distanza. In ambito AR, il posizionamento di contenuti persistenti in questo raggio è sconsigliato a causa di un fenomeno fisiologico noto come Vergence-Accommodation Conflict (VAC). Quando un oggetto virtuale è troppo vicino, gli occhi dell'utente devono convergere drasticamente verso il naso per allineare le immagini, ma il cristallino rimane accomodato sulla distanza focale fissa del display del visore. Questa discrepanza sensoriale genera rapidamente affaticamento visivo e nausea, rendendo la zona intima adatta esclusivamente ad avvisi rapidi o interazioni aptiche temporanee, ma interdetta alla lettura prolungata (Hoffman et al., 2008). Allontanandosi dall'utente, si entra nella zona "personale", che si estende dai 45 centimetri a circa 1,2 metri. Questa è la distanza ergonomica ottimale per la manipolazione diretta (Direct Manipulation), assimilabile allo spazio di lavoro di una scrivania fisica. È interessante notare come le linee guida di Apple

[2.1.2] Sistemi di riferimento

Vision Pro suggeriscano di posizionare le finestre delle applicazioni a una distanza apparente di circa 1,5 metri, una scelta progettuale che simula la posizione di un monitor, sufficientemente lontana da minimizzare il conflitto di vergenza, ma abbastanza vicina da mantenere una sensazione di controllo privato e accessibilità immediata (Apple, 2024). Oltre questa soglia si aprono le zone "sociale" e "pubblica", ideali per dashboard informative ambientali o oggetti tridimensionali che richiedono un'ispezione a distanza, spostando la modalità di interazione dal tatto diretto allo sguardo (gaze) o al puntamento remoto (Ray-casting).

Una volta definita la distanza, è necessario stabilire come l'interfaccia reagisca al movimento dell'utente nello spazio. Le Mixed Reality Design Guidelines di Microsoft (2019) identificano tre principali sistemi di riferimento, ognuno con profonde implicazioni cognitive. Il primo approccio è quello dell'ancoraggio alla testa (Head-Locked), dove l'interfaccia si muove solidalmente con la rotazione del capo dell'utente, simile a un Heads-Up Display (HUD) aeronautico. Sebbene utile per avvisi critici, questo sistema è cognitivamente faticoso per la lettura: ogni micro-movimento del collo sposta il testo, costringendo gli occhi a un continuo inseguimento e occludendo pericolosamente la visione del mondo reale. Un'evoluzione più ergonomica è rappresentata dall'ancoraggio al corpo (Body-Locked), utilizzato ad esempio nei menu di sistema di HoloLens 2. In questo caso, l'interfaccia segue l'utente con un movimento fluido e ritardato (lazy follow): se l'utente ruota la testa, il menu rimane inizialmente fermo per poi fluttuare gentilmente nel nuovo campo visivo. Questo rispetta la propriocezione umana, creando la sensazione di una "cintura degli attrezzi" invisibile ma sempre accessibile. Tuttavia, il paradigma dominante

[2.1.3] La gerarchia dell'asse Z e il ruolo delle ombre



[2.1.4] Il paradosso della luce: design additivo e sottrattivo

dello Spatial Computing è l'ancoraggio al mondo (World-Locked). In questo scenario, l'oggetto digitale è fissato a coordinate spaziali fisiche, come l'angolo di un tavolo reale. Questo conferisce all'oggetto digitale la proprietà psicologica della permanenza dell'oggetto: l'applicazione non risiede "nel computer", ma diventa un elemento di arredo della stanza, permettendo all'utente di costruire una mappa mentale dello spazio ibrido (Microsoft, 2019).

L'introduzione di una reale profondità (asse Z) pone infine la sfida dell'occlusione e della gerarchia visiva. Affinché un ologramma risulti credibile, deve rispettare le leggi dell'ottica fisica: se una mano reale si interpone tra l'occhio e l'oggetto virtuale, quest'ultimo deve essere correttamente nascosto. Il fallimento di questa illusione crea il cosiddetto "Ghost effect", un conflitto percettivo che distrugge l'immersione e rende difficile la valutazione delle distanze. Per mitigare questi problemi e ancorare gli oggetti alla realtà, i sistemi operativi moderni come visionOS adottano l'uso di materiali adattivi e ombre dinamiche. Invece di utilizzare superfici opache che nasconderebbero potenziali ostacoli fisici, l'interfaccia utilizza materiali vetrosi traslucidi (glassmorphism) che sfocano lo sfondo e reagiscono all'illuminazione ambientale. Un dettaglio cruciale, sottolineato nelle HIG di Apple (2024), è la proiezione di ombre virtuali sulle superfici reali. L'ombra non è un orpello estetico, ma una funzione cognitiva di "grounding": essa permette al cervello di triangolare istantaneamente la posizione dell'oggetto nello spazio tridimensionale, riducendo il carico cognitivo necessario per orientarsi e conferendo peso e presenza alla materia digitale. (Apple, 2024; Google ARCore, 2023)

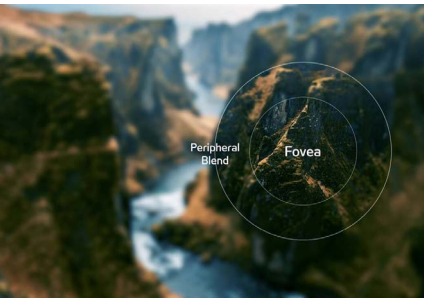
Un aspetto critico che differenzia radicalmente il design per schermi dal design spaziale è la fisica stessa della luce. Fino ad oggi, i designer hanno operato in un regime sottrattivo (stampa) o di retroilluminazione (schermi LCD/

OLED), dove è possibile visualizzare il colore nero semplicemente spegnendo i pixel o usando inchiostri scuri. Nello Spatial Computing, in particolare nei dispositivi optical see-through (come HoloLens o il prototipo Orion di Meta) che proiettano la luce su lenti trasparenti, vige la fisica del colore additivo. In questo regime, il "nero" non esiste: il nero è semplicemente trasparenza. Non è possibile proiettare un'ombra scura, si può solo aggiungere luce alla realtà. Questa limitazione fisica ha implicazioni enormi sulla leggibilità. Un testo nero su uno sfondo reale luminoso risulterebbe invisibile. Pertanto, il designer deve abbandonare le palette scure (Dark Mode) tipiche delle dashboard moderne per abbracciare contrasti basati sulla luminosità pura. Per garantire la leggibilità del testo in ambienti caotici e luminosi, diventa imperativo l'uso di "Text Plates" (sfondi volumetrici semitrasparenti o sfocati) che fungano da strato di mediazione tra il testo digitale e il mondo fisico, simulando artificialmente il contrasto che la fisica ottica non può garantire nativamente. (Billinghurst et al., 2001)

[2.2] Embodied interaction: Il corpo come cursore

Abbandonando la mediazione fisica del mouse e la superficie bidimensionale del touchscreen, l'Interaction Design entra nel territorio della Embodied Interaction. Questo concetto, formalizzato da Paul Dourish nel suo testo fondamentale "Where the Action Is" (2001), suggerisce che il significato dell'interazione non risieda nel sistema informatico in sé, ma nel modo in cui l'azione fisica dell'utente si manifesta nel mondo (Dourish, 2001). In un sistema post-schermo, il corpo umano cessa di essere un mero operatore esterno che invia comandi a una macchina, il corpo diventa il dispositivo stesso. Le mani, gli occhi e la voce si trasformano in cursori e strumenti di selezione, eliminando l'astrazione strumentale ma introducendo nuove sfide legate alla fisiologia, alla fatica e all'ambiguità dell'intenzione.

[2.2.1] Lo sguardo come intenzione



Il principio del Foveated Rendering. Sfruttando l'Eye Tracking, il sistema rende- rizza ad alta risoluzione solo la piccola porzione di schermo osservata dalla fovea, riducendo il carico computazionale nella visione periferica.

Tra tutte le modalità di input natu- rale, lo sguardo (Eye Tracking) rappresenta la forma più immediata e potente di indicazione dell'interese. Psicologicamente, guardare un oggetto è il preludio all'interazione con esso. Tuttavia, trasformare l'oc- chio in un mouse comporta la risoluzione di complessi problemi neurofisiologici. Il primo riguarda l'anatomia stessa della visione uma- na. L'occhio non percepisce la realtà in alta definizione in modo uniforme; la visione dettagliata avviene solo all'interno della fovea, un'area minuscola della retina che copre circa due gradi del campo visivo (l'equivalente della larghezza di un pollice a distanza di braccio). Tutto il resto è visione periferica, sfocata e sensibile principalmente al movimento.

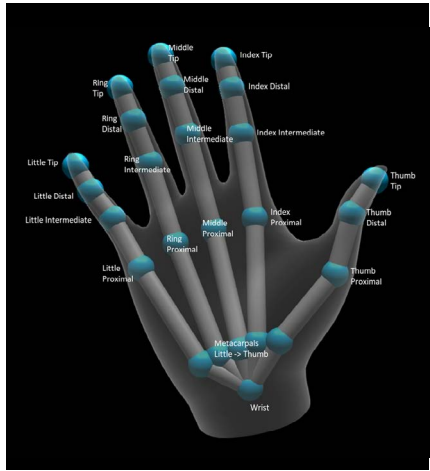
Questa caratteristica biologica ha portato allo sviluppo di tecniche di "Foveated Rendering", in cui il siste- ma computazionale concentra la potenza di calcolo solo dove l'uten- te sta guardando, sfocando il resto. Dal punto di vista dell'interfaccia, ciò significa che il designer deve progettare elementi che siano rico- noscibili anche nella visione periferica (tramite colore o sagoma) per attirare l'attenzione foveale dell'utente. La sfida più grande nell'uso dello sguardo come input è tuttavia il cosiddetto problema del "Midas Touch Problem", identifica- to da Jacob Nielsen già negli anni Novanta (Nielsen, 1993). Poiché l'occhio è un organo di percezione e non di azione, l'essere umano spo- sta lo sguardo continuamente e inconsciamente (movimenti sacca- dici) per scansionare l'ambiente. Se ogni sguardo venisse interpretato dal sistema come un "clic", l'utente attiverebbe involontariamente qual- siasi cosa osservi, rendendo l'interfaccia inutilizzabile. Per decenni, la soluzione è stata il Dwell Time (tempo di permanen- za): per cliccare, l'utente doveva fissare un bottone per un tempo pre- definito (es. 1 secondo). Questa meccanica, tuttavia, risulta innatu- rale e faticosa, trasformando l'osservazione in un esercizio di fis- sità forzata. La soluzione contemporanea, adottata

[2.2.2] Manipolazione diretta e il problema del "Gorilla Arm"

magistralmente in sistemi come Apple Vision Pro, consiste nel sepa- rare il puntamento dalla conferma: l'occhio seleziona (mira), ma è un altro segnale (come il tocco delle dita) a confermare l'azione. In que- sto paradigma, l'occhio funge da cursore balistico estremamente rapi- do, mentre la conferma manuale agisce come il "click" del mouse, risolvendo l'ambiguità tra il sempli- ce guardare e il voler agire.

Se lo sguardo definisce l'intenzione, le mani definiscono l'azione. Nello Spatial Computing, l'uso delle mani si divide in due filosofie progettuali distinte che riflettono approcci ergo- nomici opposti: la manipolazione diretta e la manipolazione indiretta. La manipolazione diretta (Direct Manipulation) è l'approccio più intuitivo: l'utente allunga il braccio e tocca l'ologramma come se fosse un oggetto fisico. Questo paradigma soddisfa il desiderio umano di tangi- bilità e offre un alto grado di precisione per compiti come preme- re pulsanti, ruotare manopole o spostare finestre. Tuttavia, questa modalità si scontra con un limite fisiologico invalicabile noto come Gorilla Arm Effect.

Il termine, coniato negli anni Ottanta durante i primi esperimenti con gli schermi touch verticali, descrive l'affaticamento muscolare che nasce quando si mantengono le braccia sollevate davanti al corpo per periodi prolungati senza suppor- to. Il cuore umano deve pompare sangue contro gravità e i muscoli deltoidi entrano rapidamente in ten- sione isometrica. Pertanto, un'interfaccia che costringa l'utente a "toccare l'aria" costantemente è destinata al fallimento per sessioni d'uso superiori ai pochi minuti. (Carmody, 1989) La risposta del design contempora- neo è la manipolazione indiretta (Indirect Manipulation). In questo modello, le mani dell'utente riman- gono comodamente in grembo o lungo i fianchi. Le telecamere del visore, dotate di un ampio campo visivo verso il basso, tracciano lo scheletro della mano e interpretano "micro-gesti". Il gesto archetipico di questa fase è il pinch (l'unione di



Visualizzazione del Hand Tracking scheletrico. Il sistema ricostruisce la posa della mano in tempo reale mappando circa 25 punti articolari (joints), permettendo l'interazione diretta senza controller fisici.

[2.2.3] La voce come acceleratore contestuale

pollice e indice). Questo gesto funge da clic astratto. Non è necessario toccare l'oggetto virtuale; basta guardarlo (puntamento oculare) e pizzicare le dita ovunque esse si trovino (azione manuale). Questo disaccoppiamento tra il luogo dell'oggetto e il luogo dell'azione rappresenta una conquista ergonomica fondamentale: permette sessioni di lavoro prolungate riducendo lo sforzo fisico quasi a zero. È interessante notare come questo approccio richieda al sistema una comprensione sofisticata della struttura della mano (Hand Tracking). Il software deve ricostruire in tempo reale la posizione di oltre 25 giunture articolari e gestire il problema dell'occlusione: quando il dorso della mano nasconde le dita alle telecamere, l'intelligenza artificiale deve calcolare e stimare la posizione delle falangi nascoste per non interrompere l'interazione (Microsoft, 2019).

Sebbene spesso trattata separatamente, la voce è parte integrante dell'interazione corporea. Nelle architetture post-schermo, la voce assume un ruolo diverso rispetto agli assistenti vocali domestici. Non è più solo uno strumento per impartire comandi ("Che tempo fa?"), ma diventa un acceleratore per operazioni complesse che richiederebbero troppi passaggi manuali. L'uso della voce nello Spatial Computing si scontra però con barriere sociali e cognitive, da un lato l'imbarazzo sociale di parlare da soli in pubblico limita l'uso dell'input vocale in contesti condivisi (uffici, mezzi pubblici) dall'altro la voce è un medium lineare e temporale, mentre lo spazio è simultaneo. Il design efficace integra la voce come modalità di "input pigro" o correttivo. Ad esempio, digitare un testo su una tastiera virtuale sospesa a mezz'aria è un'esperienza frustrante a causa della mancanza di ritorno tattile (haptic feedback): le dita attraversano i tasti senza incontrare resistenza, rallentando la digitazione. In questo scenario, la dettatura vocale diventa l'input primario per il testo (Text-to-Speech), mentre le mani sono relegate a

funzioni di editing e posizionamento del cursore. Inoltre, la voce è l'unico input capace di gestire l'astrazione. Mentre le mani possono manipolare solo ciò che è visibile (spostare questa finestra), la voce può richiamare ciò che è invisibile ("Mostrami le foto dell'anno scorso"). Questa complementarità tra la concretezza del gesto e l'astrazione della parola è alla base dei sistemi multimodali più avanzati.

Infine, l'assenza di schermi fisici pone il problema della privazione sensoriale tattile. Quando un utente tocca un pulsante in realtà aumentata non sente nulla. Questa dissonanza cognitiva riduce la fiducia nell'interazione. Per compensare, l'Interaction Design si avvale di sostituti sensoriali, principalmente audio spaziale e feedback visivo. Un "clic" sonoro nitido e localizzato spazialmente esattamente dove si trova il dito può ingannare il cervello facendogli credere di aver toccato una superficie (sinestesia audio-tattile). Parallelamente, la ricerca si sta muovendo verso tecnologie di Mid-Air Haptics (come gli ultrasuoni focalizzati che creano pressione sulla pelle) o dispositivi indossabili minimali (anelli o bracciali vibranti) per restituire al corpo la sensazione del contatto, chiudendo l'anello sensorimotorio che è alla base della nostra percezione di realtà.

[2.2.4] Il ritorno del tatto: feedback aptico fantasma

[2.3] Multimodalità: la convergenza dei sensi

Se l'analisi condotta finora ha sezionato le singole modalità di input, lo sguardo, il gesto, la voce, come canali discreti, la realtà dell'esperienza umana è intrinsecamente sinestetica e simultanea. Nella comunicazione interpersonale, noi non usiamo mai un solo canale alla volta: indichiamo mentre parliamo, moduliamo lo sguardo mentre gesticoliamo. L'architettura dell'interazione post-schermo, per risultare veramente fluida e "invisibile", deve abbandonare la logica sequenziale (prima clicco, poi scrivo) per abbracciare la multimodalità, ovvero la fusione simultanea di segnali sensoriali diversi per costruire un unico

[2.3.1] L’archetipo storico: “Put-That-There” e la deissi

intento comunicativo. La teoria alla base di questo approccio, sviluppata ampiamente da Sharon Oviatt (1999), suggerisce che l’interazione multimodale non sia una semplice somma delle parti, ma un sistema di elimina le ambiguità: dove la voce è imprecisa, il gesto chiarisce; dove il gesto è ambiguo, la voce specifica. (Oviatt, 1999)

Il fondamento teorico di tutta l’interazione multimodale contemporanea risale a un esperimento visionario condotto nel 1980 presso l’Architecture Machine Group del MIT (il precursore del Media Lab). Richard Bolt realizzò il sistema “Put-That-There”, che permetteva a un utente seduto su una sedia, in una stanza dotata di grandi schermi proiettati, di manipolare forme geometriche semplici usando la voce e il puntamento. L’importanza di questo esperimento risiede nella risoluzione del problema linguistico della deissi (dal greco deixis, indicazione). Nel linguaggio naturale, parole come “questo”, “quello”, “lì” o “qui” sono prive di significato semantico intrinseco, acquisiscono senso solo se ancorate a un contesto spaziale. Nel sistema di Bolt, se l’utente diceva “Sposta quello...”, la frase era semanticamente incompleta per il computer. Tuttavia, se contemporaneamente l’utente indicava con il braccio un cerchio blu sullo schermo, il sistema incrociava il timestamp della voce con le coordinate vettoriali del braccio (tracciato magneticamente), risolveva l’ambiguità del pronome “quello” associandolo al cerchio blu, e attendeva la destinazione. Quando l’utente completava la frase “...lì”, indicando un nuovo punto, l’azione veniva eseguita. Bolt dimostrò che la combinazione di due input imperfetti e ambigui (una frase incompleta e un gesto vago) poteva generare un comando perfettamente chiaro.

Questo paradigma, ignorato per decenni durante il dominio del mouse (che è uno strumento unimodale preciso), è tornato prepotentemente centrale con lo Spatial Computing, dove la precisione del pixel è impossibile e l’ambiguità diviene la norma (Bolt, 1980).

[2.3.2] Sinergie: lo sguardo come puntatore, la mano come conferma

Nell’era contemporanea, l’applicazione più raffinata di questo principio si trova nel paradigma di interazione di Apple Vision Pro, che fonde Eye Tracking e Hand Tracking. Come analizzato nella sezione precedente, lo sguardo è un indicatore di interesse, ma soffre del “Midas Touch” (clicca tutto ciò che vede), mentre le mani soffrono della fatica fisica se usate per puntare a lungo. Il sistema operativo visionOS risolve il conflitto creando una dipendenza gerarchica tra i due sensi, un approccio che possiamo definire “puntamento oculare, conferma manuale”.

In questo modello, gli occhi fungono da cursore balistico: l’utente guarda un’icona e il sistema la evidenzia impercettibilmente (feedback di hover), confermando che l’intenzione è stata registrata. A quel punto, le mani, che possono trovarsi comodamente in grembo in una posizione di riposo, eseguono un micro-gesto di pizzicamento (tap). Il sistema non ha bisogno di sapere dove sono le dita nello spazio, perché le coordinate dell’azione sono fornite dagli occhi. Questa sinergia crea un’illusione psicomotoria. La multimodalità maschera la complessità tecnologica: il cervello dell’utente percepisce un’unica azione fluida (“volevo aprirlo e si è aperto”), ignorando che sono stati necessari due flussi di dati biometrici separati e sincronizzati per realizzarla. L’interfaccia diventa invisibile proprio perché imita il funzionamento del sistema nervoso, dove l’attenzione visiva guida l’azione motoria in modo subcosciente (Apple, 2024).

[2.3.3] La convergenza visione – linguaggio: Multimodal AI

Una seconda declinazione della multimodalità, emersa con forza negli ultimi anni grazie ai progressi dell’intelligenza artificiale, riguarda la fusione tra visione artificiale (Computer Vision) e linguaggio naturale. Questo approccio è incarnato dai dispositivi screenless come i Ray-Ban Meta Smart Glasses. Qui la multimodalità non serve a “cliccare” su oggetti virtuali, ma a interrogare la realtà. L’utente indossa occhiali dotati di telecamere e microfoni. Quando pronuncia una

[2.3.4] Resilienza ed errore: il principio di ridondanza

frase come “guarda e dimmi cosa posso cucinare con questi ingredienti”, il sistema cattura simultaneamente l’audio (la domanda) e l’immagine soggettiva (il tavolo con il cibo).

L’AI multimodale (Large Multimodal Model) elabora i due segnali insieme. L’immagine fornisce il contesto (“vedo uova, farina, latte”) che manca alla voce; la voce fornisce l’intento (“voglio una ricetta”) che manca all’immagine. Questo rappresenta un’evoluzione del paradigma di Bolt: non è più solo “Put-That-There” (comando spaziale), ma è “Explain-This-To-Me” (comprensione semantica). In questo scenario, l’interfaccia cessa di essere un pannello di controllo per diventare un mediatore cognitivo che vede ciò che vediamo noi e ci aiuta a decodificarlo. La convergenza dei sensi artificiali (camera + microfono) potenzia i sensi naturali dell’utente, offrendo risposte uditive a stimoli visivi, chiudendo un cerchio di feedback che non richiede alcuno schermo (Meta, 2023).

Infine, è fondamentale sottolineare un vantaggio tecnico dell’approccio multimodale spesso trascurato: la robustezza. Secondo la teoria dell’interazione multimodale di Oviatt, gli utenti tendono spontaneamente a cambiare modalità quando incontrano un errore o una difficoltà ambientale. Se un ambiente è rumoroso e l’assistente vocale non capisce il comando, l’utente istintivamente tenderà a indicare o toccare lo schermo. Se le mani sono occupate o sporche, l’utente passerà alla voce. Un sistema ben progettato accetta input ridondanti: permette di eseguire la stessa operazione (es. chiudere una finestra) sia con un comando vocale (“chiudi”), sia con un gesto manuale, sia con lo sguardo prolungato. Questa flessibilità architetturale è ciò che rende l’interfaccia resiliente. In un contesto post-schermo, dove le variabili ambientali sono imprevedibili (luce variabile, rumore di fondo, movimento), affidarsi a un unico canale di input è un rischio di design inaccettabile. La convergenza dei sensi garantisce che, se un canale fallisce,

l’altro possa subentrare, mantenendo il flusso dell’interazione ininterrotto (Oviatt, 1999).

[2.4] Interazione situata e context-awareness: l’ambiente come interfaccia

Finora abbiamo analizzato come l’utente agisce sul sistema (input corporeo) e come il sistema occupa lo spazio (prossemica). Tuttavia, l’architettura post-schermo introduce una terza variabile critica: il contesto. Nelle interfacce tradizionali (desktop/mobile), il contesto è quasi irrilevante: un foglio di Word funziona allo stesso modo sia che l’utente sia a casa, sia che sia in treno. Nello Spatial Computing e nell’Ubiquitous Computing, il contesto determina il funzionamento stesso dell’interfaccia. Un paio di occhiali AR non possono mostrare le stesse notifiche mentre l’utente guida un’auto o mentre conversa con un amico. L’interfaccia deve diventare Context-Aware, trasformandosi da strumento passivo a sistema adattivo che “comprende” la situazione in cui è immerso.

[2.4.1] Definire il contesto: oltre le coordinate GPS



Semantic scene understanding. Attraverso l’uso di sensori di profondità (LiDAR) e AI, il sistema segmenta lo spazio fisico riconoscendo pavimenti, pareti e mobili per integrarvi coerentemente gli oggetti digitali.

Per progettare un sistema consapevole, bisogna prima definire cosa sia il “contesto”. Nel loro paper fondamentale Context-Aware Computing Applications, Bill Schilit, Norman Adams e Roy Want (1994) definiscono il contesto non come un singolo dato, ma come una matrice complessa composta da tre vettori:

- dove sei (posizione semantica): non bastano le coordinate GPS (latitudine/longitudine). Il sistema deve capire il significato del luogo. Essere in “cucina” attiva le interfacce per le ricette; essere in “camera da letto” attiva la modalità “Non disturbare”. Questa comprensione semantica richiede che la macchina classifichi l’architettura circostante.
- con chi sei (contesto sociale): il sistema deve rilevare la presenza di altre persone. Se i sensori rilevano un interlocutore nel campo visivo, l’interfaccia AR deve entrare in modalità “sociale”, riducendo le notifiche o rendendole traslucide per non interrompere il contatto visivo.
- cosa stai facendo (riconoscimento dell’attività): grazie all’accelerometro e all’AI, il dispositivo deve

[2.4.2] Interazione implicita ed esplicita

distinguere se l’utente è seduto, sta camminando o sta correndo. Un’interfaccia leggibile da fermi diventa illeggibile e pericolosa in movimento. Questa matrice trasforma l’interfaccia in un’entità liquida: non esiste più un layout fisso, ma un layout che muta in tempo reale in risposta alle variabili ambientali.

La context-awareness abilita un nuovo paradigma definito da Albrecht Schmidt (2000) come interazione implicita.

- interazione esplicita: l’utente ordina al sistema di fare qualcosa (es. “Siri, accendi la luce”).
- interazione implicita: il sistema agisce basandosi su un cambiamento di contesto, senza comando diretto. Un esempio concreto nell’Interaction Design post-schermo è la gestione della privacy dei contenuti. Immaginiamo che l’utente stia leggendo un’email riservata proiettata virtualmente sul muro del suo ufficio tramite un visore AR. Se un collega entra nella stanza (cambio di contesto sociale), il sistema dovrebbe oscurare automaticamente l’email o trasformarla in un’icona generica. Questa azione non è stata comandata dall’utente, ma è stata dedotta dal contesto. Progettare queste interazioni richiede un delicato equilibrio: se il sistema è troppo proattivo, diventa invadente e fastidioso (come nel caso “Clippy”); se è troppo passivo, perde di utilità. Il designer deve mappare le soglie di confidenza (confidence thresholds) dell’AI: il sistema deve agire autonomamente solo quando è sicuro al 99% dell’intento contestuale.

Un caso studio pratico di design context-aware riguarda l’adattamento dell’interfaccia alla mobilità. Le linee guida di Google per la realtà aumentata (ARCore) e gli studi di ergonomia cognitiva evidenziano come il carico cognitivo aumenti drasticamente durante la deambulazione. Camminare richiede attenzione visiva al terreno e agli ostacoli. Pertanto, un’interfaccia situata deve cambiare stato in base all’attività motoria:

[2.4.3] La modalità “walking” e “sitting” nella realtà aumetata

[2.4.4] Intelligenza distribuita

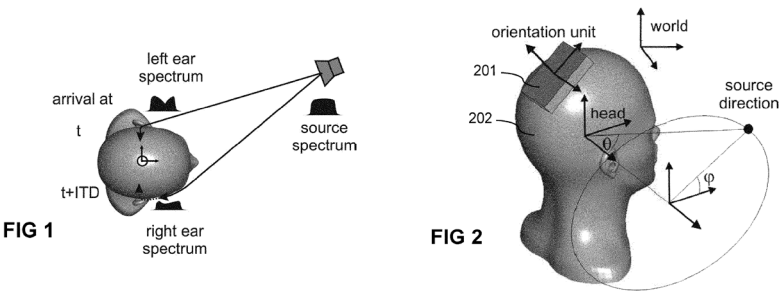
- (High Density): l’utente è fermo, il sistema può mostrare testi densi, finestre multiple e controlli complessi, l’interfaccia è immersiva.
- (Low Density): non appena l’accelerometro rileva il passo, l’interfaccia deve “evaporare”. Le finestre si chiudono o si ancorano alle pareti (World-locked), lasciando libero il campo visivo centrale (Heads-up). Le informazioni diventano minimali (solo frecce direzionali o avvisi di pericolo). Questo adattamento dinamico è fondamentale per la sicurezza. Un’interfaccia che non reagisce al contesto motorio non è solo “scomoda”, è fisicamente pericolosa per l’utente, violando il principio primo del design etico.

Infine, la context-awareness spinge l’interazione oltre il singolo dispositivo indossabile, verso l’ecosistema dell’intelligenza ambientale (Ambient Intelligence o AmI). Come teorizzato da Adam Greenfield nel suo libro *Everyware* (Greenfield, 2006), l’interfaccia post-schermo non risiede solo negli occhiali, ma nel dialogo tra gli occhiali e l’ambiente. Immaginiamo uno scenario di Spatial Computing domestico: l’utente entra in salotto indossando il visore. Il visore non deve calcolare tutto da solo; comunica con le lampadine intelligenti (IoT) per abbassare le luci fisiche quando l’utente guarda un film virtuale. Comunica con il termostato per regolare la temperatura. In questo scenario, l’interfaccia diventa ubiqua e distribuita. Non c’è un centro focale. L’utente interagisce con la casa come un’unica entità senziente. Il designer deve quindi progettare non solo l’interfaccia grafica (UI), ma l’orchestrazione (Orchestration) dei vari dispositivi. Il contesto funge da direttore d’orchestra che sincronizza l’output di schermi, luci, suoni e temperature per creare un’esperienza coerente.

[2.5] L’architettura del suono: audio spaziale e sonificazione

Spesso, nello studio delle interfacce visive, si trascura il fatto che l’esperienza spaziale è per il 50% acustica. Nel design post-schermo, dove gli oggetti visivi possono essere fuori

[2.5.1] HRTF e la psicoacustica della presenza



Rappresentazione della HRTF (Head-Related Transfer Function). La forma anatomica della testa e dei padiglioni auricolari modifica le frequenze sonore, fornendo al cervello gli indizi necessari per localizzare la sorgente nello spazio 3D.

[2.5.2] Sonificazione dell'interfaccia: UI sounds

[2.6] Inclusive design: accessibilità nello Spatial Computing

dal campo visivo (dietro l'utente), l'audio diventa il principale strumento di orientamento e realismo.

L'audio spaziale si basa sulla tecnologia HRTF (Head-Related Transfer Function). Questa funzione matematica simula il modo in cui la forma dell'orecchio umano, la testa e le spalle modificano le onde sonore in arrivo. Applicando filtri HRTF in tempo reale, un sistema come gli AirPods Pro o i visori VR può ingannare il cervello facendogli percepire un suono come proveniente da un punto preciso nello spazio 3D (es. "in alto a destra, a 2 metri"). Per l'Interaction Design, questo è un game-changer. Permette di creare ancora sonore. Se l'utente riceve una notifica visiva fuori dal campo visivo, un suono spazializzato attira la sua attenzione in quella direzione. L'audio agisce come un "puntatore invisibile", guidando lo sguardo e riducendo il tempo di ricerca visiva. (Begault, 1994)

Oltre alla localizzazione, c'è il tema della sonificazione. Poiché gli oggetti olografici non hanno peso, il suono deve conferire loro massa. Il sound design delle interfacce spaziali deve progettare suoni per:

- texture: che rumore fa una finestra virtuale quando viene spostata? Se suona come vetro pesante, l'utente la muoverà lentamente. Se suona come carta ("fruscio"), la muoverà velocemente.
- prossimità: il volume deve cambiare man mano che la mano dell'utente si avvicina all'oggetto interattivo, fornendo un radar acustico che compensa la mancanza di tatto.

Un capitolo sulle architetture dell'interazione non può dirsi completo senza affrontare il tema dell'accessibilità. Le interfacce post-schermo, basandosi fortemente su sensi specifici (vista, udito, motricità fine), rischiano di escludere enormi fette di popolazione. Come si progetta la realtà aumentata per un non vedente o per chi ha disabilità motorie?

[2.6.1] Sostituzione sensoriale

Il principio chiave è la sostituzione sensoriale. Se l'output primario è visivo (AR), il sistema deve offrire un output parallelo uditivo o tattile per chi non vede.

- Screen Reader spaziale: nei sistemi tradizionali (come VoiceOver su iPhone), il lettore legge linearmente ciò che è sullo schermo. Nello spazio 3D, il lettore deve diventare un "esploratore". Apple ha introdotto funzionalità che descrivono la stanza: "C'è una porta a 3 metri a ore 12". L'interfaccia diventa una protesi percettiva che narra il mondo fisico.
- feedback aptico per l'udito: per gli utenti non udenti, le notifiche sonore spaziali vengono tradotte in pattern di vibrazione sui controller o sul visore stesso, indicando la direzione dell'evento.

[2.6.2] Input assistivi e Dwell Control

Per utenti con disabilità motorie che non possono gesticolare (pinch) o mantenere la testa stabile, le interfacce spaziali implementano il Dwell Control (controllo a permanenza) o input alternativi.

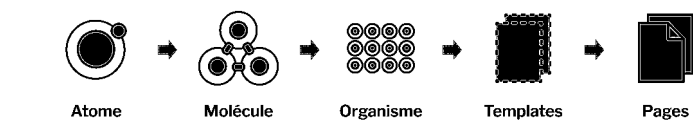
- Head Tracking alternativo: usare il movimento degli occhi o anche suoni della bocca ("Pop", "Click") per simulare i gesti delle mani.
- pointer assistivo: invece di richiedere la precisione del raggio laser, l'interfaccia espande l'area di influenza del cursore (Gravity Well), "attirandolo" verso i bottoni vicini per facilitare la selezione a chi ha tremori alle mani. Analizzare questi aspetti dimostra che l'architettura post-schermo, se ben progettata, può essere più inclusiva del mondo reale, offrendo "super-sensi" adattivi a chi ne ha bisogno (Microsoft Design, 2016).

[3] Verso l'interfaccia liquida: Generative UI e adattamento cognitivo

Se il primo capitolo ha esplorato la storia della smaterializzazione dell'hardware e il secondo ha definito le regole fisiche dello spazio e del corpo, il terzo capitolo si pone l'obiettivo di indagare il futuro del software. Ci troviamo di fronte a un paradosso progettuale: le tecnologie di visualizzazione sono diventate tridimensionali e contestuali (AR/VR), ma le metodologie di design sono ancora ancorate a metafore bidimensionali statiche (la "pagina", la "schermata", la "griglia"). L'ipotesi che guiderà questo capitolo è che l'interfaccia del futuro non sarà "disegnata" nel senso tradizionale del termine, ovvero definita a priori in ogni suo pixel, ma sarà generata in tempo reale da sistemi di intelligenza artificiale, adattandosi fluidamente al contesto, all'utente e all'ambiente. Attraverso l'analisi della letteratura recente sui sistemi adattivi (Adaptive User Interfaces) e sul design generativo, delineeremo i contorni di quella che possiamo definire interfaccia liquida. Infine, proporremo un set di linee guida uristiche speculative, intese non come dogmi, ma come principi orientativi per i designer che dovranno navigare questa transizione generazionale.

[3.1] La crisi del layout statico: oltre la metafora del contenitore

[3.1.1] La fine della “pagina” e l’atomizzazione del contenuto



Il paradigma dominante nel Digital Design degli ultimi quindici anni è stato il Responsive Web Design (Marcotte, 2010). Questa metodologia ha permesso alle interfacce di adattarsi a schermi diversi (smartphone, tablet, desktop) basandosi su un principio fondamentale: il contenitore. L’interfaccia è concepita come una scatola rettangolare (il viewport) all’interno della quale i contenuti si riorganizzano fluidamente. Nello Spatial Computing, tuttavia, questo paradigma crolla. Non esiste più un contenitore. Non esiste un viewport rettangolare che delimita il mondo; il viewport è il campo visivo dell’utente, che è continuo, sferico e privo di bordi netti.

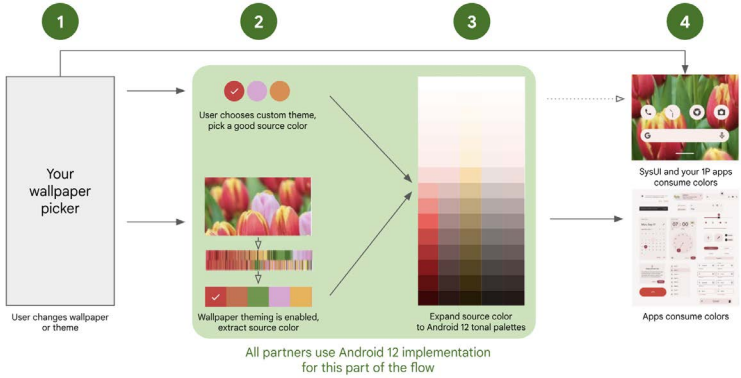
In un ambiente di realtà mista, importare il concetto di “pagina” o “finestra 2D” è una soluzione di comodo (skeuomorfismo) destinata a essere superata. Come evidenziato dai ricercatori di Google Material Design nelle recenti esplorazioni sulle interfacce adattive, il futuro appartiene a un design Object-Oriented (orientato all’oggetto) piuttosto che Screen-Oriented (orientato allo schermo). Ciò implica una radicale atomizzazione dell’interfaccia.

Prendiamo a prestito la metodologia dell’Atomic Design di Brad Frost (2016), che scompone le interfacce in atomi, molecole e organismi. Nel web design classico, questi atomi si aggregano per formare una “pagina”. Nello spazio 3D, gli atomi devono poter sopravvivere da soli, fluttuando nell’ambiente senza bisogno di una cornice di supporto. Un “atomo” potrebbe essere un bottone di riproduzione musicale che l’utente “incolla” sul tavolo; una “molecola” potrebbe essere un widget meteo ancorato alla finestra reale. La sfida progettuale, dunque, non è più impaginare questi elementi all’interno di una griglia rettangolare, ma dotare ogni singolo componente di una propria intelligenza spaziale autonoma. Ogni componente deve sapere come comportarsi se viene posizionato su un muro, su un tavolo o se viene lasciato fluttuare a mezz’aria. Non progettiamo più quadri, progettiamo sciami di oggetti intelligenti.

[3.1.2] Il paradosso della consistenza visiva

Questa esplosione del layout solleva un problema critico di consistenza. Se ogni elemento è libero, come si mantiene l’unità visiva e funzionale dell’applicazione? La letteratura accademica sulle Adaptive User Interfaces (AUI) (Gajos et al., 2010) suggerisce che la consistenza non debba più essere visiva (tutti i bottoni sono blu e allineati a sinistra), ma comportamentale. In un contesto liquido, un’interfaccia potrebbe apparire radicalmente diversa a due utenti distinti o in due momenti diversi della giornata. Mentre l’utente è seduto al divano, la notifica è una finestra espansa ricca di testo e immagini (High Fidelity). Mentre l’utente cammina verso la cucina, la stessa notifica si condensa in una piccola sfera fluttuante che segue l’utente (Low Fidelity), per poi trasformarsi in un sussurro audio se l’utente distoglie lo sguardo. L’elemento grafico è mutato completamente, ma la funzione (comunicare il messaggio) e il comportamento (seguire l’utente) sono rimasti consistenti. Questa capacità di metamorfosi richiede un approccio al design non più basato su template statici, ma su sistemi di regole dinamiche.

[3.1.3] Material You e la personalizzazione algoritmica



Un precursore di questa tendenza è rintracciabile nel linguaggio visivo Material You introdotto da Google nel 2021. Sebbene confinato agli schermi smartphone, questo sistema introduce un concetto nuovo: il designer non sceglie i colori. L’algoritmo estrae i colori dallo sfondo scelto dall’utente e rigenera l’intera palette dell’interfaccia in tempo reale. Estendendo questo concetto allo Spatial Computing, possiamo ipotizzare interfacce che non solo cambiano colore, ma cambiano forma e dimensione basandosi sulla scansione dell’ambiente. Se il sensore LiDAR rileva una stanza arredata in stile minimalista e con luci fredde, l’interfaccia generativa potrebbe adottare font sottili e materiali vetrosi. Se rileva una stanza caotica e colorata, potrebbe optare per sfondi opachi ad alto contrasto per garantire la leggibilità. In questo scenario, il designer umano perde il controllo

assoluto sull’output finale (pixel perfect), ma guadagna il ruolo di “regista” che definisce i vincoli entro cui l’AI genererà l’esperienza.

[3.2] Generative UI: l’intelligenza artificiale come co-designer in tempo reale

[3.2.1] Dal design imperativo al design dichiarativo (Intent-Based)

Se l’atomizzazione del contenuto risolve il problema della flessibilità spaziale, rimane aperta la questione della produzione. In un ecosistema di Spatial Computing dove il contesto muta ogni secondo, è umanamente impossibile per un team di designer pre-disegnare ogni possibile variazione dell’interfaccia per ogni possibile scenario d’uso. La risposta a questa impossibilità di scala risiede nella Generative UI (GenUI), un paradigma emergente in cui l’intelligenza artificiale non si limita a generare testi o immagini, ma genera l’interfaccia stessa.

Per comprendere la portata di questo cambiamento, dobbiamo analizzare l’evoluzione della logica di interazione. Fino ad oggi, il design delle interfacce ha seguito un approccio imperativo: il designer stabilisce percorsi rigidi (User Flows) e schermate fisse. L’utente deve navigare attraverso questi percorsi predefiniti per raggiungere il suo obiettivo. Se l’app di prenotazione viaggi non ha una schermata per “viaggi economici con cane al seguito”, l’utente deve arrangiarsi incrociando filtri standard. Con l’avvento dei Large Language Models (LLM) capaci di scrivere codice (come GPT-4 o modelli specializzati in React/SwiftUI), stiamo transitando verso un Design Dichiarativo o basato sull’intento (Intent-Based). In questo scenario, l’interfaccia non esiste a priori. Esiste solo una richiesta dell’utente (“voglio organizzare un weekend a Parigi ma piove”). Il sistema interpreta l’intento e, invece di rimandare l’utente a una lista generica di risultati, genera in tempo reale una micro-interfaccia (un widget, una dashboard) specifica per quella richiesta, che potrebbe contenere un grafico comparativo dei musei al coperto e una lista di hotel. Come osservato recentemente da Jakob Nielsen (2023), questo rappresenta il primo vero nuovo

[3.2.2] Interfacce effimere e “Zero-Shot Generation”

[3.2.3] Il rischio dell’allucinazione dell’interfaccia

paradigma di UI in sessant’anni: passiamo dal comanda-e-controlla (l’utente guida la macchina) alla specifica-dell’intento (l’utente dice cosa vuole, la macchina costruisce la strada).

La manifestazione tecnica di questo concetto è la cosiddetta Zero-Shot UI Generation. Il termine “Zero-Shot” in ambito AI indica la capacità di un modello di eseguire un compito per cui non è stato specificamente addestrato. Applicato al design, significa che il sistema può renderizzare un componente grafico che non è mai stato disegnato da un umano. Immaginiamo un assistente AR per la manutenzione industriale. L’operaio guarda un motore di un modello vecchio e sconosciuto. Un sistema classico non saprebbe cosa mostrare. Un sistema Generative UI, analizzando il manuale tecnico PDF e la visione della telecamera, può generare al volo etichette, pulsanti di “check” e frecce direzionali posizionate sui bulloni corretti. Questa interfaccia è effimera: viene creata nel momento del bisogno (just-in-time) e distrutta subito dopo l’uso. Questo riduce drasticamente il “debito tecnico” e il peso cognitivo: l’utente non deve imparare a navigare menu complessi pieni di opzioni inutili, perché l’unica interfaccia che vede è quella strettamente necessaria per il compito attuale. Come teorizzato da Vercel nelle loro esplorazioni sull’AI SDK (2024), l’interfaccia diventa un “flusso di coscienza” visivo tra uomo e macchina.

Tuttavia, delegare la costruzione dell’interfaccia a un algoritmo probabilistico introduce un rischio inedito: l’allucinazione di interfaccia (UI hallucination). Sappiamo che gli LLM possono inventare fatti; analogamente, possono inventare interfacce ingannevoli o non funzionanti (“dark patterns involontari”). Un sistema potrebbe generare un pulsante “cancella” che non è collegato a nessuna funzione reale, o visualizzare dati finanziari in un grafico che ne distorce visivamente le proporzioni per errore di calcolo. Questo impone una nuova

responsabilità al designer umano: non più creatore di schermate, ma creatore di guardrail (barriere di sicurezza). Il ruolo del designer si sposta verso la definizione di Design Systems rigidi e vincolanti che l'AI deve usare come “mattoncini LEGO”. L'AI può decidere come assemblare i mattoncini per rispondere alla richiesta dell'utente, ma non può inventare nuovi mattoncini o violare le regole di accessibilità e contrasto definite dal sistema. La “Generative UI” deve quindi essere sempre un sistema ibrido: creatività algoritmica imbrigliata in una griglia etica e funzionale umana. A margine delle considerazioni sulla sicurezza, esiste un ostacolo tecnico che separa l'attuale stato dell'arte dalla visione di una Generative UI fluida: la latenza. Esiste una discrepanza fondamentale tra i tempi dell'intelligenza artificiale e i tempi della realtà estesa (XR). Per evitare il malessere fisico (Motion Sickness), un sistema AR deve reagire ai movimenti della testa dell'utente in meno di 20 millisecondi (la cosiddetta Motion-to-Photon latency). Al contrario, i Large Language Models attuali richiedono dai secondi a decine di secondi per generare codice e interfacce complesse. Questa frizione temporale rende impossibile la generazione sincrona (mentre l'utente guarda). La soluzione progettuale risiede dunque nel Pre-fetching predittivo: il sistema deve indovinare l'intenzione dell'utente e pre-generare l'interfaccia nella cache prima che l'utente compia l'azione. Se l'anticipazione fallisce, il sistema deve mostrare immediatamente un'interfaccia “scheletro” (Skeleton UI) standard precaricata, per poi idratarla con i contenuti generati dall'AI in un secondo momento. Il design dell'attesa diventa quindi cruciale quanto il design dell'interfaccia stessa.

Se la Generative UI risolve il problema della forma dell'interfaccia, l'Anticipatory Design si propone di risolvere il problema del tempo e dell'attenzione. Nel paradigma tradizionale dell'interazione

uomo-macchina, il flusso è sempre stato avviato dall'utente: l'uomo sente un bisogno, formula una richiesta, naviga un menu, seleziona un'opzione e attende un risultato. Questo processo, per quanto familiare, è intrinsecamente inefficiente. Ogni passaggio rappresenta un momento di attrito cognitivo, un “costo” mentale che l'utente deve pagare per ottenere valore. In un ecosistema di Spatial Computing, dove le informazioni sono potenzialmente infinite e l'utente è spesso in movimento o impegnato in altre attività fisiche, questo costo diventa insostenibile. L'ipotesi che avanziamo in questa sezione è che l'interfaccia post-schermo debba necessariamente evolvere da un modello reattivo (che attende l'input) a un modello proattivo (che previene il bisogno).

Il fondamento teorico di questo approccio risiede nella psicologia comportamentale, specificamente nel concetto di Decision Fatigue e nel paradosso della scelta teorizzato da Barry Schwartz (2004). Schwartz dimostra che, contrariamente al senso comune, aumentare le opzioni disponibili non aumenta la libertà dell'utente, ma genera ansia, paralisi e insoddisfazione. Applicando questo principio al design delle interfacce AR, ci troviamo di fronte a un rischio concreto: se l'utente, indossando un visore, venisse bombardato da tutte le possibili interazioni disponibili in una stanza (accendere la TV, leggere le email, controllare il meteo, vedere le recensioni dei libri sullo scaffale), il suo sistema cognitivo andrebbe in sovraccarico immediato (Cognitive Overload). L'Anticipatory Design, termine reso popolare da Aaron Shapiro (2015), propone una soluzione radicale: eliminare la scelta. Un sistema perfettamente anticipatorio non chiede all'utente “Cosa vuoi ascoltare?”, ma fa partire la playlist giusta nel momento in cui l'utente entra in doccia, basandosi sulla routine storica, sull'orario e sui dati biometrici. Nello spazio aumentato, questo si traduce nella scomparsa dei menu di navigazione. L'interfaccia non è

[3.3] Anticipatory Design: la riduzione dell'attrito e il rischio della bolla

[3.3.2] La “Filter Bubble” spaziale e la perdita di agency

più una mappa da esplorare, ma un servizio che appare just-in-time. Se l’utente si siede alla scrivania, compaiono automaticamente gli strumenti di lavoro; se si sposta sul divano, questi svaniscono per lasciare posto ai controlli domotici o di intrattenimento. La migliore interfaccia, in questo paradigma, è quella che ha già preso la decisione corretta per conto dell’utente, eliminando il divario temporale tra l’insorgere del desiderio e la sua soddisfazione.

Tuttavia, l’implementazione massiva di algoritmi predittivi solleva questioni etiche di primaria importanza che il designer non può ignorare. Se il sistema decide cosa mostrarmi basandosi su ciò che pensa io voglia vedere, si crea inevitabilmente una Filter Bubble, un concetto introdotto da Eli Pariser (2011) per il web, ma che nello Spatial Computing assume connotati letterali e fisici. Nel web, la bolla ci nasconde notizie discordanti dalle nostre opinioni; nella realtà aumentata, la bolla potrebbe nasconderci pezzi di realtà.

Immaginiamo un sistema anticipatorio che, per ridurre il carico cognitivo, evidenzia solo i negozi che corrispondono ai gusti d’acquisto dell’utente e oscura o desatura gli altri. Sebbene efficiente dal punto di vista utilitaristico, questo design riduce drasticamente la serendipità, la possibilità di scoprire l’inatteso, e limita l’Agency (l’agire intenzionale) dell’essere umano. L’utente smette di essere un esploratore del mondo e diventa un passeggero passivo di una realtà curata algoritmicamente. La critica speculativa che emerge è che un design troppo anticipatorio rischi di infantilizzare l’utente. Se l’interfaccia rimuove ogni attrito, rimuove anche ogni opportunità di scelta consapevole. Il rischio è passare da un’interfaccia “User-Centric” a un’interfaccia “System-Centric”, dove l’efficienza computazionale prevale sull’autonomia umana.

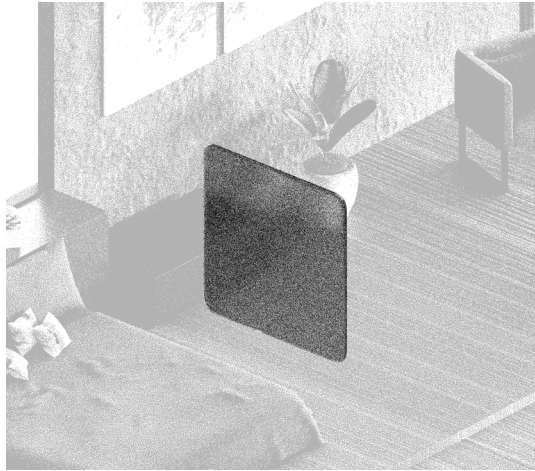
[3.3.3] Verso un’anticipazione trasparente: explainable interfaces

Per mitigare questi rischi senza rinunciare ai benefici dell’anticipazione, è necessario introdurre nel design principi di negoziazione ed explainability. L’interfaccia non deve agire come una “scatola nera” magica, ma deve comunicare le sue intenzioni prima di eseguirle o mentre le propone. Questo approccio è definito XAI (Explainable Artificial Intelligence) applicato alla UI. In termini pratici di design, ciò significa che quando un sistema genera un’interfaccia o suggerisce un’azione anticipatoria, deve fornire sottili indizi sul “perché”. Ad esempio, se un paio di occhiali smart fa apparire un percorso di navigazione senza che l’utente l’abbia chiesto, dovrebbe accompagnarlo con una micro-copy o un’icona che segnali: “basato sul tuo appuntamento in calendario tra 20 minuti”. Questa trasparenza restituisce il controllo all’utente: trasforma l’anticipazione da un’imposizione occulta a un suggerimento informato. L’obiettivo del design post-schermo non deve essere l’automazione totale dell’esperienza umana, ma il potenziamento delle capacità decisionali. L’anticipazione deve servire a liberare la mente dai compiti di basso livello (configurare il Wi-Fi, cercare un file) per permettere all’utente di concentrarsi su decisioni di alto livello, mantenendo sempre una via d’uscita (“Opt-out”) chiara e accessibile per disattivare il pilota automatico e riprendere i comandi manuali.

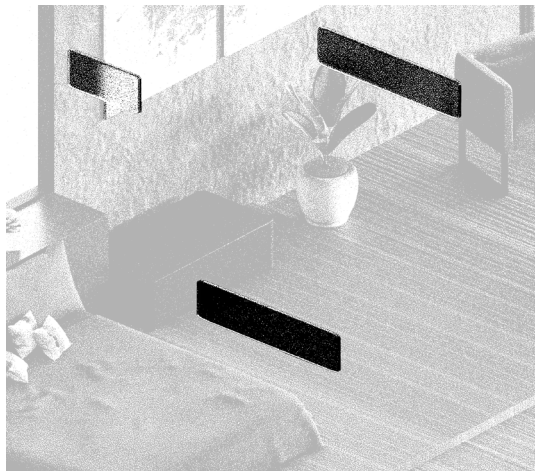
[3.4] Guidelines per l'interfaccia post-schermo: un framework euristico

Alla luce delle trasformazioni analizzate, dalla crisi del layout statico all'avvento dell'intelligenza generativa e anticipatoria, appare evidente che le euristiche di usabilità tradizionali, codificate da Jakob Nielsen negli anni Novanta per il paradigma WIMP (Windows, Icons, Menus, Pointer), non siano più sufficienti a coprire la complessità dello Spatial Computing. Sebbene principi come la "visibilità dello stato del sistema" rimangano validi, la natura volumetrica e agentiva delle nuove interfacce richiede un'evoluzione del quadro normativo. In questa sezione conclusiva, si propone un set di cinque principi guida speculativi, intesi non come regole dogmatiche, ma come orientamenti progettuali per il design di interfacce liquide, contestuali e guidate dall'AI.

1. Principio di coerenza fisica e ancoraggio



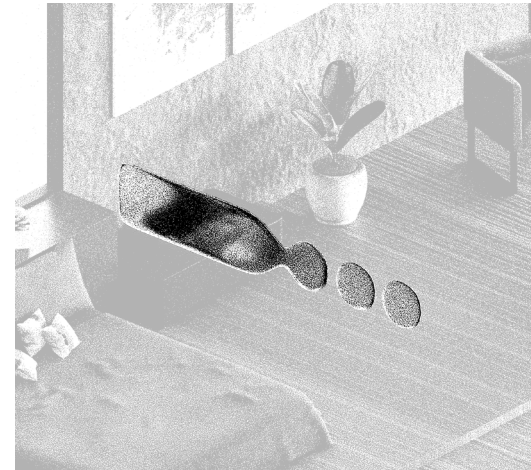
2. Principio di minima intrusione cognitiva (negative space as reality)



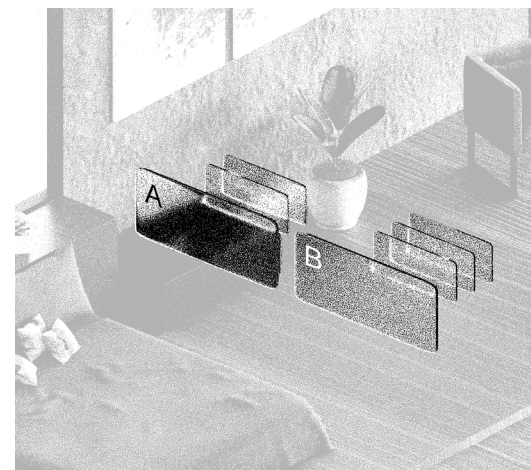
Il primo principio fondamentale riguarda la relazione tra l'oggetto digitale e le leggi della fisica. In un'interfaccia 2D, la "fisica" è metaforica (le finestre si sovrappongono, le ombre simulano la profondità). Nello spazio aumentato, la fisica deve essere percepita come reale per garantire la sospensione dell'incredulità e l'usabilità. Un'interfaccia post-schermo deve rispettare il principio di coerenza fisica: gli oggetti virtuali devono reagire alla luce ambientale, proiettare ombre coerenti sulle superfici reali e rispettare l'occlusione (nascondersi dietro agli oggetti fisici). Un menu che "fluttua" senza proiettare ombra o che attraversa un tavolo rompe la propriocezione e disorienta l'utente, rendendo difficile valutare la distanza per l'interazione manuale. Il design deve quindi privilegiare materiali che simulano la materia (come il vetro smerigliato o la plastica) rispetto a texture puramente digitali, e deve garantire che ogni elemento abbia un chiaro ancoraggio spaziale. L'utente deve sempre capire se l'oggetto è "appeso" al muro, "appoggiato" al tavolo o "fluttuante" rispetto al proprio corpo. Senza questa chiarezza gravitazionale, l'interfaccia diventa rumore visivo caotico.

Il secondo principio affronta il problema del sovraccarico informativo. Nel design per schermi (smartphone), l'obiettivo è massimizzare l'uso dei pixel: ogni millimetro di schermo è prezioso. Nel design spaziale, vige la regola opposta: la risorsa più preziosa è la realtà stessa. Le linee guida per l'interfaccia liquida devono adottare il paradigma del "negative space as reality": lo spazio vuoto non è bianco, è il mondo reale dell'utente. L'interfaccia deve occupare la minima porzione possibile del campo visivo (Field of View), appearing solo quando strettamente necessario e dissolvendosi immediatamente dopo l'uso. Questo principio suggerisce di evitare grandi pannelli opachi che ostruiscono la vista, favorendo invece componenti frammentati e traslucidi. Inoltre, impone una gerarchia dell'attenzione

3. Principio di adattabilità semantica (liquidità)



4. Principio di trasparenza dell'intento (explainability)



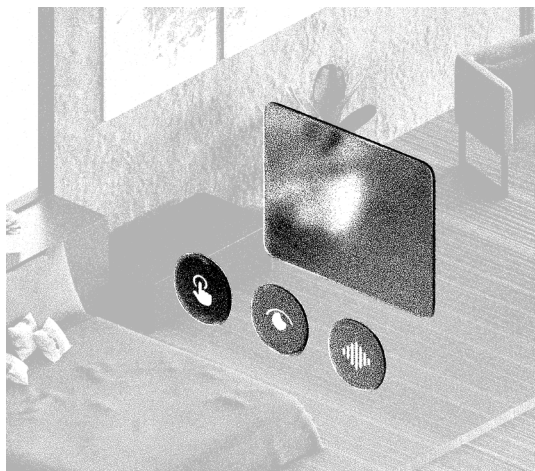
rigorosa: le notifiche non urgenti devono essere posizionate nella visione periferica o comunicate tramite segnali ambientali (suono o luce), lasciando la fovea libera per l'interazione con il mondo fisico e sociale. Un design etico post-schermo è un design che, per la maggior parte del tempo, non si vede.

Come affrontato nella sezione sulla Generative UI, l'interfaccia non può avere un'unica forma. Il terzo principio stabilisce che l'interfaccia debba possedere adattabilità semantica: deve cambiare non solo dimensione, ma struttura e modalità in base al contesto. Il designer non deve progettare una "schermata", ma definire gli stati della materia dell'interfaccia:

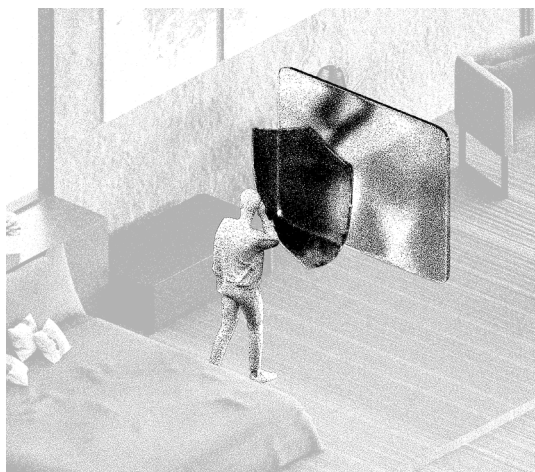
- stato gassoso (ambientale): quando l'utente cammina o non è focalizzato, l'interfaccia è diffusa, composta da suoni o luci periferiche.
- stato liquido (assistivo): quando l'utente compie un'azione, l'interfaccia si condensa in strumenti fluttuanti che seguono l'utente (Body-locked).
- stato solido (immersivo): solo quando l'utente si siede e decide di lavorare, l'interfaccia si cristallizza in finestre rigide e complesse (World-locked). Questa transizione fluida tra stati deve avvenire automaticamente, guidata dai sensori e dall'AI, senza richiedere micro-gestione da parte dell'utente.

Considerando il ruolo crescente degli algoritmi predittivi, il quarto principio è di natura etica e fiduciaria. Se l'interfaccia agisce in anticipo (Anticipatory Design), deve sempre esplicitare il "perché". Il principio di trasparenza dell'intento impone che ogni azione generata automaticamente dall'AI (come suggerire un percorso o evidenziare un oggetto) sia accompagnata da un significante che ne spieghi l'origine. L'utente non deve mai chiedersi "Perché vedo questo?". Inoltre, deve essere sempre garantita una via di fuga chiara (Kill Switch o Override), l'utente deve poter annullare l'azione dell'agente intelligente con un gesto

5. Principio di ridondanza sensoriale (input agnostico)



6. Principio di sovranità cognitiva (neuro-privacy)



semplice e immediato. La sensazione di controllo è inversamente proporzionale all'ansia tecnologica; in un mondo in cui la realtà è mediata, mantenere la percezione di Agency è fondamentale per l'accettazione del sistema.

Infine, l'accessibilità e l'ergonomia impongono il principio della ridondanza sensoriale. Non è ammissibile progettare interazioni che si basino su un unico canale di input (solo mani o solo voce), poiché le condizioni ambientali nello Spatial Computing sono imprevedibili. Il design deve essere input agnostico: ogni azione critica deve poter essere eseguita indifferentemente con lo sguardo, con le mani o con la voce. Se ho le mani occupate, devo poter chiudere una finestra con la voce. Se sono in un ambiente rumoroso, devo poterlo fare con un cenno del capo. Questa ridondanza non è un "di più", ma è l'essenza della robustezza del sistema. Un'interfaccia che funziona solo in condizioni ideali (stanza silenziosa, utente seduto) è un'interfaccia fallimentare. Il sistema deve accogliere l'imperfezione umana e ambientale, offrendo molteplici strade per raggiungere lo stesso obiettivo.

Con l'introduzione di input neurali diretti (come l'EMG di Meta) e interfacce cervello-computer (BCI), il confine ultimo della privacy viene valicato. Le guidelines future devono incorporare il principio di sovranità cognitiva, ispirato ai neuro-rights proposti dalla fondazione di Rafael Yuste. L'interfaccia non deve mai estrarre, analizzare o archiviare dati biometrici neurali senza un consenso esplicito, granulare e revocabile per ogni singola sessione. In termini di design, questo significa evitare i "Dark Patterns neurali": l'utente non deve essere spinto a cedere i propri dati mentali per sbloccare funzionalità essenziali. Deve esserci una chiara separazione visiva tra la modalità "lettura passiva" (dove il sistema legge solo i gesti volontari) e la modalità "lettura intima" (dove il sistema accede a stati emotivi o intenzioni latenti). Un indicatore di

stato inequivocabile (es. un led virtuale rosso sempre visibile) deve segnalare all'utente quando i sensori neurali sono attivi, garantendo il diritto fondamentale alla disconnessione mentale.

Conclusioni

Con la definizione di questi principi, chiudiamo l'analisi teorica e strutturale. Abbiamo attraversato la crisi del design statico, abbracciato le potenzialità generative dell'AI e delineato i rischi etici dell'anticipazione. L'interfaccia del futuro, così come emerge da questa trattazione, non è più un artefatto grafico disegnato su una tela, ma un ecosistema comportamentale. Il compito del designer si sposta dalla definizione dell'aspetto (lo styling) alla definizione delle regole di ingaggio (il comportamento). Le guidelines qui proposte rappresentano un primo tentativo di mappare questo territorio inesplorato, consapevoli che, come ogni cartografia di frontiera, saranno soggette a continue revisioni man mano che la tecnologia, e la nostra relazione simbiotica con essa, evolverà.

[4] Validazione sperimentale: scenari di design fiction

Dopo aver dissezionato l'evoluzione storica dell'interfaccia [1], analizzato i vincoli fisiologici dello spazio tridimensionale [2] e formalizzato le linee guida euristiche per un design post-schermo [3], emerge la necessità di una validazione empirica. Tuttavia, progettare per lo Spatial Computing e l'intelligenza artificiale generativa pone una sfida metodologica: come si può testare l'efficacia di un'interfaccia che ancora non esiste nella sua forma matura? Come si valuta l'ergonomia di un sistema che promette di essere invisibile?

La teoria dell'Interaction Design ci insegna che un'interfaccia non vive nel vuoto asettico di un laboratorio, ma si scontra quotidianamente con la complessità disordinata della vita reale: la fretta del mattino, l'affollamento di una metropolitana, la dinamica sociale di una riunione. È in questi interstizi di attrito che la tecnologia fallisce o trionfa.

Pertanto, questo capitolo abbandona l'analisi teorica per abbracciare la metodologia del Design Fiction. Attraverso la costruzione di "prototipi diegetici" (Kirby, 2010; Sterling, 2009), non ci limiteremo a immaginare il futuro, ma lo simuleremo.

Verranno presentati tre scenari applicativi situati in un orizzonte temporale prossimo (2030), progettati come stress-test per le sei linee guida formulate in questa tesi.

L'obiettivo di questa narrazione progettuale non è la previsione tecnologica, ma la dimostrazione che un'interfaccia spaziale ben progettata cessa di essere un dispositivo da utilizzare per diventare un ambiente da abitare.

[4.1] Metodologia: il prototipo diegetico come strumento di verifica

Al fine di evitare la mera speculazione futurologica, si è scelto di adottare la metodologia del Design Fiction. Tale approccio, formalizzato nell’ambito della ricerca sull’Interaction Design da autori quali Bruce Sterling (2009) e Julian Bleecker (2009), non mira alla predizione tecnologica, bensì alla costruzione di “prototipi diegetici”. Questi scenari fungono da laboratori narrativi in cui le tecnologie emergenti vengono situate in contesti quotidiani, permettendo al ricercatore di osservare le frizioni, le implicazioni etiche e le dinamiche sociali che emergono dall’uso prolungato. I tre scenari seguenti, definiti rispettivamente domestico (Ambient Computing), mobilità (Urban Privacy) e professionale (Collaborative Agency), simulano l’interazione di un utente tipo (Soggetto A) con un ecosistema di Spatial Computing maturo (orizzonte temporale 2030). Ogni scenario è strutturato in tre fasi: definizione del contesto ambientale, descrizione dell’interazione (protocollo d’uso) e analisi critica basata sulle sei linee guida formulate in questa tesi: coerenza fisica, minima intrusione, adattabilità semantica, trasparenza dell’intento, ridondanza sensoriale e sovranità cognitiva.

[4.2] Scenario A l’ambiente domestico e la “Calm Technology”

Lo scenario si svolge in un ambiente residenziale privato durante la fase di transizione sonno-veglia (ore 07:00). Il problema progettuale centrale riguarda la gestione del carico cognitivo iniziale. Nei paradigmi attuali (smartphone-centrici), l’utente è soggetto a un’immediata esposizione a flussi informativi ad alta densità (notifiche, email, social media) appena sveglia, generando un picco di cortisolo non fisiologico. L’obiettivo dello scenario è dimostrare come un’interfaccia spaziale possa mediare questa transizione applicando i principi della Calm Technology teorizzati da Mark Weiser.

[4.2.1] Protocollo di interazione (descrizione diegetica)

Il sistema di domotica integrata, rilevando tramite sensori biometrici la conclusione del ciclo di sonno REM del Soggetto A, avvia una sequenza di risveglio graduale, modulando l’apertura delle schermature solari fisiche. Il Soggetto A indossa il dispositivo di visualizzazione (occhiali AR a guida d’onda leggera). Il sistema operativo si avvia in modalità low information density. Il campo visivo (FOV) centrale rimane completamente libero da elementi grafici. L’interfaccia si manifesta esclusivamente nella visione periferica e ambientale: una luminescenza diffusa proiettata sulla parete frontale, renderizzata con una texture che imita la carta da parati sottostante, informa l’utente delle condizioni meteorologiche attuali tramite codici cromatici astratti, senza l’uso di testo o numeri. Nel momento in cui il Soggetto A transita dalla zona notte alla zona cucina, l’interfaccia esegue una migrazione spaziale (Spatial Handoff). Le informazioni, prima ancorate alla parete della camera, svaniscono per ricomparire, con maggiore dettaglio, sulla superficie del frigorifero. L’interazione diventa attiva solo quando il Soggetto A focalizza l’attenzione sulla macchina del caffè. Il sistema di Computer Vision, riconoscendo l’oggetto e rilevando un’anomalia di manutenzione, sovrappone un indicatore di stato pulsante (affordance aumentata) direttamente sul componente fisico interessato. Mantenendo lo sguardo sull’indicatore per 1,5 secondi (Dwell Time), l’interfaccia si espande (“unfolding”), proiettando una guida tridimensionale passo-passo per la manutenzione, ancorata rigidamente alla geometria della macchina. Applicazione della minima intrusione cognitiva. L’interfaccia dimostra una rigorosa applicazione del principio di Progressive Disclosure. Invece di presentare l’intera dashboard informativa al risveglio, il sistema dosa la quantità di dati in base allo stato psicofisiologico dell’utente. L’informazione è

	presente ma non richiede attenzione focale (“attunement”), rimanendo alla periferia della percezione finché non è attivamente ricercata. Coerenza fisica e occlusione. La proiezione delle informazioni meteorologiche sulla parete valida la guideline della coerenza fisica. Il render non appare come una finestra fluttuante (che creerebbe una dissonanza cognitiva nota come “effetto sticker”), ma rispetta le proprietà illuminotecniche della stanza. L’uso di ombre virtuali e la simulazione della texture muraria ingannano il sistema percettivo, facilitando l’accettazione dell’oggetto digitale come parte dell’arredo fisico. Adattabilità semantica. Il comportamento del sistema evidenzia una comprensione contestuale avanzata. L’interfaccia muta la sua natura ontologica in base alla localizzazione: è “ambientale/estetica” nella camera da letto (zona di riposo) e diventa “strumentale/operativa” in cucina (zona di azione). Questa fluidità dimostra il superamento del concetto di “App” chiusa a favore di un’intelligenza di sistema diffusa.
--	--

[4.3] Scenario B mobilità urbana e la bolla di privacy

[4.3.1] Protocollo di interazione (descrizione diegetica)

Lo scenario si colloca in un vagone della metropolitana ad alta densità di occupazione (ore 08:30). Il focus progettuale verte sulla dicotomia tra la necessità di produttività in mobilità e la tutela della privacy in spazi pubblici. Il paradigma tradizionale (laptop/tablet) espone i dati sensibili al rischio di “Visual Hacking” (lettura da parte di terzi) e richiede posture ergonomicamente svantaggiose.

Il Soggetto A, seduto in posizione ristretta, attiva la modalità workstation. Il visore genera uno schermo virtuale curvo equivalente a 30 pollici, posizionato a una distanza ergonomica di 60 cm. Il sistema di sensori ambientali (LiDAR e telecamere laterali) rileva la presenza di altri passeggeri all’interno della zona prossemica “intima” (< 45 cm) e “personale” (< 1,2 m). In risposta, il sistema attiva autonomamente il protocollo di privacy ottica.

[4.3.2] Analisi critica e validazione

Dal punto di vista soggettivo dell’utente, i contenuti sono nitidi e leggibili. Tuttavia, per qualsiasi osservatore esterno o telecamera di sorveglianza che tentasse di intercettare il riflesso sulle lenti o la proiezione, il contenuto risulterebbe offuscato o sostituito da pattern di rumore bianco, grazie alla tecnologia di polarizzazione selettiva delle lenti a guida d’onda.

Per interagire con il documento (scrolling e selezione), il Soggetto A adotta una postura socialmente neutra: braccia conserte o mani nelle tasche. Un dispositivo indossabile al polso basato su elettromiografia (EMG) rileva i segnali elettrici dei motoneuroni diretti alle dita. Il Soggetto A esegue micro-movimenti di frizione tra pollice e indice, invisibili all’esterno. Il sistema traduce questi impulsi neurali in comandi di scorrimento fluido sull’interfaccia virtuale. La selezione del testo avviene tramite la sincronizzazione di Eye Tracking (puntamento grossolano) e conferma EMG (click).

Sovranità cognitiva e privacy contestuale. Questo scenario rappresenta l’applicazione estrema della guideline sulla sovranità cognitiva. La privacy cessa di essere una mera impostazione software (password) per diventare una proprietà fisica dello spazio percepito. Il sistema crea una “bolla di realtà privata” (rivate reality bubble), permettendo all’utente di manipolare dati sensibili in pubblico senza compromettere la sicurezza, risolvendo uno dei paradossi fondamentali dell’Ubiquitous Computing.

Ridondanza sensoriale (Input Agnostico). La validazione della ridondanza sensoriale è qui cruciale per l’usabilità sociale. In un contesto in cui i comandi vocali violerebbero la privacy e i gesti ampi (Mid-air gestures) violerebbero lo spazio sociale altrui, il sistema permette una degradazione elegante verso modalità di input a bassissima visibilità (EMG).

L’interfaccia è “agnostica” rispetto al canale di input: accetta comandi vocali quando l’utente è solo, e passa automaticamente ai micro-gesti quando rileva un contesto sociale

affollato. Ergonomia cognitiva. L'uso combinato di sguardo e micro-gesti riduce drasticamente il Gorilla Arm Effect (affaticamento degli arti superiori), permettendo sessioni di lavoro prolungate in mobilità che sarebbero impossibili con interfacce touch verticali o gestuali pure.

Lo scenario si svolge in uno studio di progettazione architettonica. Tre utenti collaborano sincronicamente attorno a un tavolo fisico vuoto. L'obiettivo è dimostrare come la Generative UI e l'intelligenza artificiale possano integrarsi nel flusso di lavoro non come strumenti passivi, ma come agenti collaborativi trasparenti, superando il problema della "Black Box".

Al centro del tavolo, visibile a tutti i partecipanti tramite visori MR, è ancorato un modello olografico ad alta fedeltà di un complesso residenziale (ancoraggio World-Locked condiviso). Il Soggetto A identifica una criticità nell'illuminazione della facciata Nord. Invece di manipolare manualmente i vertici del modello o navigare menu complessi, attiva l'Agente AI vocale: "genera tre varianti per la facciata che ottimizzino l'ingresso di luce naturale, mantenendo il budget invariato". Il sistema non modifica il modello principale istantaneamente, azione che creerebbe disorientamento nel team. Al contrario, l'AI genera tre istanze semitrasparenti ("Ghost Objects") che fluttuano in orbita attorno al modello centrale. Ogni istanza è accompagnata da una D\data card volumetrica che esplicita i parametri decisionali: Variante alpha: aumento superfici vetrate (+25% luce, +15% costo termico). Variante beta: inserimento lucernari strutturali (+10% luce, costo invariato). Variante gamma: riorientamento moduli (+30% luce, modifica strutturale radicale). I partecipanti possono camminare attorno alle varianti ("6 Degrees of Freedom"), ispezionandole da

diverse angolazioni. Il Soggetto A indica fisicamente la variante beta. Grazie alla sincronizzazione spaziale a bassa latenza, tutti i partecipanti vedono un raggio visuale partire dalla mano del Soggetto A, eliminando l'ambiguità deittica. Confermato il comando, l'AI fonde la variante scelta nel modello principale, aggiornando la fisica e l'illuminazione in tempo reale.

Trasparenza dell'intento e Explainable AI (XAI). Lo scenario valida la guideline della trasparenza dell'intento. L'interfaccia evita l'esecuzione opaca dei comandi. Visualizzando le opzioni come ipotesi distinte accompagnate dai dati che hanno portato alla loro generazione (trade-off costi/benefici), il sistema mantiene l'essere umano al centro del ciclo decisionale (Human-in-the-loop). L'interfaccia diventa uno strumento di negoziazione tra l'intento umano e la capacità generativa della macchina. Coerenza fisica multi-utente. La tecnologia di Spatial Anchoring condiviso permette di replicare le dinamiche della collaborazione fisica. La capacità di indicare, girare intorno e guardare lo stesso oggetto digitale dal proprio punto di vista fisico ripristina la prossemica collaborativa naturale, che viene persa nelle interfacce 2D (dove ogni utente guarda il proprio schermo privato). Design per la deissi. L'interazione multimodale (voce + puntamento) risolve le ambiguità linguistiche tipiche del lavoro di gruppo. Il sistema comprende che il pronome "questa" pronunciato dal Soggetto A si riferisce alla variante beta perché incrocia il dato audio con il vettore di puntamento della mano e la direzione dello sguardo, realizzando una fusione sensoriale completa. L'analisi di questi tre scenari di Design Fiction conferma che il passaggio dall'interfaccia bidimensionale allo Spatial Computing non è una mera evoluzione estetica, ma un cambio di paradigma ontologico.

[4.4] Scenario C ambito professionale e Agency collaborativa

[4.4.1] Protocollo di interazione (descrizione diegetica)

[4.4.2] Analisi critica e validazione

[4.5] Sintesi dei risultati

Le linee guida proposte in questa tesi si sono dimostrate utili a risolvere le frizioni emergenti: senza coerenza fisica, l'oggetto digitale viene percepito come un'intrusione fastidiosa (Scenario A). Senza ridondanza sensoriale e sovranità cognitiva, l'uso in pubblico diventa socialmente inaccettabile (Scenario B). Senza trasparenza dell'intento, l'AI generativa rischia di alienare l'utente riducendone l'agenzia (Scenario C).

In conclusione, la validazione tramite scenari suggerisce che l'interfaccia del futuro cesserà di essere un "luogo" dove si va per eseguire compiti (il desktop, la app), per diventare una "membrana" intelligente, invisibile e contestuale, che si adatta fluidamente alle intenzioni dell'utente e alla natura dell'ambiente fisico circostante.

[Outro]

Nel 1960, J.C.R. Licklider profetizzò l'avvento di una “simbiosi uomo-computer”, un rapporto di mutua dipendenza in cui la macchina non fosse più un mero strumento da comandare, ma un partner capace di estendere le facoltà cognitive umane. Ripercorrendo il cammino tracciato in questa tesi, appare evidente che siamo giunti alla soglia di realizzazione di quella visione. L'analisi condotta ha dimostrato che la storia dell'interfaccia è, in ultima istanza, la storia della rimozione degli intermediari. Dalle schede perforate al mouse, dal touch screen agli occhiali intelligenti, ogni passo avanti ha ridotto la distanza fisica e cognitiva tra il pensiero umano e l'azione digitale. Con lo Spatial Computing e le tecnologie generative discusse nella tesi, questa distanza si sta azzerando: il mondo stesso diventa l'interfaccia e l'intenzione diventa l'input. Tuttavia, come emerso nella parte speculativa del lavoro, questa “invisibilità” tecnologica comporta rischi inediti. La dissoluzione del layout statico a favore di interfacce liquide e anticipatorie sposta il potere decisionale dall'utente all'algoritmo. Se l'interfaccia decide cosa mostrarci prima ancora che lo chiediamo, il confine tra assistenza e manipolazione diventa labile. Inoltre, l'ingresso dei sensori nel dominio della biometria neurale pone l'urgenza di integrare nel design principi di “sovranità cognitiva”, affinché la simbiosi non si trasformi in parassitismo dati-centrico. In conclusione, il ruolo del designer nel paradigma post-schermo subisce una mutazione profonda. Non siamo più chiamati a disegnare rettangoli perfetti o icone accattivanti, ma a progettare sistemi di regole, comportamenti ed etiche. L'interfaccia del futuro non sarà un oggetto da guardare, ma un ambiente da abitare; sarà fluida, multimodale. La sfida per la prossima generazione di progettisti sarà garantire che questa magia rimanga al servizio dell'uomo, potenziandone le capacità senza mai oscurarne la volontà. La fine dello schermo non è la fine del design, ma l'inizio della sua fase più umana e complessa.

Apple Inc. (2024). Human interface guidelines: Designing for visionOS. Apple Developer.
<https://developer.apple.com/design/human-interface-guidelines/>

Bardini, T. (2000). Bootstrapping: Douglas Engelbart, coevolution, and the origins of personal computing. Stanford University Press.

Begault, D. R. (1994). 3-D sound for virtual reality and multimedia. AP Professional.

Berners-Lee, T. (1999). Weaving the Web: The original design and ultimate destiny of the World Wide Web. HarperCollins.

Billinghurst, M., Kato, H., & Poupyrev, I. (2001). The MagicBook: A transitional AR interface. Computers & Graphics, 25(5), 745–753.

Bleecker, J. (2009). Design fiction: A short essay on design, science, fact and fiction. Near Future Laboratory.

Bolt, R. A. (1980). “Put-that-there”: Voice and gesture at the graphics interface. ACM SIGGRAPH Computer Graphics, 14(3), 262–270.

Bush, V. (1945). As we may think. The Atlantic Monthly, 176(1), 101–108.

Buxton, B. (2007). Multi-touch systems that I have known and loved. Microsoft Research.
<http://www.billbuxton.com/multitouchOverview.html>

Campbell-Kelly, M., & Aspray, W. (1996). Computer: A history of the information machine. Basic Books.

Card, S. K., Moran, T. P., & Newell, A. (1983). The psychology of human-computer interaction. Lawrence Erlbaum Associates.

Carmody, M. A. (1989). Touch screen fatigue. Human Factors Society.

Ceruzzi, P. E. (2003). A history of modern computing. MIT Press.

Cheng, K. (2023). The end of the GUI: How AI transforms design. Rosenfeld Media.

Dourish, P. (2001). Where the action is: The foundations of embodied interaction. MIT Press.

Engelbart, D. C. (1968). A research center for augmenting human intellect. AFIPS Fall Joint Computer Conference.

Frost, B. (2016). Atomic design. Brad Frost Web.
<https://atomicdesign.bradfrost.com/>

Gajos, K. Z., Weld, D. S., & Wobbrock, J. O. (2010). Automatically generating personalized user interfaces with Supple. Artificial Intelligence, 174(12-13), 910–950.

Gladwell, M. (2011, November 14). Creation myth: Xerox PARC, Apple, and the truth about innovation. The New Yorker.

Google. (2023). ARCore design guidelines. Google Developers.
<https://developers.google.com/ar/design>

Google Design. (2021). Material You: Designing for expressiveness and personalization. Google Design Blog.

Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., & Wang, M. (2011). Proxemic interactions: The new ubicomp? Interactions, 18(1), 42–50.

Greenfield, A. (2006). Everywhere: The dawning age of ubiquitous computing. New Riders.

Hall, E. T. (1966). The hidden dimension. Doubleday.

Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. Journal of Vision, 8(3), 33.

Hoober, S. (2013). How do users really hold mobile devices? UX Matters.

Isaacson, W. (2011). Steve Jobs. Simon & Schuster.

Jo, A. (2023). Explainable AI in user experience design. Springer.

Kare, S. (2019). Susan Kare icons. Kare Prints.

- Kirby, D. A. (2010). The future is now: Diegetic prototypes and the production of visible futures. *Social Studies of Science*, 40(1), 41–70.
- Krishna, G. (2015). The best interface is no interface: The simple path to brilliant technology. New Riders.
- Krug, S. (2000). Don't make me think: A common sense approach to web usability. New Riders.
- LaValle, S. M. (2023). Virtual reality. Cambridge University Press.
- Licklider, J. C. R. (1960). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, 1, 4–11.
- Manovich, L. (2001). The language of new media. MIT Press.
- Manovich, L. (2013). Software takes command. Bloomsbury Academic.
- Marcotte, E. (2010). Responsive web design. *A List Apart*, 306.
- Meta. (2023). Ray-Ban Meta smart glasses: Multimodal AI capabilities. Meta Newsroom.
- Meta Reality Labs. (2024). Inside the lab: Building the future of AR. Meta.
- Microsoft. (2019). Mixed reality design guidelines: Interactions. Microsoft Learn. <https://learn.microsoft.com/en-us/windows/mixed-reality/design/>
- Microsoft Design. (2016). Inclusive design toolkit. Microsoft. <https://inclusive.microsoft.design/>
- Musk, E., & Neuralink. (2019). An integrated brain-machine interface platform with thousands of channels. *Journal of Medical Internet Research*, 21(10), e16194.
- Neuralink. (2024). Livestream demonstration with Noland Arbaugh. X.
- Nicolelis, M. A. (2001). Actions from thoughts. *Nature*, 409(6818), 403–407.
- Nielsen, J. (1993). Usability engineering. Morgan Kaufmann.
- Nielsen, J. (1999). Designing web usability: The practice of simplicity. New Riders.
- Nielsen, J. (2023). AI: The first new UI paradigm in 60 years. Nielsen Norman Group. <https://www.nngroup.com/articles/ai-paradigm/>
- Noessel, C. (2017). Designing agentive technology: AI that works for people. Rosenfeld Media.
- Norman, D. A. (1988). The design of everyday things. Basic Books.
- Norman, D. A. (2010). Living with complexity. MIT Press.
- OpenAI. (2023). GPT-4 system card. OpenAI.
- Oviatt, S. (1999). Ten myths of multimodal interaction. *Communications of the ACM*, 42(11), 74–81.
- Pariser, E. (2011). The filter bubble: What the Internet is hiding from you. Penguin Press.
- Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, 106(4), 643–675.
- Rashid, J. (2024). Generative UI and the future of interaction. Vercel Blog.
- Rosenfeld, L., & Morville, P. (1998). Information architecture for the World Wide Web. O'Reilly Media.
- Saffer, D. (2013). Microinteractions: Designing with details. O'Reilly Media.
- Schilit, B., Adams, N., & Want, R. (1994). Context-aware computing applications. IEEE Workshop on Mobile Computing Systems and Applications.
- Schmidt, A. (2000). Implicit human computer interaction through context. *Personal Technologies*, 4(2), 191–201.
- Schwartz, B. (2004). The paradox of choice: Why more is less. Harper Perennial.
- Scoble, R., & Israel, S. (2017). The fourth transformation: How augmented reality & artificial intelligence will change everything. Patrick Brewster Press.
- Shapiro, A. (2015). The next big thing in design is less choice. Fast Company.
- Smith, D. C., Irby, C., Kimball, R., Verplank, B., & Harslem, E. (1982). Designing the Star user interface. *Byte*, 7(4), 242–282.
- Stephenson, N. (1999). In the beginning... was the command line. Avon Books.
- Sterling, B. (2009). Design fiction. *Interactions*, 16(3), 20–24.
- Tegenbos, J., & Beyien, W. (2023). Generative UI: A new era of design. *International Journal of Human-Computer Interaction*.
- Tesler, L. (1981). The Smalltalk environment. *Byte*, 6(8), 90–147.
- Tognazzini, B. (1992). Tog on interface. Addison-Wesley.
- Victor, B. (2011). A brief rant on the future of interaction design. Worrydream. <http://worrydream.com/ABriefRantOnTheFutureOfInteractionDesign/>
- Weiser, M. (1991). The computer for the 21st century. *Scientific American*, 265(3), 94–104.
- Wiener, N. (1948). Cybernetics: Or control and communication in the animal and the machine. MIT Press.
- Yuste, R., Goering, S., Bi, G., Carmena, J. M., Carter, A., Fins, J. J., ... & Wolpaw, J. (2017). Four ethical priorities for neurotechnologies and AI. *Nature*, 551(7679), 159–163.
- Zuckerberg, M. (2024). Meta Connect Keynote. Meta.

[Ringraziamenti]

Grazie Rebecca per avermi indicato la via; per la presenza, l'ascolto, la pazienza, l'amore.
Grazie alla mia famiglia per l'affetto; per aver sempre creduto in me e nelle mie ambizioni.
Grazie ai miei amici per avermi fatto sentire a casa; per aver reso il tragitto leggero e ricco di ricordi indelebili.
Grazie Milano per tutto quello che mi hai regalato.
Grazie Parigi per quello che mi stai regalando.

Alessandro Cesa
MSc in Digital and Interaction Design
A.A. 2025 — 2026
Politecnico di Milano