



UNIONE EUROPEA
Fondo Sociale Europeo



REACT EU



UNIVERSITÀ DEGLI STUDI DI MILANO

Dottorato in Filosofia e Scienze dell'Uomo
XXXVII Ciclo

Dipartimento di Filosofia
Piero Martinetti

TESI DI DOTTORATO DI RICERCA

La Superficie di Mezzo
L'Ecosistema delle Interfacce Digitali: una Tassonomia Critica

M-FIL/04

Alessandro Costella

Tutor: Prof. Andrea Pinotti

Co-tutor: Prof. Alfio Ferrara

Coordinatore: Prof. Niccolò Guicciardini

A.A. 2024/25

La borsa di dottorato è stata cofinanziata con risorse del
Programma di finanziamento PON REACT-EU
Azione IV.4 – Dottorati su tematiche dell'innovazione

Questa tesi riflette solo i punti di vista e le opinioni degli autori, né l'Unione Europea né
la Commissione Europea possono essere considerate responsabili per essi.

Ringraziamenti

Il percorso che conduce alla tesi di dottorato è lungo, intenso, tortuoso ed entusiasmante come solo un viaggio nell'ignoto sa esserlo. È a tratti emozionante, spaventoso, energetico, sovrachiaro e confuso. Non sorprende, dunque, che un simile viaggio non si compia da soli. Se al risultato, formalmente, è una sola persona ad arrivare, le persone che vi hanno contribuito in maniera imprescindibile sono molte, ed è opportuno esprimere almeno un po' della profonda gratitudine che ho provato, provo e ancora proverò per loro.

Innanzitutto, ringrazio dal profondo i miei supervisori, Andrea e Alfio. In modalità, misure e forme diverse, devo a entrambi ben più della conclusione del dottorato: sapete bene che, senza ognuno di voi, questo dottorato non sarebbe nemmeno mai iniziato, e non è dir poco. A entrambi devo, dunque, una parte fondamentale della mia vita e della mia identità da prima che iniziasse a formarsi. Grazie di cuore.

Inevitabilmente, un simile debito lo devo a tutta la mia famiglia. A mia madre e a mio padre, per l'eccezionale capacità di supporto umano e materiale nonostante i limiti della distanza, attraverso ogni fase del percorso. A mia sorella, che conosce l'efficacia di forme di supporto morale che pochi capirebbero. A Mara, che eroicamente ha saputo farsi carico delle mie necessità anche nei momenti più difficoltosi. Grazie infinitamente.

In ultimo, ringrazio quei colleghi che hanno rappresentato un esempio, che hanno saputo fornire consigli eccezionalmente preziosi, che hanno contribuito più o meno direttamente al lavoro: la squadra di AN-ICON, ognuno di voi. Un gruppo di eccellenze professionali e umane in egual misura, a cui va il merito dei miei più recenti pregi. Sulla coda del team AN-ICON, ringrazio anche l'asse HCI dell'ENAC di Tolosa e Catherine Letondal, la cui accoglienza e il cui contributo entusiasta hanno saputo arricchire il mio lavoro su moltissimi fronti. Grazie, e *merci*.

Un individuo di piccola e immensa statura, una volta, espresse molto bene ciò che provo in questo momento: *conosco la metà di voi soltanto a metà; e nutro, per meno della metà di voi, metà dell'affetto che meritate*. A voi vanno, in questo momento di immensa importanza, il mio pensiero e tutta la mia gratitudine.

Indice

Prefazione.....	7
Introduzione	10
I. Definire l'interfaccia.....	13
II. Storia e archeologia dell'interfaccia grafica.....	30
III. Classificare l'interfaccia	61
<i>A. Interfacce hardware</i>	64
<i>Interfacce tattili</i>	69
<i>Interfacce cinetiche</i>	77
<i>Interfacce a tracciamento oculare</i>	86
<i>Interfacce organiche</i>	89
<i>Interfacce a pulsantiera</i>	94
<i>Interfacce vocali</i>	102
<i>B. Interfacce formali</i>	109
<i>Interfacce grafiche tattili</i>	114
<i>Interfacce a linea di comando</i>	119
<i>Interfacce a linguaggio naturale</i>	122
<i>Interfacce geospaziali</i>	126
<i>C. Interfacce teleologiche</i>	130
<i>Interfacce tangibili</i>	134
<i>Interfacce multimodali</i>	138
<i>Interfacce naturali</i>	142
<i>Interfacce simulative</i>	147
<i>La diffusività della Realtà Aumentata</i>	154
Conclusione.....	161
Bibliografia	165

Prefazione

Il presente lavoro si colloca, sia sul piano concettuale che su quello metodologico, nello spazio tra discipline molto diverse tra loro, appartenenti ai mondi delle scienze umanistiche e delle scienze cosiddette più dure.

È innanzitutto doveroso inquadrare il lavoro all'interno della sua propria cornice: il progetto di ricerca ERC Advanced AN-ICON dell'Università degli Studi di Milano. Il punto cardine della ricerca della squadra guidata dal Professor Andrea Pinotti è lo studio delle tecnologie digitali immersive, tanto secondo il punto di vista pratico quanto secondo quelli storico e teorico.

Tale inquadramento è più di una formalità: la ricerca qui presentata è stata fortemente informata da quella dei numerosi membri del gruppo AN-ICON, informazione fortunata per la presenza di spiccata interdisciplinarietà, grande competenza professionale e ottimi consigli – caratteristiche che l'autore spera di aver assorbito almeno in parte negli anni del suo dottorato di ricerca.

La collocazione multidisciplinare è anche, forse poco sorprendentemente, il risultato di un percorso fortemente multidisciplinare dell'autore stesso, non solo nella varietà delle sue parti ma anche (e forse soprattutto) nella progressione che quel percorso ha compiuto, dalle scienze dure alle scienze umanistiche.

Introdurremo dunque rapidamente questo percorso, poiché riteniamo che l'esposizione dell'autore e dei suoi bias sia una pratica di grande valore scientifico – sia per fornire al lettore una cornice più completa in cui inquadrare il lavoro, sia per non privarlo del piacere di esprimere puntuali giudizi disciplinari, come: “Ecco, qui si vede proprio l'ingegnere!”. Riconosciamo che questa sarebbe una grave mancanza nei confronti di chi dovrà fruire del testo.

Le peregrinazioni disciplinari che hanno condotto a questa ricerca scientifica, metodologica e (soprattutto) umana cominciano dall'ingegneria dell'informazione, in particolare con un forte interesse per i temi dell'automazione, proseguono con l'informatica, affrontata con una vena fortemente ludologica e un orientamento al design e alla progettazione, e terminano con la filosofia estetica e la filosofia della

tecnica. Non dubitiamo che le infiltrazioni dei diversi approcci scientifici e le diverse *formae mentis* che caratterizzano ambiti così diversi facciano capolino in questo lavoro, e così anche i conflitti che emergono dall'accostamento di campi epistemici così solitamente lontani e difficili all'accordo.

La Superficie di Mezzo si offre come un tentativo di colmare una lacuna foriera di confusione, quella tra le differenti categorie e sottocategorie di ciò che nel discorso sulle tecnologie digitali prende il nome di interfaccia. Il concetto, seppur molto giovane, accompagna le nostre giornate, le nostre comunicazioni interpersonali, le nostre pratiche lavorative, il nostro tempo libero. Le sue genealogie e i suoi usi sono molteplici e non banali, e meritano di essere esplorati per comprendere quali meccanismi spingano diverse forme e diversi livelli ontologici di interfacce verso differenti meccanismi.

Per operare una simile analisi è necessario armarsi di differenti attrezzi: la storia del concetto, l'archeologia mediale delle tecnologie che esso descrive e la tassonomia, per citare i più evidenti.

È inoltre opportuno comprendere, nel pieno spirito antropocentrico proprio della filosofia della tecnica, alcuni dei meccanismi che le interfacce innescano negli individui e tra gli individui. Poiché siamo ben coscienti che discuterli tutti sia il lavoro della filosofia della tecnica nella sua interezza, nel presente lavoro si è scelto di focalizzarsi sugli aspetti spiccatamente critici, osservando come diverse forme e diversi livelli di interfaccia possano escludere individui e gruppi o riorientare forzatamente le loro pratiche. A questo scopo, il presente lavoro si è rivolto ai *disability studies*, e in particolare al loro orientarsi a ciò che sta ai margini e che rappresenta una deviazione dalla norma per comprendere le dinamiche tramite cui tale norma si costituisce, poggiando su diversi casi studio categorici e materiali¹.

Le categorie prese in analisi sono state presentate secondo un ordine determinato più dalla loro vicinanza concettuale che dall'ordine cronologico, accostando gruppi di dispositivi che presentino forti analogie o evidenti contrapposizioni con quelli

¹ R. Garland Thomson, *Extraordinary Bodies: Figuring Physical Disability in American Culture and Literature*, Columbia University Press, New York (1997), p.8.

che li precedono o li seguono, per quanto possibile. Per quanto riguarda invece i singoli dispositivi, l'autore si è rivolto all'esperienza diretta del loro uso ogni qualvolta è stato possibile o ha, in alternativa, fatto ricorso a esperienze d'uso pregresse. Tuttavia, il diverso livello di accessibilità di diverse tecnologie e l'importanza del loro uso situato e inteso, doveroso in un'analisi significativa, ha reso necessario, per alcune di esse, rivolgersi ad altri fruitori regolari di simili tecnologie. È questo, ad esempio, il caso di dispositivi protesici per disabilità non funzionali come la cecità.

Lo strumento utilizzato nella raccolta di dati da altri utenti è stato, durante la ricerca che ha condotto a questo testo, quello dell'intervista contestuale: una tecnica di intervista che prevede di situare ogni utilizzo di una data tecnologia rivolgendosi sempre alla memoria e alla prassi, appreso dall'autore durante un visiting di ricerca presso l'ENAC di Tolosa tra il 2023 e il 2024 grazie alla generosa supervisione della professoressa Catherine Letondal e alle sue ricerche sull'interazione uomo-macchina. Durante la permanenza nella città di Tolosa sono state condotte simili interviste sia per apprenderne le procedure, sia per raccogliere dati inerenti ad alcuni dispositivi candidati all'analisi in questo testo. Ai ricercatori dell'asse interazione uomo-macchina dell'ENAC è anche dovuta la conoscenza di più di uno dei dispositivi presi in analisi in questa tesi, parte del lavoro di ricerca sugli utenti affetti da disabilità e una particolare attenzione alla ricerca delle pratiche della progettazione di interfacce fisiche e digitali.

Introduzione

Viviamo nell'era dell'interattività. La diffusione esplosiva dell'utilizzo dei computer, intesi nel senso più ampio di "tecnologie basate sull'uso del transistor e di circuiti integrati", ha travolto a velocità spaventose la civiltà umana con la logica pervasiva dell'interruttore. La partecipazione dell'individuo allo svolgimento dell'attività macchinica ha raggiunto la sua forma, per così dire, più "orizzontale": l'universalizzazione dell'agire con la macchina tramite scambio di input e output, quantitativamente comparabili. Senza la rivoluzione digitale sarebbe difficile concepire la stesura di un testo come attività simile a un calcolo matematico, e insieme simile alla partecipazione a un'attività ludica e a una conversazione.

Vi sono tuttavia molti presupposti a questo sconvolgimento. I più evidenti sono la generale attribuzione della capacità di agire a una macchina, che implica un elevato livello di autonomia per uno strumento tecnico, e in particolare la sua capacità di porsi parte di una relazione o di una transazione, che è una forma specifica di interazione, una forma di comunicazione.

Evidentemente, però, i partecipanti a questi scambi comunicativi parlano lingue diverse, e necessitano di forte mediazione. Come sarà approfondito nella presente tesi, tale mediazione ha molte forme ed è rimasta fuori dalla portata del grande pubblico per molto tempo. Ma proprio dalla diversità di forme che questa mediazione ha assunto è emerso un fenomeno ben noto nei primi anni del terzo millennio: il fenomeno dell'interfaccia digitale.

"Interfaccia" è un termine che appartiene a una famiglia di etichette curiose, la cui caratteristica è quella di comprendere un insieme di fenomeni o enti di complessa definizione, ma di tale familiarità da non scatenare spesso il problema del dare loro tale definizione in una forma precisa. Non è questa la categoria del termine "arte", che porta con sé un sacro e storico senso di incertezza. "Interfaccia" è più simile a "divertimento": tutti conoscono il divertimento, così intimamente da non preoccuparsi di doverlo definire per poterne discutere o riconoscerlo quando si

presenta in una qualunque delle sue innumerevoli forme¹. Tutti sanno parlare dell'interfaccia di un editor di testo, di un videogioco, dell'interfaccia di YouTube o di Facebook. Ci confrontiamo giornalmente con una pletora di interfacce per svolgere le attività più disparate.

L'interfaccia è un fenomeno strettamente legato al concetto di medium, anch'esso dalle forme molteplici e difficili da afferrare, e similmente difficile da definire. Sono stati compiuti, a tal proposito, alcuni notevoli tentativi di importanza storica e significativi per comprendere l'evoluzione del concetto, che saranno riportati ed esaminati nel primo capitolo. Nel primo capitolo "I. Definire l'interfaccia" verranno esplorate differenti definizioni di interfaccia a partire dall'origine del termine nell'ambito della chimica, saranno delineati i differenti significati che esso assume in diverse discipline a partire da differenti autori e differenti periodi della storia del concetto di interfaccia, e verranno infine discussi tentativi di superare i possibili limiti delle definizioni più comuni di interfaccia in ambito informatico, partendo dal rapporto tra interfaccia e superficie. Nel secondo capitolo "II. Storia e archeologia dell'interfaccia grafica" verrà condotta una ricerca delle origini dei dispositivi e degli oggetti tecnici all'origine delle interfacce contemporanee, sia da un punto di vista prettamente cronologico ed evolutivo, sia dal punto di vista dell'archeologia dei media. Questo secondo approccio permetterà di osservare l'emergere delle dinamiche che oggi consideriamo proprie dell'interfaccia digitale in media che l'hanno preceduta o accompagnata, così come in sue anticipazioni e rappresentazioni in diversi media, e come i contributi alla sua origine possano spaziare ben al di là della storia dell'informatica, fino al mondo del non-umano. Il terzo capitolo "III. Classificare l'interfaccia" presenta la costruzione tassonomica che costituisce il nucleo della tesi, costruzione gerarchicamente strutturata in due

¹ Ci si riferisce qui all'utilizzo comune del concetto di divertimento. Sul fronte scientifico, tuttavia, l'incertezza rimane: diversi studi hanno cercato di individuare i caratteri e le diversissime forme del divertimento, giungendo a diverse possibili classificazioni. Si prendano ad esempio C. McManus e A. Furnham. "Fun, fun, fun": *Types of Fun, Attitudes to Fun, and Their Relation to Personality and Biographical Factors*, in *Psychology*, Vol. 1 No. 3 (2010) e T. T. Oh, M. T. Pham. *The Psychology of Fun: Inquiry into Consumers' Fun Experiences*, in *Advanced Consumer Research*, Vol. 46 (2018).

livelli. Un primo livello, di ampie categorie, definito secondo i differenti livelli ontologici in cui l'interfaccia si presenta, viene discussa e operata nel quotidiano come negli ambienti accademici e professionali. Seguendo le orme dei *disability studies*, per ognuno di questi livelli sarà definita una norma vigente come costituente di ogni forma non normata di interfaccia nella stessa categoria ontologica, che si costituisce come tale proprio in funzione della sua distinzione da tutte le categorie non normali. All'interno delle superclassi di primo livello sarà poi delineato un secondo livello, più granulare, che analizza differenti tipi di interfaccia iscrivibili allo stesso reame ontologico proponendo di volta in volta casi studio e osservazioni critiche riguardo i processi di normalizzazione corporea, sociale o culturale che tali interfacce, per motivi storici, tecnici, politici o economici portano con sé.

I. Definire l'interfaccia

Parafrasando Aristotele, l'interfaccia si dice in molti modi. Non solo: l'interfaccia si dice in molti tempi e in più di una disciplina. Le diverse definizioni di interfaccia, però, non sempre esauriscono la più comune intuizione che le tecnologie digitali richiamano alla mente. Il termine, così ampiamente diffuso nell'ambito informatico, precede quello stesso ambito in realtà di più di mezzo secolo. Il vocabolario Merriam-Webster ne colloca l'origine nella seconda metà del XIX secolo, rintracciandone il primo uso nell'anno 1882¹ in ambito chimico. La definizione relativa a tale primo uso è la seguente: “a surface forming a common boundary of two bodies, spaces, or phases e. g. *an oil-water interface*”. Il termine è in effetti di origine inglese, evidentemente composto dalle parti *inter-* e *-face*.

Il termine non fornisce alcuna informazione sulla natura di ciò che il descrive, salvo la sua collocazione spaziale: si tratta di qualcosa che si colloca tra due superfici. L'ambito originario del nome è infatti quello chimico, e più ampiamente quello fisico, in cui si indica con il termine *interfaccia* la superficie di contatto tra due corpi, due fasi o due porzioni differenti di materia che siano tra loro immiscibili. Ciò può indicare due liquidi, come nell'esempio riportato dal Merriam-Webster, ma anche un liquido e un gas che non sia solubile in tale liquido, un fluido e un solido e via discorrendo. La perdita del carattere di immiscibilità causerebbe la sparizione di una superficie di contatto rintracciabile. Interfaccia è dunque, sin dalla sua origine, un termine che descrive un punto di contatto tra realtà diverse, il cui unico carattere certo è la persistenza di tale diversità. Peraltro, tale punto di contatto non ha una sua sostanza indipendente. L'interfaccia esiste in funzione di ciò che la circonda, e non può in alcun modo darsi senza una coppia di altri, indipendenti da essa. La sua materialità è quantomai curiosa, per non dire sfuggente: non è composta della materia di alcuna delle due fasi che pone in contatto, che non possono miscelarsi tra loro dando forma ad alcunché di misto, eppure è equamente

¹ <https://www.merriam-webster.com/dictionary/interface>, ultima consultazione 23/09/2024.

parte di ambedue. L'interfaccia non appartiene al reame del solido, del liquido, del gassoso o di qualunque altro stato fisico della materia. L'interfaccia è l'effetto di un contatto, emerge da un contatto, e in ciò si evidenzia una delle poche caratteristiche certe di ogni interfaccia. Sin dalla sua origine, l'interfaccia non è. L'interfaccia avviene.

Nella libertà delle regole linguistiche inglesi il termine, la cui natura relazionale porge il fianco a innumerevoli usi, è successivamente divenuto verbo. Rifacendosi ancora una volta al Merriam-Webster, il primo uso noto del termine *to interface* risale all'anno 1962, in forma transitiva. Il significato corrisponde a quello italiano di interfacciare, connettere tramite un'interfaccia. Il 1962 fu di certo un anno molto significativo per la storia del termine interfaccia, che apparve dotato di un nuovo significato ricco di accezioni ne *La Galassia Gutenberg* di Marshall McLuhan². Il filosofo la definì come il concetto, appartenente alla fisica moderna, “dell'incontro e metamorfosi di due strutture”. Nello spiegare l'uso del termine, McLuhan sostiene che “due culture o tecnologie possono, come le galassie astronomiche, penetrare l'una dentro l'altra senza entrare in collisione: ma non senza che muti la configurazione di ciascuna”. Si tratta di un paragone che descrive una grande potenza trasformativa: l'immiscibilità tra le fasi della materia necessaria alla definizione originale viene mantenuta, ma prende una nuova forma. L'incontro che presume l'attraversamento e lo sconvolgimento reciproco, rendendo impossibile identificare una singola superficie di contatto. In tale definizione la forma verbale del termine è quantomai naturale: il particolare incontro descritto da McLuhan è un agire e allo stesso tempo un subire l'azione dell'altro, un processo nel tempo e un processo nello spazio, che trasforma le parti dell'incontro mentre le attraversa. Si dovrà tuttavia riconoscere che, seppur la definizione di McLuhan sia di grande impatto per il concetto di interfaccia ed estremamente significativa per comprendere il significato contemporaneo del termine, essa non ci permette davvero di identificare immediatamente che cosa sia un'interfaccia digitale.

² M. McLuhan, *La galassia Gutenberg. Nascita dell'uomo tipografico*, S. Rizzo (trad.), Armando, Roma (2011), p. 190.

Comprendere quale parte di un dispositivo digitale sia un'interfaccia appare alquanto arduo con questa definizione.

Nel caso si volesse individuare, ad esempio, l'interfaccia tra chi fa uso di un computer portatile e il suo strumento di lavoro, sarebbe necessario individuare il luogo del loro incontro. Si potrebbe individuarlo nel reame fisico, sulla superficie di incontro tra la plastica del mouse e della tastiera e la pelle dei polpastrelli. Si potrebbe individuarlo nel reame estetico, dove i sistemi percettivi del soggetto ricercano gli stimoli della macchina e dell'interazione con le sue parti. Si potrebbe individuarlo nel reame semiotico, dove la parte umana dell'incontro riceve e interpreta segni che la macchina offre, e nel mentre offre alla macchina segni che essa sa interpretare³ in tempo reale. E in tutti questi casi si starebbe dalla parte della ragione. L'interfaccia può avvenire a ognuno di questi livelli, e tende a farlo simultaneamente con specificità diverse in base al luogo. Inoltre, il computer può essere allo stesso tempo interfaccia tra chi lo sta usando e altre realtà: informazioni, individui e pratiche. L'interfaccia non viene mai sola. Comincia dunque ad apparire evidente un legame estremamente stretto tra l'interfaccia e il *medium*. Non sarà qui discussa la differenza tra due concetti già difficili da definire di per sé, ma il loro stretto rapporto, la loro comune proprietà di essere in mezzo, di essere insieme una divisione e un canale, genera un forte parallelismo: tecnologie drasticamente medianti come i computer, gli smartphone e internet, che esibiscono perfettamente la caratteristica dei *media* di contenere sempre un altro *medium*⁴, sono inevitabilmente strutturati come catene non lineari di interfacce. Nel mondo della tecnologia dell'informazione, che è tecnologia della trasmissione, l'interfaccia è il luogo dove l'informazione viene tradotta da una forma a un'altra, e può assumere tante forme quante sono le possibili traduzioni dell'informazione tra due diverse forme di essa.

³ Abbracciamo qui un'accezione ampia dei termini offrire e interpretare, nel primo caso per segni che qualche progettista ha fisicamente inserito nei sistemi della macchina, e nel secondo per rispondere ai gesti del soggetto in maniera rapida e aderente a un qualche criterio coerente.

⁴ McLuhan, M., *Gli strumenti del comunicare*, Capriolo, E. (trad.), Garzanti, Milano (1967).

Giungiamo così infine a un'ultima definizione di interfaccia, fornita da Lev Manovich nel suo celeberrimo *The Language of New Media*:

In termini semiotici, l'interfaccia del computer è una sorta di [*acts as a*, n.d.a.] codice che porta dei messaggi culturali in una varietà di media. [...] Nella comunicazione culturale, un codice non è quasi mai solo un meccanismo di trasporto neutrale, di solito influenza anche i messaggi che viaggiano su di esso. Per esempio, può facilitare la comprensione di alcuni messaggi e renderne altri addirittura inconcepibili.⁵

Ancora una volta, l'interfaccia è definita secondo il suo agire: prima dal suo separare, poi dal suo trasformare e infine dal suo trasportare messaggi tramite codifica e decodifica. A differenza della traduzione in lingua italiana, non vi è nel testo originale verbo essere, e l'interfaccia non è descritta a partire da altre forme come la superficie o l'incontro: "the computer interface acts". Il campo di applicazione di questa definizione non è evidentemente lo stesso delle due precedenti, eppure la forma dell'interfaccia è la meno specifica in assoluto: il codice di trasmissione in una varietà di media. La semiotica non è certo il campo dell'autore del presente testo, ma la verità secondo cui una codifica altera il contenuto del messaggio che trasforma si presta – almeno in una sua forma banale – all'intuizione, e mantiene il suo valore di verità anche nell'ambito delle telecomunicazioni, nel quale diverse necessità di trasmissione hanno dato origine a differenti protocolli di codifica e decodifica.

Il processo di codifica e decodifica delle interfacce digitali e quello delle telecomunicazioni sono molto simili nella pratica: la loro principale differenza è nello scopo con cui il testo che si vuole comunicare viene codificato.

I processi di codifica che compongono la trasmissione elettromagnetica sono molti e molto diversi tra loro. Per essere trasmesso, un messaggio (sia esso verbale o multimediale) deve essere sottoposto a una serie di traduzioni successive. Prendiamo ad esempio il messaggio vocale, il cui uso è ormai universale su

⁵ L. Manovich, *Il linguaggio dei nuovi media*, Merlini, R. (trad.), Edizioni Olivares, Milano (2022), p.90.

piattaforme di messaggistica come WhatsApp. Il suo formato e il suo contenuto devono innanzitutto venir tradotti in una forma compatibile con il computer (un testo alfabetico, visivo, sonoro, tattile...), in questo caso in un discorso sufficientemente chiaro rispetto al rumore di sottofondo. Dopodiché quella forma viene inserita nel computer tramite una qualche operazione di registrazione o digitazione. Parlando vicino al microfono l'utente fa vibrare la sua membrana, che con il suo movimento genera segnali elettrici nei circuiti del telefono. Quest'input viene trasformato dalla macchina in una sua rappresentazione digitalizzata, composta di lunghe sequenze binarie: nel caso del messaggio vocale, si tratta di trasformare il flusso continuo di tensione elettrica in una sequenza di valori che ne descrivano l'andamento, rappresentati in codice binario. Una volta che questo pacco digitale è pronto per la spedizione, inizia l'applicazione di un protocollo di codifica: gli si aggiungono, sempre in codice binario, informazioni relative al mittente e al destinatario, l'orario di invio, delle ridondanze per resistere meglio all'errore di trasmissione, secondo regole che dipendono dallo specifico protocollo applicato. Una volta arricchito di questi equivalenti digitali del francobollo e delle informazioni di spedizione, la sequenza di cifre che lo compone viene tradotta, grazie alla Trasformata di Fourier, in una serie di parametri numerici con cui generare un segnale elettromagnetico periodico che si trasmette tramite cavi o si diffonde tramite un'antenna.

Con la ricezione di quel segnale periodico e la sua traduzione in segnali elettrici e scomposizione nelle sue componenti, inizia il processo inverso di decodifica. Le componenti di questo processo, così come i fattori di disturbo e di incertezza, sono evidentemente innumerevoli. Ad esempio, poiché i segnali tendono a venir trasmessi in un tempo finito (a causa dello sconveniente fatto che le nostre vite, così come le nostre agende, presentano un forte carattere di finitezza), la banda dei segnali trasmessi sarà infinita⁶, e dunque il segnale originale dovrà essere approssimato da chi lo riceve. Il motivo per cui il segnale viene codificato secondo

⁶ Per tutto ciò che riguarda l'elaborazione del segnale, ci riferiamo in questa tesi a N. Benvenuto, M. Zorzi, *Principles of Communications Networks and Systems*, John Wiley & Sons Ltd., Chichester, UK (2011).

un dato protocollo è per renderlo più robusto agli errori del canale su cui si intende trasmettere il messaggio; per questo motivo il processo viene anche detto Codifica di Canale (Channel Coding⁷).

Analogamente, e soprattutto contemporaneamente, si stratifica la codifica culturale operata dalle interfacce fisiche e digitali di un computer: il messaggio, che è stato precedentemente tradotto in una forma testuale digitalizzabile (in forma alfabetica, visuale, uditiva, tattile), viene tradotto a seconda dei metadati contestuali nella lingua corretta o nella forma audiovisiva appropriata, per poi fungere da indice, metafora o metonimia di una determinata funzione o contenuto, che verrà inserita in un contesto secondo le regole dell'interfaccia grafica, e infine riprodotta da un dispositivo di output per venir ricevuta e decodificata dal destinatario.

Tali regole hanno origine storica. Come evidenzia lo stesso Lev Manovich, le strutture ideologiche imposte dall'interfaccia sono altamente specifiche: l'organizzazione dei files in cartelle rende inconcepibile una loro fruizione non gerarchica, che non è invece assolutamente pensabile per la fruizione dei siti internet veicolata da un browser⁸. A queste aggiungeremo l'ideologia fortemente accentratrice che le Graphical User Interfaces impongono, circondando un contenuto con tutti i possibili strumenti per operare su di esso, disposti per la maggior parte nella parte alta o laterale dello schermo, esibendo alcune funzioni e nascondendone altre dietro intricati labirinti di menù e opzioni, creando bolle agentive in cui rinchiudersi per operare. La storicità di tali meccanismi sarà discussa più a fondo in questa tesi.

Il dibattito sulla natura dell'interfaccia non si è, ovviamente, esaurito con Manovich. Nonostante i tentativi di delineare esattamente un concetto così vasto, mutevole e trasversale rischi di sfociare in nient'altro che una lunga e (finora) infruttuosa diatriba su cosa meriti di venir definito interfaccia e cosa no, la messa in discussione degli stessi fondamenti della definizione di interfaccia può presentare

⁷ Ibid.

⁸ L. Manovich, *Il linguaggio dei nuovi media*, Merlini, R. (trad.), Edizioni Olivares, Milano (2022), p.91.

un significativo incentivo alla riflessione sulla presenza dell'interfaccia nel nostro relazionarci ai media.

In una pubblicazione del 2015 Janlert e Stolterman⁹ hanno fornito un insieme di possibili categorie di interfaccia, arricchito del ruolo che ogni tipo di interfaccia ricopre e “dove si trova” basandosi sul concetto di *thought style* di Ludwig Fleck¹⁰. A questo scopo, i due autori hanno categorizzato l'interfaccia secondo quattro differenti *thought styles*, che ne producono quattro diverse definizioni secondo i differenti modi di pensarla:

These thought styles are not distinctly related to a particular period, particular technology, or type of designs, but they have evolved over time and can be seen as stemming from different traditions. [...]

The four thought styles are: the interface thought of as

1. a *surface of contact* between matching objects (from the tradition of industrial machine making),
2. a *boundary* of an independent (self-sustained) object (from biology and traditional artifact design),
3. a *means for controlling* (operating, checking, steering) an object (from design of complex machines), and
4. a *means for expressions and impressions*, a target of interpretations and affectations (from human communication, architecture, and art).

A fifth thought style might be added, in which the interface is seen as a *channel of communication*, that is, an opening or conduit through which “stuff” passes. This way of thinking about the interface is quite common, especially in HCI. Stuff passes *through* the interface from one side to the other side, often in both directions, and usually selectively (some will pass, some will not be able to get through).

Se la prima e la seconda modalità di descrizione dell'interfaccia richiamano la definizione fisico-chimica originale e la quarta e la quinta rimandano alla definizione di Lev Manovich, la terza si avvicina molto più delle precedenti alla definizione più colloquiale di interfaccia digitale. La presenza del verbo “checking”

⁹ L.E. Janlert, and E. Stolterman, *Faceless Interaction — a Conceptual Examination of the Notion of Interface: Past, Present, and Future*, in *Human-Computer Interaction*, 30 (2015) pp.507-539.

¹⁰ L. Fleck, *Genesis and Development of a Scientific Fact*, F. Bradley, T.J. Trenn (trad.), University of Chicago Press, Londra (1979).

tra le azioni veicolate dall'interfaccia permette di includere ogni dispositivo di output come necessario al controllo di un dispositivo nella sua interfaccia. Queste differenti categorizzazioni, seppur parzialmente manchevoli dell'aspetto più "mcluhaniano" di reciproca influenza tra le realtà che l'interfaccia separa e connette, sono focalizzate su una caratteristica dell'interfaccia e la sua possibile universalità, che gli autori desiderano mettere in discussione: la presenza di una superficie. Il problema che Janlert e Stolterman desiderano evidenziare emerge in particolare dalla terza interpretazione del concetto di interfaccia, per estendersi alla quarta:

Even though instrument panels were, and still are, a common way of implementing control, nothing in the basic concept of control requires it to be exercised through some physical surface. Indeed, theremin [sic] players just use their ears and hands in free gestures to control pitch and volume; blind conductors have been known to successfully lead an orchestra. Similarly, with Thought Style 4—to express and receive impressions, affect and be affected—you do not really need the mediation of a physical surface: A common, everyday example is an ordinary telephone conversation with another person. Faced with the current interface conundrum, we have decided to go literal and—just as the term implies—restrict “interface” to apply to real surfaces and surface-bound interaction, and not use it when there is no actual surface involved. [...] An important implication of this move is that *interaction* does not necessarily require an *interface*.

Ad oggi, le interfacce non tattili di interazione digitale sono di certo al loro picco. Il progressivo miglioramento delle tecnologie di visione computazionale e l'integrazione pervasiva dell'intelligenza artificiale, e in particolare di modelli in grado di processare il linguaggio naturale, rendono più che mai pertinente il dibattito sulla natura planare dell'interfaccia. Eppure, le nostre interazioni con telefoni e computer sono, ad oggi, ancora veicolate principalmente tramite schermo. La natura di tali schermi varia, in particolare nella loro prestazione tattile, ma la loro presenza è indiscussa.

Va qui anticipato che qualunque discorso sullo schermo¹¹ rischia di indispettire, e a buon titolo, tutte le categorie dei mediologi e dei teorici e degli storici del cinema, dotti in aspetti della realtà schermica di cui l'informatica classica tende occuparsi troppo poco. Per questo motivo ci limiteremo qui a evidenziare due cose: la prima, banale, è che lo schermo è un eccezionale esempio di superficie comunicativa. La seconda è che lo schermo del computer tende a moltiplicare le interfacce superficiali che contiene. La metafora dominante nell'informatica contemporanea è, infatti, quella dell'ufficio. A causa della sua origine, essa è stata concepita come un supporto a forme lavoro fortemente testuale, seppur anticipando un giorno in cui il computer avrebbe supportato anche altre forme visuali, come l'ispezione di modelli tridimensionali di edifici nell'ufficio di un architetto¹². Per questo motivo, i nostri computer sono pieni di fogli di carta. Nell'esperienza comune, ossia al di fuori della produzione di origami o di sgradevoli tagli ai polpastrelli, il foglio di carta si presenta come pura superficie: esso funziona e si definisce in quanto superficie, in grado di esporre le informazioni che vi sono sopra riportate o offrendo i suoi spazi bianchi al nostro operare. Analogamente importante è la natura planare dell'altra importantissima metafora d'ufficio dell'interfaccia grafica, ossia il desktop. Nella schermata con cui l'utente digitale di questi tempi si confronta più di frequente, ogni cosa è disposta su una superficie. Nonostante il nome richiami la scrivania, essa è più propriamente una parete o una lavagna, decorata con una o più immagini e a cui sono appese a diverse altezze icone di varia forma e colore.

In questo paradigma estremamente bidimensionale si inseriscono ora forme di interazione che non presentano superfici visuali. Il microfono, la telecamera cinetica per la lettura dei gesti, i giroscopi e gli accelerometri in grado di leggere

¹¹ Pur essendo impossibile fornire una panoramica della vastità dello studio degli schermi, riteniamo adeguato indicare la presenza nel panorama italiano del volume *I poteri degli schermi. Contributi italiani a un dibattito internazionale*, a cura di M. Carbone, A.C. Dalmasso e J. Bodini edita da Mimesis nel 2020, che include significativi contributi di noti autori di cultura visuale e schermologia.

¹² Il lavoro pionieristico di Douglas Engelbart, padre dell'interfaccia grafica, verrà discusso più approfonditamente nel presente lavoro. Il testo qui riferito è il monumentale *Augmenting Human Intelligence: A Conceptual Framework* del 1962. Il testo è disponibile online, in forma integrale, grazie al Doug Engelbart Institute, sia tramite scansioni del documento originale sia tramite una più leggibile trascrizione. <https://www.dougenelbart.org/pubs/augment-3906.html> Ultima consultazione 05/06/24.

gli spostamenti di un dispositivo e utilizzarli come *input*. Janlert e Stolterman evidenziano come siano spesso gli utenti affetti da disabilità a fruire per primi di simili tecnologie, a causa di peggiori condizioni visive, tattili o motorie. L'interazione visuale e tattile con le superfici implicano e normalizzano determinate prestazioni visuali e motorie, e forme di interazioni differenti sono rimaste per molto tempo dispositivi di nicchia, il cui imperfetto funzionamento rappresentava il fardello da pagare per soddisfare la necessità di interagire digitalmente nell'impossibilità di esercitare determinate funzioni secondo una forte norma corporea. Le interazioni prive di superfici sia nel processo di espressione sia in quello di impressione sono definite dagli autori “faceless interactions in the strong sense”, contrapponendosi all'espressione “faceless interactions in the weak sense” che identifica gli scambi tra utente e macchina veicolati, ad esempio, da un *input* vocale e un *output* schermico. Tali interazioni non sarebbero, dunque, veicolate da interfacce.

Che cosa, dunque, veicola queste interazioni? Una possibile risposta a questa domanda è stata elaborata da alcuni ricercatori nell'ambito dell'interazione uomo-macchina della Scuola Nazionale dell'Aviazione Civile di Tolosa (che prende l'acronimo di ENAC in francese) appoggiandosi al quinto *thought style* proposto da Janlert e Stolterman, quello che identifica l'interfaccia con un canale di comunicazione. Il concetto proposto dal gruppo, alla cui guida è la professoressa Catherine Letondal, prende il nome di *transfaccia*¹³. Il punto focale del concetto proposto da Letondal, con cui ho avuto il privilegio di collaborare e di discutere l'idea della transfaccia durante un soggiorno di ricerca nella città di Tolosa, è la sua porosità. La transfaccia propone un modello di canale comunicativo con diversi livelli di attraversabilità e differenti resistenze all'attraversamento di segnali, immagini, meccanismi di attenzione e sguardo e dei loro effetti sull'interazione che i loro rispettivi livelli di attraversamento veicolano. Il modello è dunque triadico, e

¹³ La proposta concettuale è stata presentata a diverse conferenze ed è fruibile sul sito dell'ENAC, all'indirizzo <https://lii.enac.fr/projects/transfaces/>. Seppur non esista una traduzione precedente di *transface*, la cui scrittura rimane invariata in lingua francese e in lingua inglese, ci siamo presi la libertà di adottare *transfaccia* come variante italiana.

descrive l'interazione in termini di un'area attraversabile (“the traversable: an interaction area”), un flusso attraversante (“the flow: what must be able to pass through”) e un'esperienza di attraversamento (“the experience of the traversal”). Nell'attraversamento di tale canale, il flusso può essere trasdotto, tradotto o trasformato.

Otherness, alterity	the concept of transface separates what is other and links to it, in the manner of an inter-face
Effects	The “traversable”
	an interaction area (ex: a 3D virtual area, a touchscreen, an ultrasonic area)
	Traversing: the flow
	what must be able to pass through, to traverse (ex: gestures, light, sound, haptic signal, electric signal)
	The experience of the traversal
	experience for the user who is traversed by what they are traversing
Operations on the flows	transduction, translation, transformation

Fig. 1: Le componenti della transfaccia proposte dal gruppo di ricerca dell'ENAC.

La superficie fisica viene dunque riconfigurata come spazio metaforico. Questo, però, non ci appare sufficiente a eliminare dal dibattito la natura planare dell'interfaccia digitale. Al di sotto delle metafore, infatti, permane la materialità della macchina. L'hardware è sempre costituito di oggetti nello spazio, dotati di una loro massa, e che occupano necessariamente un dato volume che presenta l'inevitabile caratteristica di essere finito. Ognuna di queste parti hardware è dunque delimitata da superfici, che non cessano di presentarsi a noi e alle nostre interazioni quand'anche tali interazioni cessassero di essere visuali o tattili. Quando facciamo uso di comandi vocali, infatti, non ci limitiamo a parlare aspettando che la membrana di un microfono capti il segnale che abbiamo emesso. Quando ci rivolgiamo a un dispositivo, siamo ben coscienti della presenza del microfono e cerchiamo di rivolgerci ad esso nello spazio. Analogamente, quando ascoltiamo ciò che un dispositivo elettronico emette, che si tratti di musica, della voce della

persona che stiamo chiamando o una lezione registrata, orientiamo attivamente il nostro sistema uditivo verso la superficie che sta emettendo quel suono. Come la maggior parte dei comportamenti legati alle tecnologie complesse anche questi meccanismi, che non richiedono genericamente di avvenire consciamente o con alcuna insistenza a causa della disposizione spaziale dei dispositivi (ad esempio il microfono di fronte alla bocca), divengono molto evidenti non appena le tecnologie o le interazioni non funzionano correttamente. Non appena la macchina non soddisfa le nostre aspettative, e ci sentiamo dire dalla persona con cui stiamo parlando che “non ci sente bene”, avviciniamo immediatamente la bocca alla superficie del microfono. Non appena la telecamera cinetica non rileva correttamente i nostri gesti, ci rivolgiamo con insistenza alla lente dell’obiettivo, cercando di assecondare un suo allineamento spaziale determinato dalla sua percepita bidimensionalità, che si estende di fronte a noi. Fintanto che possiamo agire intenzionalmente nei confronti di un dispositivo, non si dà interazione priva di superficie.

Differente è, in questo senso, il caso dell’interazione con la macchina tramite segnali biologici: l’elettromiografia, in grado di farci controllare dispositivi tecnici con il solo pensiero, non mette a nostra disposizione superfici a cui rivolgerci all’infuori di noi. I sensori che captano il segnale si trovano nel nostro corpo. Ciò nondimeno, è il movimento muscolare indirizzato a una precisa area che permette di inviare segnali alla macchina, e tale movimento si rivolge alla superficie dei sensori. L’autore di questo lavoro non ha mai, sfortunatamente, fatto uso di dispositivi elettromiografici, per cui non gli è possibile riportarne una fenomenologia più dettagliata vissuta in prima persona, riportando quanto l’intenzione sia rivolta al muscolo o alla superficie del dispositivo, o eventualmente ad altre superfici con cui relazionarsi. Ma qualunque sia la distribuzione dell’attenzione nell’agire, la cui natura è evidentemente cosciente e intenzionale, essa passa per la superficie di un dispositivo che dovrà poi tradurre quei segnali.

Perfino nei casi di interazione tramite tecnologia pervasiva o di simulazione della stessa, come in una casa intelligente interamente riempita di microfoni e altoparlanti

o una gigantesca stanza virtuale all'interno della quale si interagisce acusticamente o tramite tracciamento del movimento oculare, la nostra coscienza dei dispositivi non viene meno. Anzi, essa conduce piuttosto a un estremo caso di identificazione del dispositivo tecnico con l'ambiente: nel parlare con una casa intelligente i cui microfoni siano impossibili da localizzare, ci rivolgiamo alla casa stessa e alle sue pareti.

Se dunque possiamo dichiararci pienamente d'accordo con l'idea che non esista interfaccia senza superficie, siamo anche dell'idea che non esista interazione tecnica il cui presentarsi non implichi la presenza di una superficie. Dall'impugnatura dell'ascia acheuleana alle tecnologie cerebrali, non si dà oggetto tecnico privo di una superficie di interfaccia.

Alla luce dell'analisi condotta finora e della diversità degli esempi che l'hanno accompagnata, e in particolare col metonimico riferimento all'ascia acheuleana, va riconosciuta al concetto di interfaccia una grande potenza. Il concetto, seppur nuovo rispetto ai tempi della Storia, ha riscosso un tale successo dalla sua diffusione da permettere una rilettura di fenomeni, periodi e tecnologie che non appartengono a quest'epoca secondo la chiave di lettura dell'interfaccia stessa. In questo modo, l'interfaccia in senso ampio può divenire l'oggetto di innumerevoli discorsi sulla medialità, sulla strumentalità, sull'abitudine e il gesto. Ne riportiamo qui un notevole esempio: quello delle *Brain-Artifact Interfaces*, come vengono intese da Lambros Malafouris in un omonimo articolo del 2010¹⁴. Il concetto di Interfaccia Cervello-Artefatto è stato proposto dal pioniere dell'archeologia cognitiva come strumento utile a spostare l'attenzione della ricerca sugli effetti antropologici e cognitivi dell'uso di determinati strumenti nell'evoluzione umana, nel contesto della teoria del Material Engagement formulata dallo stesso Malafouris. Il precetto della teoria centrale dell'archeologo greco-britannico è quello secondo cui il pensiero umano avviene solo con e attraverso le cose materiali, che sono da considerarsi parte della realtà pensante che emerge nell'ingaggio dell'umano con la

¹⁴ L. Malafouris, *The Brain-Artifact Interface (BAI): a Challenge for Archaeology and Cultural Neuroscience*, in *Social Cognitive and Affective Neuroscience*, 2-3 (5 Jun 2010), pp.264-273.

materia¹⁵. L'ipotesi è quella secondo cui l'atto di pensare non possa mai avvenire nel vuoto, inteso come assenza di materia, ma sia esso stesso sempre e comunque parte di tale ingaggio: il *thinking* sarebbe dunque un *thinging*, un processo proprio del rapporto trasformativo con le cose che ci circondano¹⁶. Le Interfacce Cervello-Artefatto divengono così una speciale categoria che descrive l'effetto di alcune delle cose che ci circondano, ossia "il tipo di mediazione tecnologica (strutture materiali, processi, oggetti o altri apparati o pratiche socio-materiali) che permettono la configurazione di un allineamento dinamico o sintonia (*tuning*) tra plasticità neurale e culturale." Si tratta dunque di interfacce, o "mediazioni tecnologiche", con determinati effetti sulla cognizione umana. L'identità tra interfaccia e mediazione tecnologica in Malafouris è chiarita dallo stesso archeologo:

In its broad sense an interface is essentially any natural or artificial mediational means or prosthesis that enables, constrains and in general specifies communication and interaction between entities or processes. The human body, language and gesture are some obvious examples of such an interface, as it is also, the handheld rake used by a monkey to retrieve distant food [...].¹⁷

L'interesse espresso dalle definizioni fornite dallo studioso non è rivolto all'attraversamento, ma alla riconfigurazione dell'informazione e del pensiero che esse operano. Gli effetti di queste interfacce possono essere distinti secondo tre principali categorie: "mediational", "temporal" e "plastic". La prima categoria si

¹⁵ La scelta terminologica nel presente passaggio ("cose materiali", "realtà pensante") è dettata da un tentativo di coerenza allo scopo, dichiarato da Malafouris nelle sue pubblicazioni, di superare una dicotomia soggetto-oggetto che separa parti di un'entità pensante costituita tanto dalla componente umana quanto dagli strumenti che non si può davvero scindere in essere senza perdere la sua sostanza.

¹⁶La formulazione della Material Engagement Theory a cui questo lisergico riassunto si riferisce è quella pubblicata in L. Malafouris, *Mind and Material Engagement*, in *Phenomenology and the Cognitive Sciences*, Vol. 18 (2019), pp.1-17. Per gli effetti del rapporto tra gli esseri umani e loro strumenti nella teoria del Material Engagement si veda anche D. Ihde, L. Malafouris, *Homo Faber Revisited: Postphenomenology and Material Engagement Theory*, in *Philosophy and Technology* 32 (2019), pp. 195-214.

¹⁷ L. Malafouris, *The Brain-Artifact Interface (BAI): a Challenge for Archaeology and Cultural Neuroscience*, in *Social Cognitive and Affective Neuroscience*, 2-3 (5 Jun 2010), pp. 264-73.

riferisce alla capacità delle Interfacce Cervello-Artefatto di alterare il rapporto tra i soggetti e tra soggetti e ambiente, e di trasformare la struttura delle informazioni ambientali in modo da massimizzarne la salienza in una maniera impossibile ai soli organi di senso quando sprovvisti di potenziamento protesico. Gli effetti di tipo plastico sono quelli più corporali e strettamente neurologici: per Malafouris, queste interfacce possiedono la capacità di causare dei *rewiring* neurali, sia alterando circuiti neurologici pre-esistenti sia aprendo nuovi percorsi all'interno del cervello, testimonianze biologiche della plasticità cerebrale indotta da cultura e attività. Inoltre, le attività cognitive influenzate dalle tecnologie appartenenti alla grande famiglia delle BAI vengono radicalmente trasformate, alterando i percorsi di attivazione delle diverse aree cerebrali coinvolte in tali attività. Questo significa che la nostra trasformazione ad opera delle BAI non avviene solo a livello neurale, ma anche al livello delle pratiche e degli strumenti. Infine, vi è l'aspetto temporale: quell'effetto di integrazione e coordinazione di cui le BAI sono foriere può influenzare processi che avvengono su scale temporali estremamente diverse, da quella neurale a quella evolutiva, passando per i tempi dei processi corporei e di quelli culturali.

Se il concetto di interfaccia è in grado di descrivere i *media* e gli strumenti tramite cui gli esseri umani si relazionano all'ambiente, con i quali l'animale-uomo si codetermina, e attraverso alcuni dei quali riconfigura completamente il proprio rapporto con se stesso e con il mondo, allora quel concetto descrive una porzione della realtà della nostra specie che costituisce una delle condizioni a priori della realtà umana come la conosciamo. Quando i nostri strumenti e il nostro linguaggio si mostrano a noi come interfacce, non sono solo essi a prestarsi all'attraversamento della nostra operatività, ma noi stessi appariamo come evoluti appositamente per attraversare interfacce estremamente differenti tra loro, con prestazioni altrettanto differenti, riconfigurando l'intera storia umana come una storia di relazioni tramite interfacce.

Questa strada ci appare, nell'utilizzo dell'interfaccia come strumento di analisi della realtà e di analisi delle interfacce stesse, foriera di pericolo. Parafrasando Anne

Friedberg nel suo brillante testo sulla vetrina e il postmoderno¹⁸, se il termine interfaccia dovesse giungere a descrivere gli elementi grafici di un software, la scelta iconologica di un sistema operativo, una lingua parlata in un tempo e uno spazio circoscritti, il linguaggio biarticolato, l'intera gestualità umana, un mestolo da cucina, una finestra, un insieme di regole e un effetto fisico, rischieremmo di trasformare il termine in una *buzzword*. Si tratta di un rischio insito nella primissima estrazione del termine dal suo originale contesto chimico: intendere per interfaccia ciò che sta in mezzo a due realtà determinando, almeno in parte, il rapporto tra di esse è destinato a trasformare "interfaccia" in un sinonimo meno altisonante del *medium*. Non di meno, l'utilizzo dell'Interfaccia Cervello-Artefatto di Malafouris è prezioso e di grande potenza per l'archeologia cognitiva proprio perché ne investiga le caratteristiche e gli effetti, più dell'ontologia. L'interfaccia è la chiave di lettura di una realtà concepita in termini di trasferimento di informazione (non necessariamente in forma elettromagnetica), informazione in grado di trasformarsi, ripetersi o disperdersi. L'informazione numericamente misurabile, quella espressa in bits, è una funzione inversa delle diverse probabilità che si verifichi uno tra differenti eventi possibili¹⁹. Se trattare l'intera storia umana come storia di rapporti con l'interfaccia è eccessivo, è possibile d'altra parte considerare il nostro tempo come un tempo di interfacce. È il tempo in cui gli strumenti tecnologici vengono progettati *come* interfacce, in cui le tecnologie indossabili e i dispositivi di tracciamento oculare interagiscono con il nostro sguardo per farsene attraversare e, nel farsi attraversare, trasformare ciò che le attraversa, è il tempo in cui oggetti d'uso domestico che fino a pochi anni fa attendevano il momento della loro

¹⁸A. Friedberg, *Window Shopping. Cinema and the Postmodern*, University of California Press, Berkley (1994).

¹⁹La definizione, alla base dell'intera teoria dell'informazione, fu coniata da Claude Shannon in C.E. Shannon, W. Weaver, *The Mathematical Theory of Communication*, stampato per la prima volta nel 1949. L'esempio più semplice per chiarire il legame tra eventi e informazione è quello del lancio della moneta non truccata: i due risultati, testa e croce, sono equiprobabili, per cui la descrizione del lancio della moneta può essere ridotta a un singolo bit, che con i suoi valori di 0 e 1 informerà se il risultato del lancio sia stato testa o croce. Man mano che la probabilità di un evento diminuisce sarà necessario far uso di un numero sempre maggiore di bit per poterla esprimere, aumentando l'informazione che tale evento porta con sé. Questo ha l'effetto di rendere più efficiente la comunicazione di eventi che siano più frequenti.

chiamata sono ora, attraverso le loro interfacce, in uno stato di costante scambio di informazione con chi vive nelle loro vicinanze.

II. Storia e archeologia dell'interfaccia grafica

Non è da dibattersi il fatto che la più diffusa e variegata tipologia di interfacce digitali nel momento della stesura del presente lavoro, così come nei decenni che l'hanno preceduto, sia di natura grafica. Miliardi di individui si confrontano giornalmente con schermi che fungono da cornice a innumerevoli ecosistemi di immagini dinamiche, le cui complesse realtà sfidano le passate teorie dei media e dell'immagine¹. Tuttavia, questa immensa iconosfera digitale non è composta di puro caos. Schemi e paradigmi ricorrenti sono necessari alla fruizione di interfacce sempre nuove, segni e icone dotate di senso comune dalle convenzioni o da altri meccanismi permettono agli utenti di riconoscere agentività già fruite in passato, e protocolli imposti da aziende e distributori rendono riconoscibili i prodotti e i software grazie alle differenze stilistiche che generano. Se l'interfaccia agisce da codice sul piano semiotico, è naturale che gli elementi di tale codice debbano essere apprendibili e riproponibili. Da principio l'interfaccia grafica (o Graphical User Interface, spesso abbreviata in GUI), composta di icone, menù, documenti e cartelle e gestita tramite un mouse, si è evoluta a partire dagli editor di testo di Xerox PARC (Palo Alto Research Center)². Come già accennato in questo testo, questo paradigma grafico – che prende il nome di WIMP, dalle iniziali dei suoi elementi Window, Icon, Menu e Pointer – tende a circondare un contenuto con una finestra adornata da serie di strumenti finalizzati alla segmentazione, modifica e controllo di tale contenuto. L'idea precede di gran lunga l'interfaccia e lo stesso concetto di

¹ Definire lo studio dei nuovi media un campo sarebbe estremamente riduttivo, e qualunque elenco di fonti sarebbe qui insufficiente a fornire anche solo un panorama della quantità e varietà di discipline ad oggi impegnate nell'analizzare l'enorme complessità e mole di trasformazioni che gli schermi digitali hanno prodotto nel nostro rapporto con le immagini. Si proporranno qui dunque, a titolo di esempi su tutti, il testo di R. Eugeni, *La condizione postmediale*, La Scuola, Brescia (2015) e gli studi sulle an-icone (immagini che negano la loro iconicità rifiutando la cornice e ponendosi come presenziali e immediate) condotti dal gruppo di ricerca AN-ICON guidato dal professor Andrea Pinotti, presso l'Università Statale di Milano.

² Seppur il fatto sia genericamente noto, la fonte primaria dell'autore era originariamente un esaustivo articolo dal titolo *PARC History, PARC – A Xerox Company*, presente sul sito del PARC. Dal 2023, tuttavia, con il cedimento del PARC da parte di Xerox SRI, il sito e l'articolo non sono più reperibili. Si riporta dunque la voce relativa al PARC presente sull'enciclopedia britannica alla data del 30/09/2024: <https://www.britannica.com/topic/PARC-company>.

computer, anche in una sua accezione sufficientemente ampia da includere il protocomputer di Charles Babbage³ e persino le macchine calcolatrici di Leibniz e Pascal⁴, e si presta dunque a una ricerca archeologica su più fronti e attraverso svariati secoli e diverse storie, apparentemente lontane tra loro.

L'utilizzo di una finestra-schermo, sulla quale fiumi di inchiostro sono stati stesi tanto dagli storici dell'arte quanto dagli estetologi e teorici della fotografia, risale al Rinascimento. La tecnica della *perspectiva*, ideata da Brunelleschi e raffinata da Leon Battista Alberti, ha significato letterale di “sguardo attraverso” e prevedeva l'utilizzo di una finestra con griglia per simulare un punto di vista unico e oggettivo, aiutando il pittore a rappresentare correttamente le proporzioni di un corpo umano⁵. La griglia viene applicata, nella teoria di Alberti, tramite un velo trasparente quadrettato da far aderire alla finestra “come tecnica di rilevazione, in grado di scomporre il corpo umano in riquadri regolari, che ne agevolano la riproduzione in scala, pezzo per pezzo.”⁶



Fig. 2: Albrecht Dürer, *Macchina prospettica* (*Underweysung der Messung*, 1525).

³ Per una storia dei primi calcolatori, ossia la macchina differenziale e la macchina analitica di Charles Babbage, si rimanda a R. A. Hyman, *Charles Babbage – Pioneer of the Computer*, Princeton University Press, Princeton, NJ (1982).

⁴ M.J. Beeson, *The Mechanization of Mathematics*, in C. Teucher (ed.), *Alan Turing: Life and Legacy of a Great Thinker*, Springer (2004), p. 82.

⁵ A. Friedberg, *The Virtual Window. From Alberti to Microsoft*, The MIT Press, Cambridge, MA (2006).

⁶ B. Grespi, L. Malavasi, *Dalla parte delle immagini – Temi di cultura visuale*, McGraw Hill Education Italy, Milano (2022), p.109.

Barbara Grespi e Luca Malavasi hanno sottolineato, in questo processo, l'importanza della trasparenza di tale finestra. Trasparenza deriva dal latino *inspicio trans*, ossia guardare attraverso. La capacità di lasciarsi attraversare è già emersa come fondamentale nelle interfacce, discutendone la definizione. La trasparenza, tuttavia, è un meccanismo molto diverso dall'invisibilità: la trasparenza può trasformare, deformare, privare o arricchire ciò che la attraversa. Far guardare attraverso di sé non significa impossibilità di essere visti, ma capacità di divenire canale. Significa concedere passaggio a uno sguardo, ma significa anche offrirsi come direzione di esso e come punto di passaggio lungo il quale arricchirlo di informazione o di finalità. L'azione di trasporto di messaggi culturali proposta da Manovich avviene necessariamente grazie a un certo livello di trasparenza.

Una griglia sovrapposta a un'immagine grazie alla sua trasparenza è evidentemente più che una sequenza di quadrati: è un insieme di linee e di punti di intersezione tra le linee che compongono la griglia. Questi due elementi sono eccellenti riferimenti, in grado di fornire al lettore della griglia informazioni estremamente precise: cosa sia a destra o a sinistra di una determinata linea, quanto distino due dettagli posti su due punti di intersezione, e quale sia la posizione di diversi elementi dell'immagine in un possibile sistema di coordinate, ottenuto semplicemente contando le linee: sappiamo tutti quanto, un secolo e mezzo dopo, una forma più precisa di tale idea abbia sconvolto il nostro mondo grazie al lavoro di René Descartes e all'introduzione del piano cartesiano. L'idea di far uso di linee e punti di intersezione come riferimenti in un'immagine è estremamente intuitiva, ma sovrapporli a un campo visivo fu un'idea brillante ed estremamente duratura. Un'immagine digitale, nel senso più ampio di "ciò che appare sullo schermo di un computer", non è altro che una griglia di coordinate. La trama di tale griglia è tanto fine che ci è solitamente impossibile, a occhio nudo, farne un uso consapevole e cosciente. È tuttavia difficile dubitare che si tratti di un'applicazione su scala microscopica del principio di misurabilità generato dalla griglia.

Ma fu proprio un giovane contemporaneo di Descartes, uno ben meno noto e studiato del matematico francese o di Leon Battista Alberti, a trasformare per primo

le tecniche di misura in dispositivi “interattivi”, grazie all’aiuto più inaspettato che potesse ricevere. Tale idea ebbe infatti origine nell’astronomia.

Nel XVII secolo gli astronomi erano soliti sperimentare con lenti e telescopi di diverse forme e fatture. Divenuti maestri nello sfruttare le proprietà deformanti di dispositivi trasparenti per ottenere diversi ingrandimenti di immagine, non era raro che combinassero diversi tipi di lente per osservare diversi corpi celesti. Ed è raccontando una di queste prove che William Gascoigne, nel 1642, racconta in una lettera all’amico William Oughtred di aver incontrato una sorpresa ricca di significato:

This is that admirable secret, which, as all other things, appeared when it pleased the All Disposer, at whose direction a spider’s line drawn in an opened case could first give me by its perfect apparition, when I was with two convexes trying experiments about the sun, the unexpected knowledge [...] and therefore resolved that if I [...] placed a thread where that glass would best discern it, and then joining both glasses, and fitting their distance for any object, I should see this at any part that I did direct it to. And so I found it (if you please to try it, the convex of a perspective and a good convex spectacle glass will let you fully see it), which the next night’s trial confirmed strangely accurate and ready for finding the altitude of any small star.”⁷

Il dispositivo che Gascoigne, con l’aiuto di un aracnide, ha ideato in quell’occasione è destinato a rivoluzionare l’astronomia e molte scienze successive: è la nascita del micrometro filare. Il filo di ragno che Gascoigne scorse, fortuitamente teso in una scatola rimasta aperta proprio sulla linea di puntamento del suo telescopio durante un tentativo di osservazione del Sole⁸, ispirò un vero e proprio balzo in avanti nelle tecniche di misura tramite tecnologie trasparenti: l’idea di un puntatore mobile pilotabile posto di fronte alla lente di vetro. Se l’idea alla base di uno strumento può apparire semplice, poiché il suo funzionamento non appare dissimile a quello di un

⁷ W. Gascoigne, *Letter to William Oughtred, c. February 1641*, in D. Sellers, *In Search of William Gascoigne*, Astrophysics and Space Science Library, Vol. 390, New York (2012), p.184.

⁸ Ibid.

più semplice parallelografo e facilmente realizzabile su carta, i suoi requisiti tecnologici non sono altrettanto facili da soddisfare. Il filo della tela di ragno, come notò Gascoigne, era contemporaneamente in grado di mantenersi teso senza cedere alla tensione e di rimanere a fuoco durante l'osservazione, e grazie a queste sue caratteristiche permetteva di calcolare con grande precisione la posizione relativa degli oggetti celesti e, confrontando le loro posizioni, di misurarne l'altitudine. Proprio per questo il ragno della sua stanza non fu un'eccezione, ma il pioniere di una forma di artigianato aracnide destinato a proseguire per più di due secoli. Se ne ritrovano tracce almeno fino al 1883, anno in cui la voce *Micrometer* dell'Enciclopedia Britannica riporta diversi riferimenti all'uso della ragnatela:

The modern filar micrometer has now assumed forms of five types.

Type A. Micrometers in which there are two webs, each movable by a fine screw with a divided head. This is the usual English form of filar micrometer. meters.

Type B. Micrometers in which one web is movable by means of a fine screw with a divided head, and the other by a screw without a divided head. The latter screw, in ordinary use, is only employed to change the coincidence-reading of the two webs, for eliminating the errors of the micrometer screw. This is the ordinary German form of micrometer as originally made by Fraunhofer and since by Merz, and employed by the Struves and other principal Continental astronomers down to the present day.

Type C. A similar form of micrometer to B, except that the coincidence-point cannot be changed, there being no second screw to alter the position of the fixed web.

Type D. A micrometer somewhat similar in general construction to form B, except that, in addition to means of changing the zero point, there is a screw head by which a fine movement can be given to the whole micrometer box, in the direction of the axis of the micrometer screw. This is the modern form of micrometer as constructed by Repsold.

Type E. Micrometers fitted with two eye-pieces for measuring angles larger than the field of view of an ordinary eye-piece.⁹

⁹ *Micrometer*, in *Encyclopaedia Britannica – A Dictionary of Arts, Sciences and General Literature*, Ninth Edition, Volume 16, Charles Scribner's Sons, New York (1883). L'autore ringrazia calorosamente il contributo dell'INAF e del Dott. Mario Alessandro Carpino per l'accesso a questa fonte bibliografica.

La voce riporta anche un estratto di un piccolo manuale di David Gill, nel quale è ben illustrata la procedura per creare micrometri filare con l'aiuto di assistenti aracnidi, appropriatamente accompagnati nella procedura di stesura dei fili.



Fig. 3: Micrometro filare, o Micrometro M, conservato presso l'osservatorio di Brera.

Il micrometro filare è di certo uno splendido proto-esempio di interfaccia per strumenti ottici, ed è all'origine di una catena di interfacce successive che hanno condotto a quelle a noi più familiari oggi, caratterizzate dalla possibilità di essere manipolate secondo la volontà di chi ne fa uso. Il micrometro filare non impedisce in alcun modo la visione di ciò che il telescopio inquadra. Esso, tramite la sua trasparenza, potenzia quella visione arricchendola di informazione senza per questo inficiarla minimamente. Lo sguardo che la attraversa viene arricchito nel preciso istante di quell'attraversamento. Questo effetto di arricchimento è la causa di ogni

interfaccia digitale, e tra il XVIII e il XIX secolo ha accompagnato innumerevoli invenzioni di grande rilievo per l'astronomia: dalla macchina parallattica di Grindel, i cui cerchi graduati permettono di determinare la posizione di pianeti e comete, ai primi spettroscopi a cristallo che dalla seconda metà del XIX secolo hanno permesso (grazie al contributo della fotografia) di identificare gli elementi chimici che compongono le stelle, ai successivi “strumenti dei passaggi” e al circolo meridiano di Ertel, che facevano uso di puntatori mobili e di un cronografo per calcolare l'altezza di un astro sul meridiano del luogo¹⁰.

Il filo in tensione fu un'idea di tale successo da sopravvivere al filo stesso. Molto tempo dopo l'invenzione del micrometro filare, la sua eredità fu raccolta dai mirini balistici, e in seguito da diverse categorie di puntatore. Nonostante le loro forme siano molte e diverse tra loro, e spesso ben differenti da una semplice croce di due segmenti, gli indicatori visuali che identificano il centro di un mirino hanno mantenuto il nome di una delle prime e più semplici forme, il *crosshair*. Il nome si è diffuso in differenti ambiti, incluso quello dell'informatica, e ha finito per significare qualunque puntatore di forma simile a una croce che serva da mirino di un'unità mobile, come ad esempio avviene nei dispositivi di precisione del settore biomedicale¹¹.

Esistono, d'altra parte, forme di trasparenza di estrema efficacia nell'individuare e trasformare l'oggetto dell'osservazione senza che ciò sia necessariamente l'intenzione cosciente dell'osservatore. Se un sistema di puntamento, in ogni sua forma, stabilisce una gerarchia di priorità all'interno del suo campo di movimento

¹⁰ Diversi esemplari dei dispositivi elencati, incluso il micrometro M, possono essere osservati presso il Museo dell'Osservatorio Astronomico di Brera.

¹¹ K. Meehan et al., *Crosshair, Semi-automated Targeting for Electron Microscopy with a Motorised Ultramicrotome*, in *eLife*, 11 (2022), online.

S. W. Hadley, K. Lam, *Light Field and Crosshair Quality Assurance Test Using a Simple Lens System*, in *Medical Physics*, Vol. 33 (2006), pp. 930-932.

o della sua inquadratura, l'inquadratura stessa deve essere messa in discussione, poiché nessuna finestra è un oggetto neutrale, e non tutto ciò che si chiama finestra presenta le stesse caratteristiche. Anche la finestra è un dispositivo originariamente causato dalla sua trasparenza. Tuttavia, a differenza di altre tecnologie architettoniche finalizzate a permettere l'osservazione come il merlo o la faretra, la finestra è in genere ostruita fisicamente da una superficie che permette il passaggio della luce, una superficie che può essere rimossa tramite l'apertura della finestra stessa. È naturale concepire una finestra come aperta o chiusa. Esistono però finestre d'eccezione, categoria in cui rientrano a pieno titolo quelle di un'interfaccia grafica digitale, così come dell'interfaccia schermica che la veicola. La principale e più significativa di queste finestre, prima di quella digitale, è senza dubbio la vetrina.

La vetrina è un dispositivo di enorme impatto sull'ambiente in cui si insedia. Nel suo testo sul cinema e il postmoderno dal titolo *Window Shopping*¹², Anne Friedberg identifica come strumento di esposizione del reale così potente da sostituire il turismo urbano in senso classico con l'escursione commerciale: il *flâneur* e la *flâneuse* si trasformano negli *shoppers*, dallo sguardo ingabbiato nei riquadri di vetro posti a far mostra dei prodotti in vendita dentro un *mall*. Nel fare ciò, la superficie della vetrina giunge ad assumere il ruolo di una vera e propria cornice: essa separa nettamente ciò che esibisce dal mondo circostante, creando una nuova gerarchia tra ciò che merita di essere osservato e ciò che deve invece essere completamente obliato. Si tratta di una cornice forse anche più violenta di quella classica, che impone (o cerca di imporre) a un segmento di realtà materiale quella riconfigurazione a "dominio sovrano" propria del quadro come la espone Georg Simmel, teorico della cornice¹³. La cornice attrae e ingabbia in sé lo sguardo, facendosi terza parte trasformativa in un'osservazione che avviene tra il soggetto e l'oggetto, una cornice in grado di "accecare completamente" l'osservatore

¹² A. Friedberg, *Window Shopping – Cinema and the Postmodern*, University of California Press, Berkeley (1994).

¹³ G. Simmel, *L'ansa del vaso. Saggio di estetica*, in G. Simmel, *Stile moderno*, B. Carnevali, A. Pinotti (a cura di), Einaudi, Torino (2020), pag.306.

(“enforcing a blindness to a range of urban blights the homeless, beggars, crime, traffic, even weather.”¹⁴), e insieme separando il mondo dell’osservatore dal mondo in cui trascina il suo sguardo e il suo pensiero.

La vetrina svolge però anche un’altra funzione: quella di attirare fisicamente a sé l’osservatore. A differenza della finestra, essa è parte attiva del processo di osservazione, e può richiamare a sé i corpi. Come un magnete, che allinea le cariche del metallo prima di poterla attirare verso di sé, la vetrina allinea l’attenzione dell’osservatore prima di attrarlo fisicamente. La vetrina è, nel museo costituito dal *mall*, un quadro che deve competere con gli altri quadri. Si tratta di un tipo di dispositivo, in un certo senso, opposta a quello messo a punto da William Gascoigne, che permetteva all’osservatore di spostare l’interfaccia per osservare oggetti differenti: la vetrina è un’interfaccia che muove diversi soggetti a osservare un oggetto immobile. Opposta non significa però incompatibile, non significa cioè che le due prestazioni dell’interfaccia non possano coesistere.

Un esempio della potenza della sinergia tra queste due facoltà può essere rintracciato in un film d’animazione di Project Itoh, pseudonimo dello scrittore giapponese Satoshi Itō, dal titolo italiano *L’Organo Genocida*¹⁵. La pellicola adatta il romanzo di Itō, i cui protagonisti sono membri delle Forze Speciali dell’esercito degli Stati Uniti equipaggiati con tecnologie di avanguardia, impiegati in un 2022 ucronico e su una Terra più tecnologicamente sorvegliata di quella che conosciamo a causa della minaccia del terrorismo. La scena di nostro interesse ha luogo al confine tra India e Pakistan, in una polverosa città semidistrutta dalla guerriglia, dove all’interno di un albergo si nascondono diversi leader di alcune minoranze in conflitto che le Forze Speciali statunitensi intendono catturare.

Prima di lanciarsi in missione nelle loro capsule, gli uomini delle Forze Speciali vengono equipaggiati con le “alternative”, delle speciali lenti a contatto liquide dotate di avanzatissime tecnologie di Realtà Aumentata. Questi dispositivi,

¹⁴ A. Friedberg, *Window Shopping. Cinema and the Postmodern*, University of California Press, Berkley (1994).

¹⁵ S. Itō, *L’Organo Genocida* (虐殺器官, *Gyakusatsu kikan*), regia di S. Murase, Geno Studio, Suginami, Tokyo (2017).

sincronizzati alle tute speciali dei soldati, sono in grado di visualizzare diverse informazioni inviate dal comando centrale, manipolare le immagini visualizzate tramite movimenti della lingua e allo stesso tempo analizzare e manipolare automaticamente l'immagine di fronte al soldato, identificando potenziali minacce e compiendo operazioni di taglio e ingrandimento in tempo reale. Ciò che rende rilevante la scena per questa riflessione è la scelta del regista di proporla buona parte in soggettiva. All'interno dell'inquadratura, ci è possibile osservare con gli occhi dell'agente Clavis Shepherd tutto ciò che la sua interfaccia indossabile mostra durante le sparatorie che accompagna la scalata dell'albergo. Per l'intera durata della soggettiva, lo schermo è dominato da un enorme mirino mobile di forma circolare, dotato di alcuni indicatori numerici sul bordo esterno, che identifica i movimenti e le minacce nel campo visivo del soldato. A ogni scatto del mirino, immediatamente il fucile dell'agente segue il puntatore: l'interfaccia si fa giudice ed emette sentenze capitali nei confronti dei guerriglieri, e il nostro sguardo, insieme al fucile del protagonista, esegue quelle condanne inseguendone la danza mortale. La scena permette di osservare, su un livello pienamente fenomenologico, l'azione sinergica dell'interfaccia del primo micrometro filare e la forza coercitiva dell'interfaccia di una vetrina, rendendo lo spettatore insieme carnefice e vittima di una sequenza la cui violenza è guidata dall'interfaccia stessa.



Fig. 4: Interfaccia delle lenti AR di Clavis Shepherd. *L'organo genocida*, Satoshi Itoh.

L'interfaccia, assumendo simultaneamente i ruoli di strumento di misura e agente in grado di istituire classificazioni assiologiche (cosa è degno di osservazione e cosa no o, nell'ultimo esempio, chi dovrà vivere e chi dovrà morire), diviene un tipo speciale di mentore permissivo. Non insegnante, ma guida a cui fare riferimento, presentificazione impersonale – in quanto numerizzata e apparentemente relativa all'oggetto da osservare e non al soggetto osservante – di una determinata proposta di azione, ottima secondo i criteri di una determinata assiologia – come avviene nel caso della vetrina.

Come storicamente è avvenuto per molte tecnologie di grande impatto, la prima applicazione della sinergia appena discussa ha visto i suoi natali in ambito bellico. Nel 1947 Paul Fitts, psicologo dell'aviazione statunitense impiegato nello studio del miglioramento delle prestazioni dei piloti tramite l'ingegneria dei velivoli, stese un rapporto all'aeronautica militare sulla complessità del pilotaggio degli aerei da caccia:

A buddy of mine was flying at 1,500 feet. He looked up and there was a bunch of trees in front of him. The funny thing was, after it was all over and the plane had crashed, he felt and saw that his bones were all OK and said, 'Well, I guess I must have hit a tree.' What had happened was that he had misread his altimeter. He was actually flying at an indicated altitude of 500 feet, which was just the altitude of the terrain at that point.¹⁶

Il rapporto, all'epoca altamente secretato, venne dischiuso al pubblico molti anni dopo. In un articolo che riassume i risultati del rapporto, il progettista statunitense affermò che l'aereo da guerra e il suo pilotaggio avevano raggiunto il picco della loro possibile complessità: Fitts riteneva che, da quel punto in poi, l'aggiunta di qualsivoglia ulteriore dispositivo, comando di bordo o indicatore parametrico avrebbe finito per incidere negativamente sulle prestazioni dei già sovraccarichi piloti. La proposta da lui avanzata era quella di progettare sistemi di visualizzazione

¹⁶ P.M. Fitts, *Psychological Research on Equipment Design - Report no. 19*, Army Air Forces Aviation Psychology Program Research Reports, Washington D.C. (1947).

secondo le abilità percettive umane¹⁷. Secondo quanto riportato nel suo rapporto, la divisione dell'attenzione dei piloti tra gli strumenti della cabina di pilotaggio e l'ambiente esterno al velivolo è la causa un numero ormai inaccettabile di incidenti. Fitts non era un'eccezione: la sua soluzione fu influenzata dalla ricerca della RAF e della Luftwaffe, che fin dai primi anni '40 erano alla ricerca di un modo per potenziare le prestazioni dei piloti, le cui capacità faticavano a tenere il passo con un fitto insieme di informazioni da monitorare durante il volo, a cui si era appena aggiunta un'interfaccia completamente nuova: il radar.

La soluzione proposta da Fitts, il cui successo è ad oggi incontestato, fu quella di visualizzare direttamente sul vetro della cabina buona parte delle informazioni normalmente disposte sugli indicatori del cruscotto tramite un dispositivo di proiezione, permettendo ai piloti di mantenere la testa sollevata e l'attenzione sull'esterno mentre consultavano i valori più frequentemente necessari al mantenimento della crociera. Il dispositivo prese il nome di Head-Up Display, ed è ancora noto con l'acronimo HUD. L'interfaccia di Fitts, che circonda un contenuto con una serie di strumenti utili a navigarlo, anticipa di un ventennio l'interfaccia grafica. La costante visualizzazione di informazioni salienti ai lati di uno schermo o del campo visivo è rimasta una caratteristica quasi imprescindibile delle interfacce della maggior parte dei software di simulazione, o che necessitano la visualizzazione del mondo fisico. Dalla fotocamera digitale, la cui interfaccia ha ereditato e posto a cornice i pulsanti delle vecchie videocamere manuali, passando per l'immensa (seppur estremamente formulare) sfera delle interfacce videoludiche con le loro barre di vita, minimappe radar e indicatori di svariati valori negli angoli dello schermo, fino agli ultimi eredi di quella tradizione, i dispositivi di Realtà Virtuale e Aumentata.

¹⁷ P. M. Fitts, *Psychological Research on Equipment Design in the AAF*, in *American Psychologist*, Vol. 2 (1947), pp.93-98.

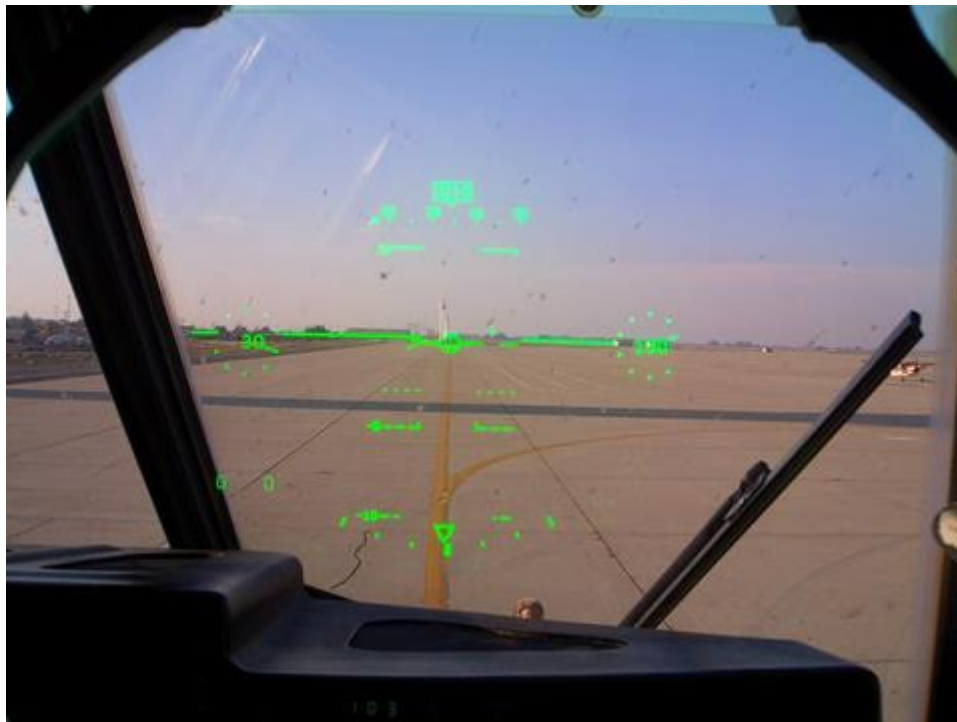


Fig. 5: Head-Up Display per il copilota di un aereo da trasporto Super Hercules.

Nell'atto di proiettare sul vetro le informazioni che permettono di integrare con informazioni numeriche il mondo esterno a quel vetro, Fitts trasforma tutti i valori forniti dal cruscotto in un'interfaccia sul mondo. Ciò che era stato progettato per ridurre il carico di attenzione, permettendo al pilota di intenzionare simultaneamente il mondo esterno e le informazioni fornite dai suoi sensori, finì per rendere quegli stessi numeri il filtro attraverso il quale il mondo viene rielaborato in tempo reale. Ogni nuvola, ogni trasformazione del terreno, ogni virata viene riconfigurata in tempo reale dai valori visualizzati sul parabrezza del velivolo, e quei valori vengono costantemente alterati dai comandi che il pilota impartisce al suo aereo da caccia.

L'HUD si muove sulla traccia del micrometro di Gascoigne e della griglia albertiana. Il telescopio equipaggiato con strumenti di misura e la cabina di pilotaggio potenziata tramite proiezione sono dispositivi di arricchimento informativo del mondo fisico: l'interfaccia emerge dall'interposizione tra osservatore e osservato di una realtà terza, un mondo-informazione meno ricco

della realtà circostante ma costituito di informazione mirata e saliente, disposto secondo una logica improntata all'efficienza e con obiettivi prestabiliti. Lo scopo di queste interfacce è precisamente quello di essere esperite mentre si osserva ciò che di materiale esiste al di là della loro superficie.

Due anni prima del rapporto di Paul Fitts, prima dell'invenzione dei transistor, aveva inizio la storia del computer come lo intendiamo oggi, agli inizi del terzo millennio. Il primo passo fondamentale porta la firma di Vannevar Bush, fondatore della National Science Foundation e dell'Office of Scientific Research and Development degli Stati Uniti, nonché uno dei principali contributori all'invenzione del radar¹⁸. L'ingegnere, specializzato in teoria del segnale e trasmissione¹⁹, pubblicava nel luglio del 1945 all'interno di un saggio la visione del futuro di innumerevoli tecnologie la cui velocità di evoluzione si era moltiplicata durante la guerra, e di come queste tecnologie avrebbero potuto servire l'umanità e dare lavoro agli scienziati una volta che il conflitto globale fosse terminato.

Nel testo, dal titolo *As we may think*²⁰, Bush anticipava i dispositivi che avrebbero visto la luce grazie all'evoluzione della fotografia, dei dispositivi a tastiera, dei dischi in cera o su pellicola: fotocamere indossabili di dimensioni inferiori al centimetro, la cui attivazione richiede solo un gesto delle dita (anche se, nell'assenza di schermi tattili, ciò sarebbe dovuto avvenire grazie a piccoli fili dentro la manica), e senza necessità di attendere lo sviluppo delle foto. Secondo Bush, avrebbe stato possibile modificare le immagini grazie a puntatori simili a penne elettriche, come quelle utilizzate all'epoca su carta iodata, il cui passaggio poteva lasciare segni scuri alterando lo stato chimico dello iodio.

¹⁸ Vannevar Bush, in *Encyclopedia Britannica*, <https://www.britannica.com/biography/Vannevar-Bush>, ultima visita 03/10/24.

¹⁹ La tesi di dottorato di Vannevar Bush presso l'MIT, dal titolo *Oscillating-current Circuits: An Extension of the Theory of Generalized Angular Velocities, with Applications to the Coupled Circuit and the Artificial Transmission Line*, è reperibile presso il sito del Mathematics Genealogy Project, a <https://genealogy.math.ndsu.nodak.edu/id.php?id=84507>, ultima visita 03/10/24.

²⁰ V. Bush, *As We May Think*, in *The Atlantic Monthly* (Luglio 1945) è stato digitalizzato e reso disponibile in rete nel 1994 da D. Duchier, per conto dell'Università di Ottawa, a <http://www.cs.unibo.it/cianca/wwwpages/dd/bush.pdf>, ultima visita 03/10/24.

Il saggio descrive anche l'evoluzione del processo di scrittura, immaginando come lo scrittore del futuro avrebbe potuto liberarsi dalla necessità di utilizzare carta e penna o macchina da scrivere e parlare direttamente con il dispositivo di registrazione del testo, la cui forma ipotetica era quella di un Vocoder invertito, che avesse al posto dell'altoparlante un microfono le cui vibrazioni avrebbero attivato i tasti di una macchina da scrivere automatica. Un ricercatore avrebbe potuto usare entrambi i dispositivi, la macchina fotografica indossabile e la macchina da scrivere vocale, ed esse avrebbero utilizzato gli orari di ogni registrazione per connettere automaticamente le foto alle annotazioni vocali appuntate durante le indagini scientifiche. Tutte queste informazioni sarebbero state poi miniaturizzate, con la possibilità di venir proiettate successivamente.

Il testo riporta anche il futuro di macchine calcolatrici che, a differenza delle macchine fotografiche, non sarebbero rimpicciolite in dimensioni ma aumentate in potenza. Simili macchine avrebbero potuto compiere operazioni sempre più complesse a velocità sempre più elevate, secondo una procedura senza la quale non era possibile, all'epoca, concepire un calcolatore:

[...] they [le macchine calcolatrici] will be far more versatile than present commercial machines, so that they may readily be adapted for a wide variety of operations. They will be controlled by a control card or film, they will select their own data and manipulate it in accordance with the instructions thus inserted, they will perform complex arithmetical computations at exceedingly high speeds, and they will record results in such form as to be readily available for distribution or for later further manipulation. Such machines will have enormous appetites. One of them will take instructions and data from a roomful of girls armed with simple keyboard punches, and will deliver sheets of computed results every few minutes. There will always be plenty of things to compute in the detailed affairs of millions of people doing complicated things.²¹

²¹ Nonostante esuli dall'obiettivo del presente capitolo, quello di tracciare una storia e un'archeologia mediale dell'interfaccia, l'invisibile lavoro delle donne nella storia del computer e dell'informatica come disciplina è un argomento degno di menzione. A partire da Lady Ada Lovelace, scrittrice del primo programma informatico (S. Charman-Anderson, *Ada Lovelace: A*

La lista dei dispositivi anticipati con grande precisione da Bush prosegue, e non sarà qui riportata per intero. Vi è però una macchina ipotetica, l'ultima, che coniuga buona parte delle capacità di tutte le macchine precedenti. Il dispositivo prende il nome di memex, nome di pura invenzione, ed è descritto come segue:

A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory.

It consists of a desk, and while it can presumably be operated from a distance, it is primarily the piece of furniture at which he works. On the top are slanting translucent screens, on which material can be projected for convenient reading. There is a keyboard, and sets of buttons and levers. Otherwise it looks like an ordinary desk.

In one end is the stored material. The matter of bulk is well taken care of by improved microfilm. [...]

Most of the memex contents are purchased on microfilm ready for insertion. Books of all sorts, pictures, current periodicals, newspapers, are thus obtained and dropped into place. Business correspondence takes the same path. And there is provision for direct entry. On the top of the memex is a transparent platen. On this are placed longhand notes, photographs, memoranda, all sort of things.

When one is in place, the depression of a lever causes it to be photographed onto the next blank space in a section of the memex film, dry photography being employed.

Il dispositivo avrebbe permesso l'indicizzazione del testo e delle immagini, la possibilità di selezionare solo determinate sezioni o pagine di questo o quel libro e l'inserimento di proprie note e commenti personali e la possibilità di accedere a tutte le informazioni riguardanti un singolo argomento di interesse direttamente dal

Simple Solution to a Lengthy Controversy, in *Patterns*, New York, (Oct 2020), online) al ruolo di innumerevoli giovani donne nelle sale computer del primo dopoguerra e di come il loro ruolo sia stato marginalizzato nei decenni (R. Varma, V. Galindo-Sanchez, *Native American Women in Computing*, Idea Group Inc. (2006), online presenta un interessante punto di vista sulla marginalizzazione delle minoranze in questo processo), l'argomento ha ricevuto solo negli ultimi anni una piccola parte dell'attenzione che merita. L'eredità della controversia è ancora ben visibile all'interno delle aule universitarie di qualunque dipartimento di Ingegneria o Informatica, nella disparità numerica tra studenti di sesso maschile e femminile.

testo che ne trattasse. L'idea di ipertesto, così come quella di copia-e-incolla all'interno di un testo editabile strutturato esattamente come il presente lavoro, prendeva forma prima di ogni elettronica digitale. Ma il dispositivo faceva ben di più. Il memex si proponeva come il primo vero sistema schermico completamente interattivo, seppur totalmente ipotetico. In esso, l'agire dell'utente era principalmente imposto a un testo su pagina rettangolare o a fotografie quadrate, tramite una superficie trasparente posta sulla superficie orizzontale di una scrivania.

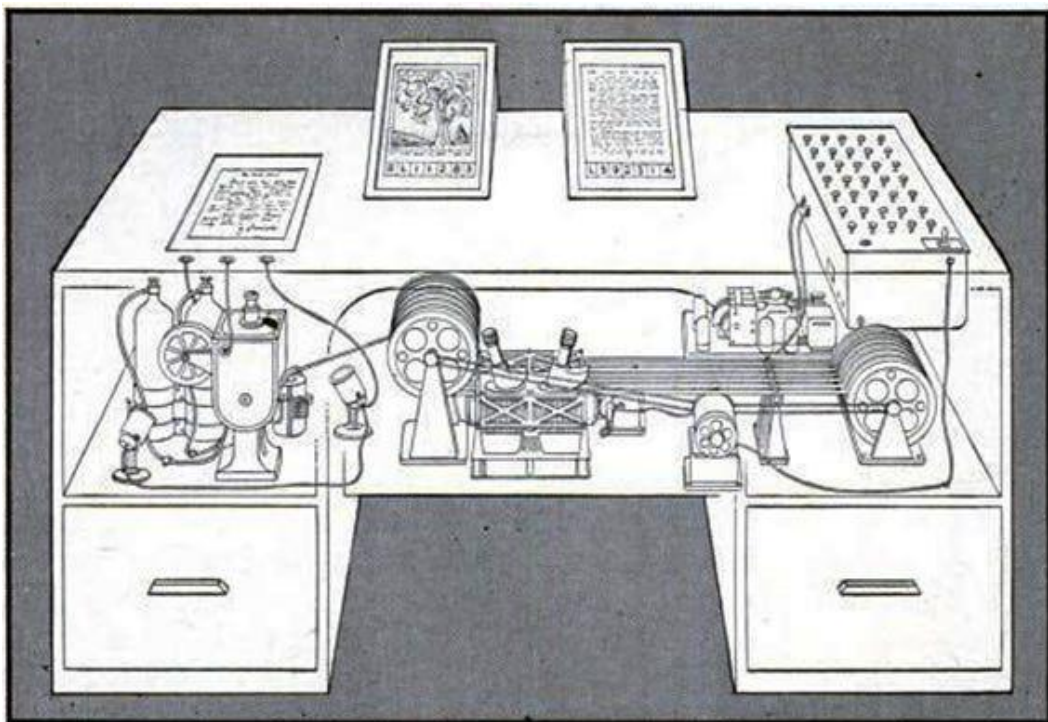


Fig. 6: Illustrazione del memex di Vannevar Bush.

Nella sua forma puramente speculativa, il dispositivo di Bush non poteva certo approfondire gli aspetti delle interazioni da esso proposte: la sola certezza era che una superficie (più precisamente due superfici schermiche, secondo i disegni di Vannevar Bush) avrebbe permesso a una persona di relazionarsi a un vastissimo contenuto informativo altrimenti inaccessibile. Si tratta in tutto e per tutto di un nuovo modo di essere in relazione con l'altra parte di uno schermo, completamente diverso da quello degli schermi dell'epoca (cinematografico e televisivo). La

pellicola del memex è realmente manipolabile sul piano materiale da parte di ogni osservatore. Questa nuova agentività nei confronti di ciò che sta dall'altra parte implica un'attraversabilità bidirezionale dello schermo stesso, uno stato di accessibilità universale e di disponibilità costante del contenuto: la superficie di vetro necessitava di un nuovo tipo di permeabilità, che permettesse di agire sul mondo al di là dello schermo, un mondo di pura informazione, quando l'utente lo desiderava. L'idea che si trattasse dell'interfaccia su un mondo, o quantomeno su una realtà oggettuale, è implicata nella spazialità della descrizione di Bush. Innanzitutto, nel descrivere l'utilizzo del memex, l'utente immaginario scorre ("runs through") un'enciclopedia per trovarvi un articolo, come fosse un grande libro da sfogliare. Dopodiché, in seguito a una seconda ricerca e al ritrovamento un altro oggetto, li lega ("ties") tra di loro. Con questo processo, egli costruisce un percorso di oggetti ("building a trail of many items"). Si tratta di un'agentività che non solo pertiene al mondo fisico, ma che appare anche spiccatamente manipolativa. Il tipo di dispositivo ipotizzato da Bush prevede la capacità di manipolare informazione "attraverso" lo schermo. Non solo tale schermo dovrà permettere l'osservazione attraverso di esso, ma anche di venir attraversato dalle mani di chi opera: uno schermo transmanuale.

La possibilità di agire sul contenuto posto al di là dello schermo è imprescindibile dallo schermo stesso. Si tratta della prima volta in cui un medium, seppur ipotetico, propone un nuovo tipo di immagine, destinata a sconvolgere il mondo: l'immagine pienamente, impunemente presenziale. Il concetto di presenza a cui ci riferiamo qui è quello che deriva dalla definizione di telepresenza. Il concetto di telepresenza introdotto negli anni '70 fu formalizzato, nel contesto della filosofia della coscienza, in un paper di Marvin Minsky nel 1980²². Con presenziale intendiamo

²² M. Minsky, *Telepresence*, 1980. Digitalizzato da Philpapers a <https://philpapers.org/rec/MINT>, ultimo accesso 31/11/2024.

Pur conoscendo la vastità delle implicazioni del concetto e le conseguenze della sua formalizzazione fino alla nascita dei Presence Studies, l'uso fatto del termine in questo testo si limita a indicare la possibilità di compiere operazioni intenzionali e significative a una distanza superiore a quella delle proprie membra e degli strumenti puramente meccanici a essi connessi, secondo un'operatività fondamentalmente manuale. Il termine potrebbe essere sostituito con "teleoperatività", ma l'uso di

qui dunque la possibilità di lasciarsi manipolare, con tutte le implicazioni corporee del verbo, a prescindere dal requisito di distanza. È da considerarsi rilevante in questo contesto la natura spiccatamente manuale di tutti i verbi utilizzati da Bush per descrivere le operazioni da compiersi sul contenuto. Ciò che intendiamo con impunemente, invece, è l'unione di due caratteristiche della manipolazione del memex. La prima, evidentemente nel paradigma tecnologico descritto nel paper, è nell'universalità delle operazioni: qualunque sezione di testo può essere spostata e appropriata più volte, senza che ciò ne deturpi irrimediabilmente la materia. Compiere le operazioni descritte da Bush con la tecnologia del tempo, che si trattasse della re-incisione di un disco o di cera o dell'utilizzo dei più mondani fogli di carta e colla, la diversa materialità di questi supporti non avrebbe permesso la loro totale manipolabilità. Non esiste, nel memex, la possibilità di incollare erroneamente una foto sopra un testo, né quella di inscrivere troppo testo all'interno di un dato supporto. La seconda forma di impunità è invece un pilastro contemporaneo della progettazione software: la normalizzazione delle operazioni e dei risultati. Tutto ciò che viene prodotto può essere letto, e tutto ciò che è stato trovato e ottenuto seguendo una determinata procedura può essere trovato e ottenuto nuovamente. Ogni frammento di testo, pagina di libro o immagine può essere trattata come uno stesso tipo di oggetto e inserito nello stesso contenitore, la cui forma discende da quella del libro e il cui funzionamento discende da quello del cinema.

Nel 1962 Douglas Engelbart, ingegnere dell'SRI, pubblicava *Augmenting Human Intellect: A Conceptual Framework*²³.

Il rapporto steso da Engelbart mirava a istituire un protocollo di ricerca per lo sviluppo di tecnologie orientate all'ottimizzazione e al potenziamento degli esseri

“telepresenza” è sufficientemente diffuso e riconosciuto nel discorso sui media elettrici ed elettronici da preferirlo in quanto riferimento noto.

²³ Il testo è disponibile online, in forma integrale, grazie Doug Engelbart Institute, sia tramite scansioni del documento originale sia tramite una più leggibile trascrizione. <https://www.dougenelbart.org/pubs/augment-3906.html>, ultima consultazione 07/10/24.

umani. Più correttamente, il testo si occupava dello studio e del potenziamento di un sistema chiamato “modello H-LAT/M”: un essere umano (rappresentato dall’H di Human), dotato di linguaggi, artefatti e addestramento (il segmento LAT, “Languages, Artifacts, Training”), organizzati secondo le metodologie (“Methodologies”) che permettono di strutturare l’attività di risoluzione di problemi secondo un approccio *goal-oriented*. Il testo di anticipava con grande precisione diverse funzioni dei moderni computer, sull’onda dell’articolo di Bush:

Let us consider an augmented architect at work. He sits at a working station that has a visual display screen some three feet on a side; this is his working surface, and is controlled by a computer (his "clerk") with which he can communicate by means of a small keyboard and various other devices.

He is designing a building. He has already dreamed up several basic layouts and structural forms, and is trying them out on the screen. The surveying data for the layout he is working on now have already been entered, and he has just coaxed the clerk to show him a perspective view of the steep hillside building site with the roadway above, symbolic representations of the various trees that are to remain on the lot, and the service tie points for the different utilities. The view occupies the left two-thirds of the screen. With a "pointer," he indicates two points of interest, moves his left hand rapidly over the keyboard, and the distance and elevation between the points indicated appear on the right-hand third of the screen.

Now he enters a reference line with his pointer, and the keyboard. Gradually the screen begins to show the work he is doing—a neat excavation appears in the hillside) revises itself slightly, and revises itself again. After a moment, the architect changes the scene on the screen to an overhead plan view of the site, still showing the excavation. A few minutes of study, and he enters on the keyboard a list of items, checking each one as it appears on the screen, to be studied later.

Ignoring the representation on the display, the architect next begins to enter a series of specifications and data—a six-inch slab floor, twelve-inch concrete walls eight feet high within the excavation, and so on. When he has finished, the revised scene appears on the screen. A structure is taking shape. He examines it, adjusts it, pauses long enough to ask for handbook or catalog information from the clerk at various points, and

readjusts accordingly. He often recalls from the "clerk" his working lists of specifications and considerations to refer to them, modify them, or add to them. These lists grow into an evermore-detailed, interlinked structure, which represents the maturing thought behind the actual design.

Prescribing different planes here and there, curved surfaces occasionally, and moving the whole structure about five feet, he finally has the rough external form of the building balanced nicely with the setting and he is assured that this form is basically compatible with the materials to be used as well as with the function of the building.²⁴

La descrizione prosegue a lungo, anticipando molti altri dispositivi e pratiche contemporanee. Chiunque abbia in vita sua avuto a che fare con dei modelli 3D potrà riconoscere molte delle pratiche più basilari della propria attività nella descrizione qui riportata. Alcuni aspetti fondamentali emergono immediatamente, aspetti che si finirono per dimostrarsi pilastri della progettazione *mainstream* di interfacce fino ai nostri giorni. Di certo, quello di più grande interesse per l'analisi di una storia e di un'archeologia dell'interfaccia è la *working surface*. Prima che il concetto di interfaccia venisse applicato all'informatica, Engelbart aveva determinato che l'area di lavoro avrebbe dovuto mantenersi su una superficie per ottimizzare i processi di interazione tra l'uomo e la macchina e la cooperazione tra individui all'interno di un sistema coordinato. L'idea non era frutto della sua sola mente: il testo è fortemente ispirato a quello di Bush, sulla cui visione Engelbart ha lavorato per più di vent'anni²⁵. Ed è proprio per mano di Engelbart e del suo gruppo di ricerca presso l'Augmenting Research Lab, e più tardi presso il Palo Alto Research Center, che lo scritto di Bush venne tradotto in dispositivi materiali. L'invenzione più nota e a noi più familiare che il laboratorio di Engelbart abbia prodotto è senza dubbio il mouse²⁶. La sua prima forma, una scatola dotata di rotelle perpendicolari (le cui rotazioni su una scrivania venivano tradotte nel movimento di un piccolo insetto digitale su uno schermo di vetro), e di un cavo simile alla coda

²⁴ Ibid.

²⁵ F. Turner, *From Counterculture to Cyberculture*, The University of Chicago Press, Chicago (2006).

²⁶ P. E. Ceruzzi, *A History of Modern Computing*, MIT Press, Cambridge, MA (1998).

di un topo (come dichiarò Engelbart scusandosi, tale nome venne scherzosamente attribuito all'oggetto per poi non venir più sostituito²⁷).

Il dispositivo, inventato nel 1963 e presentato per la prima volta nel 1967²⁸, venne ripresentato un anno dopo come parte di un sistema più complesso, in uno degli eventi riconosciuti come più importanti nella breve storia dell'informatica contemporanea: la demo di presentazione dell'oN-Line System, universalmente noto come NLS²⁹. La presentazione, il cui titolo originale era *A Research Center for Augmenting Human Intellect*, venne ribattezzata nel 1994 *The Mother of all Demos* da Steven Levy³⁰ per la sua monumentale importanza ed è da allora nota con il suo secondo nome. La demo ebbe luogo l'8 Dicembre 1968 al Fall Joint Conference Center nella città di San Francisco, come riporta una delle prime slide della registrazione della presentazione di Engelbart:

This movie captures directly a technical-session presentation made at the Fall Joint Conference Center in San Francisco, on December 8, 1968.

The movie screen will show what was projected by a high-powered TV projector onto a 22' x 18' screen mounted at the front of the 2000-chair Convention-Center Arena, and the soundtrack will reproduce what came over the loudspeakers.

Engelbart propose il suo sistema come ben rodato: si trattava del suo strumento di lavoro, come ne esistevano altri cinque nel loro ufficio, presto raggiunti da altre sei macchine analoghe. Si trattava di una macchina pensata per collaborare all'interno di un gruppo, e allo stesso tempo, forse, del primo computer davvero personale, in contrapposizione alle enormi macchine cooperative che l'avevano preceduto, in

²⁷ La frase proviene dalla prima presentazione del mouse come parte del sistema NLS, nota come *The Mother of All Demos*, di cui vennero mostrati alcuni spezzoni durante la celebrazione del 1998 *Engelbart's Unfinished Revolution*, istituita al trentennale dalla data della demo stessa. Le registrazioni dell'evento del 1998, incluse le clip della demo originale mostrate in quell'occasione, sono disponibili sul canale YouTube del Doug Engelbart Institute. Ultima consultazione 07/10/2024. Una versione "remastered" della demo nella sua versione integrale è visionabile allo stesso canale YouTube.

²⁸ S. Levy, *Insanely Great*, Viking Penguin, New York (1994).

²⁹ P. E. Ceruzzi, *A History of Modern Computing*, MIT Press, Cambridge, MA (1998).

³⁰ S. Levy, *Insanely Great*, Viking Penguin, New York (1994).

confronto alle quali i dispositivi di volume simile all’NLS erano stati soprannominati *mini-, micro- o desktop computers*³¹.

L’interfaccia fisica della macchina era costituita da un monitor per l’output e da un trittico innovativo per l’input: un mouse e due differenti tastiere. Il mouse pilotava un puntatore secondo una logica non dissimile da quella che ancora ne governa il funzionamento, seppure il sistema meccanico a rotelle rendesse più facile ruotare il mouse per causare un movimento lineare del puntatore, la cui forma grafica era quella di un piccolo insetto nero. Il suo funzionamento, secondo quanto riportato da Engelbart, era sufficientemente intuitivo da permettere a chiunque di familiarizzare molto rapidamente con la corrispondenza tra il moto del dispositivo fisico e quello del puntatore a schermo, rimuovendo il bisogno di preoccuparsi del movimento della mano, che diveniva presto un automatismo. Al centro della postazione, Engelbart esibì una tastiera QWERTY. La descrisse come “essenzialmente una tastiera standard per la macchina da scrivere salvo alcuni tasti speciali sul lato. Il computer sa immediatamente quando la usi e produce una risposta appropriata.” Infine, un secondo tastierino di cinque tasti chiamato Chord Keyset, la cui forma ricordava quella di un piccolo xilofono, permetteva di produrre testo usando una sola mano, premendo i tasti singolarmente o in differenti combinazioni.

L’interfaccia digitale era fondamentalmente testuale: sulla schermata bianca apparivano differenti istruzioni e comandi possibili, controllati dalla tastiera e dal mouse, in base al set di operazioni disponibili in quel momento. L’interfaccia era condivisibile: era infatti possibile permettere a un’altra macchina nella stessa rete l’accesso al proprio schermo, visualizzando il puntatore di un collega che desiderasse indicare differenti parti dell’interfaccia per discutere, e simultaneamente videochiamare quel collega (nel caso della demo, situato nel laboratorio SRI a Menlo Park).

Era anche possibile fornire al collega permessi di lavoro sulla propria macchina.

Infine, sulla scia del saggio di Bush, Engelbart introdusse il concetto di ipertesto: collegamenti tra differenti informazioni presenti nel sistema tramite parole chiave,

³¹ F. Turner, *From Counterculture to Cyberculture*, The University of Chicago Press, Chicago (2006).

che potevano essere ricercate o utilizzate direttamente come portale per schermate differenti. Si trattava in tutto e per tutto della prima interfaccia basata sulla coordinazione oculo-manuale in tempo reale.

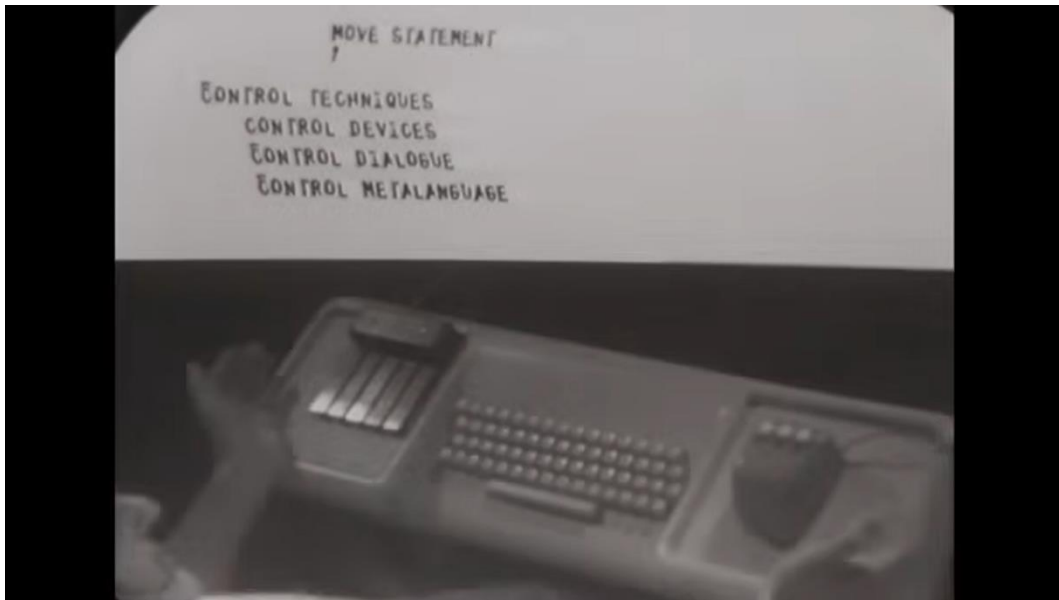


Fig. 7: Nella metà superiore, una parte del monitor di Douglas Engelbart durante la demo. Nella metà inferiore, i tre dispositivi di input impiegati da Engelbart.

Se fino ad allora le interfacce digitali erano costituite da vocabolari di comandi testuali, e dunque operate tramite un linguaggio autonomo parlato solamente dai professionisti e dalle macchine. Tramite la metafora del puntatore e la manipolazione del mouse Engelbart aveva aperto le porte a un'agentività più corporea, basata su facoltà motorie comuni (seppur non universali) piuttosto che su un enorme corpus di conoscenze, spianando la strada alla diffusione del computer presso un pubblico più ampio o, forse più correttamente, iniziando a trasformare il pubblico in utilizzatori di computer.

Il mouse non fu in realtà la prima forma di puntatore: la pistola ottica (*light gun*), poi rinominata penna ottica (*light pen*), era già ampiamente diffusa nell'uso delle tecnologie radar³², settore con cui Engelbart aveva grande familiarità, avendo

³² T. Bardini, *Bootstrapping: Douglas Engelbart, Coevolution, and the Origins of Personal Computing*, Stanford University Press, Stanford, CA (2000).

operato come tecnico radar nella marina americana tra il '44 e il '46³³. All'epoca dell'invenzione del mouse la penna ottica riscuoteva un certo grado di successo, ed erano diverse le realtà industriali impegnate nella produzione di calcolatori che la includessero³⁴. Il dispositivo di puntamento nacque durante il progetto Whirlwind I dell'MIT, sede della realizzazione del primo computer con la capacità di esibire testo e immagini in tempo reale in risposta a degli input, ideato per generare una simulazione di volo per la marina americana e successivamente evoluto nel sistema di difesa antiaereo SAGE (Semi-Automatic Ground Environment) impiegato nella Guerra Fredda³⁵. Gli operatori radar del sistema Whirlwind I vennero dotati di penne ottiche per richiedere informazioni su un dato velivolo sullo schermo radar toccando l'icona corrispondente. Le differenze con il mouse sono molte ed evidenti, a partire dalla natura prettamente fisica della *light pen*, che non presentava alcun tipo di puntatore all'interno di un'interfaccia e conseguentemente mancava del carattere di costante presenza digitale del mouse, del suo coesistere nel mondo fisico e macchinico ponendosi come ponte tra le due realtà parallele. L'idea di utilizzare la penna ottica in maniera continuativa per manipolare oggetti si può rintracciare fino al 1963, nella tesi di dottorato di Ivan Edward Sutherland (Sketchpad, il software associato a quella tesi, gli valse un premio Turing nell'88)³⁶, ma per eguagliare sul grande mercato la prestazione di costanza d'uso del mouse, la penna avrebbe dovuto perdere il suo carattere ottico e attendere l'avvento degli schermi tattili capacitivi.

Oltre alla semplificazione d'uso attraverso la coordinazione oculo-manuale, il mouse portò con sé altri due fondamentali vantaggi rispetto alla penna ottica. Il primo è di natura spaziale: il posizionamento del dispositivo sulla superficie di un

³³ Douglas Engelbart, in *Encyclopedia Britannica*, <https://www.britannica.com/biography/Douglas-Engelbart>. L'informazione è presente anche nel Curriculum Vitae di Douglas Engelbart, reperibile sul sito del Doug Engelbart Institute, presso <https://dougengelbart.org/content/view/151/> ultima consultazione di entrambe le fonti 09/10/24.

³⁴ T. Bardini, *Bootstrapping: Douglas Engelbart, Coevolution, and the Origins of Personal Computing*, Stanford University Press, Stanford, CA (2000).

³⁵ R. R. Everett, *The Whirlwind I Computer*, in *International Workshop on Managing Requirements Knowledge conference proceedings*, Philadelphia (1951), pp. 70.

³⁶ I. Sutherland, *The TX-2 Computer and Sketchpad*, in *Lincoln Laboratory Journal*, Vol. 19 (2012), online.

tavolo. Con o senza necessità di ruote, posando il puntatore materiale su una superficie di supporto è possibile ridurre drasticamente il carico sul braccio durante l'attività al computer, rendendo molto più facile farne un uso continuo e prolungato senza dover mantenere a lungo posture o sforzi muscolari particolarmente faticosi. La sua forma deve accomodare a sua volta un simile uso, senza indurre calli sulle dita o tendiniti, e deve pertanto essere libero di muoversi e trasdurre l'intenzione dell'utilizzatore con uno sforzo minimo, tramite l'utilizzo di superfici scorrevoli e tasti estremamente sensibili al tocco. Inoltre, la posizione sulla superficie del tavolo implica una maggiore vicinanza a un qualunque altro dispositivo di input su cui possa essere necessario trasferire la mano, e in particolare alla tastiera. Il secondo vantaggio, complementare al primo, consiste nella rimozione dell'oggetto dallo spazio antistante lo schermo. Non è possibile, infatti, guardare uno schermo attraverso una penna ottica. Nel relazionarsi con il computer, e per accedere al suo mondo, Douglas Engelbart si era reso conto dell'ostacolo costituito dal dover interrompere costantemente la lettura e le operazioni in corso per sollevare un grosso oggetto di fronte al sé e smettere di vedere lo schermo. La corrispondenza tra il puntatore fisico e un'icona digitale all'interno dello spazio schermico, necessaria a sostituire la punta della penna, fu (e ancora oggi è) la risposta di Engelbart alla necessità di una prestazione dell'interfaccia uomo-macchina che all'epoca non si prestava ad alcuna semplice soluzione: la necessità di trasparenza. Cinque anni dopo la demo di San Francisco, nell'Aprile del 1973, lo Xerox Palo Alto Research Center, arricchito di un numero non trascurabile dei membri del gruppo di ricerca di Douglas Engelbart all'SRI, rilascia lo Xerox Alto. Si tratta, a pieno titolo, del primo personal computer: uno schermo a bitmap, una tastiera e un mouse, il cui avatar schermico poteva essere modificato tramite software³⁷. È su queste macchine che, l'anno successivo, fece la sua comparsa l'editor di testo Bravo, progettato da Butler Lampson e Charles Simonyi³⁸.

³⁷ T. Bardini, *Bootstrapping: Douglas Engelbart, Coevolution, and the Origins of Personal Computing*, Stanford University Press, Stanford, CA (2000).

³⁸ D. K. Smith, R. C. Alexander, *Fumbling the future. How Xerox Invented, then Ignored, the First Personal Computer*, William Morrow and Company Inc., New York (1988).

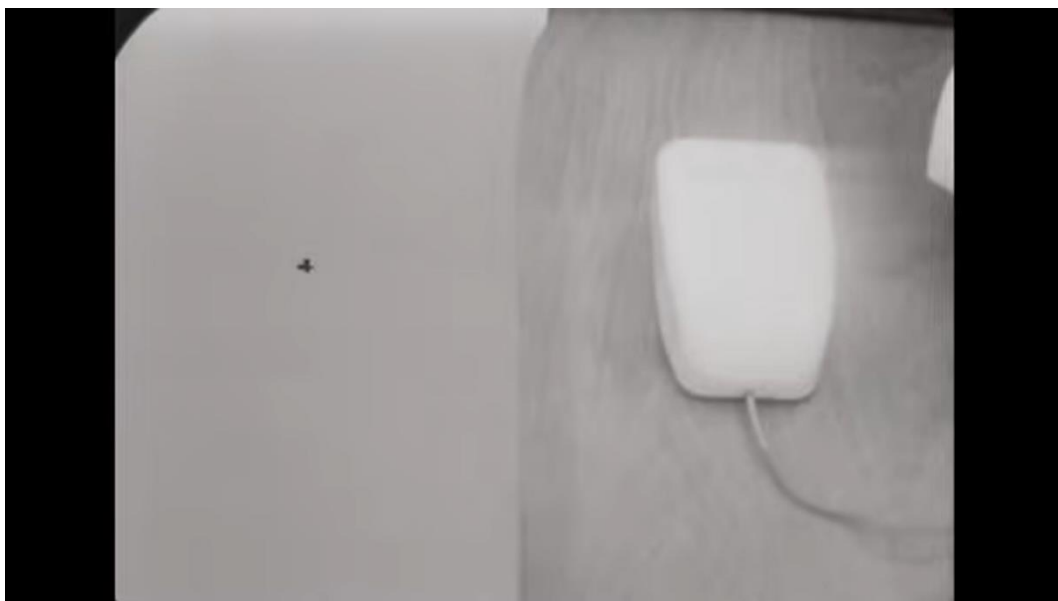


Fig. 8: A destra, il mouse di uno dei colleghi di Engelbart, trasmesso in live dal laboratorio durante la demo. A sinistra, il suo puntatore nello schermo.

Bravo è considerato il primo vero e proprio software di tipo WYSIWYG, acronimo di What You See Is What You Get. L'origine dell'espressione è colloquiale e risale alla prima metà degli anni '70, ma vuole il folklore che Chuck Thacker, uno degli ingegneri di Xerox, l'abbia coniato dopo che sua moglie, fortemente tecnofobica, gli abbia rivolto la domanda: "So what I see is what I get?" nell'osservare il funzionamento di Bravo su un computer Xerox Alto³⁹. Il significato dell'espressione è riferito a ciò che l'utente ottiene nel momento della stampa, la cui forma è riprodotta il più fedelmente possibile dall'interfaccia dell'editor di testo. L'editor di Xerox, che fu per molti aspetti genitore di Microsoft Word, dedicava al processo di stampa diverse pagine nel suo lungo manuale di istruzioni⁴⁰.

L'impatto di un simile cambiamento è simile a quello che un traduttore automatico ha nella conversazione tra due nativi di lingue differenti. Prima di Xerox Alto, i

³⁹ Una sintetica versione dei fatti è disponibile all'archivio online del New York Times, presso <https://archive.nytimes.com/bits.blogs.nytimes.com/2007/10/18/the-real-history-of-wysiwyg/> ultima visita 16/10/24.

⁴⁰ *XEROX DOCUMENT SYSTEM, Reference Manual*, versione 5.1, Xerox office product division (revisione dell'1 Novembre 1980).

sistemi operativi Unix si presentavano all'utente attraverso schermate nere ricche di opzioni di testo verde, un'estetica preservata nell'immaginario popolare dai film della serie Matrix. Non si trattava solamente di un'interfaccia differente, ma di un numero di passaggi nella catena di interfacce preposte a connettere il senso dell'utente alla materia di cui l'utente doveva farsi carico. La complessa catena di interfacce necessaria, strutturata come una sequenza di differenti traduzioni di linguaggio in linguaggio, Unix richiedeva che la parte umana si facesse carico di formare il documento nella propria mente, segmentarlo in parti manipolabili secondo le modalità agitive imposte dalla macchina e tradurre tali modalità e tali parti in comandi per la macchina per poterli impartire correttamente. L'utente non aveva alcun modo di osservare il documento futuro: non vi era un'interfaccia direttamente posta tra il suo sistema visuale e il destino del suo agire. La sovrapposizione di complesse interfacce era dunque opaca.

Evidentemente, questa affermazione è possibile solo a posteriori: non c'era motivo di concepire un'interfaccia sul futuro attraverso un calcolatore elettronico durante il suo quotidiano uso d'ufficio, e se ciò era possibile per l'operatore della stampa a caratteri mobili, la più comune e moderna macchina da scrivere poteva esibire il proprio frutto solamente in tempo reale. La ristrutturazione della catena di interfacce in una configurazione che permettesse di vedere in anticipo il frutto del proprio lavoro era però proprio il fine della ricerca di Engelbart, come emergeva dalla sua descrizione dell'operare dell'architetto nel suo rapporto del 1962. La macchina di Vannevar Bush, il memex, non era uno strumento progettuale, a differenza dell'editor Bravo. E se ancora non era possibile manipolare immagini che si modificassero per mostrare i diversi lati di un progetto tridimensionale come anticipato in *Augmenting Human Intellect*, la superficie schermica si prestava con molta più facilità a offrire un'altra superficie attraverso di sé. Ciò che l'utente vede è ciò che l'utente ottiene: una superficie su cui *avrà agito*.

READY: Select operand or type command
Last command was LOOK
{A_substa...!_way} {Computer... XEROX}S

**Personal Distributed Computing
The Alto and Ethernet Software**

Butler W. Lampson
Digital Equipment Corp. Systems Research Center

Abstract

The personal distributed computing system based on the Alto and the Ethernet was a major effort to make computers help people to think and communicate. A complex and diverse collection of software was built to pursue this goal, ranging from operating systems, programming environments, and communications software to printing and file servers, user interfaces, and applications such as editors, illustrators, and mail systems.

1. Introduction

A substantial computing system based on the Alto [Thacker et al.

Computer Science Laboratory
Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, California 94304

XEROX

Glen J. Culler
608 Litchfield Lane
Santa Barbara, CA 93109

Dear Glen:

This is a follow-up to earlier correspondence you received from Alan Perlis regarding the ACM Conference on the History of Personal Workstations. As you know, the conference is scheduled for January

Fig. 9: L'interfaccia dell'editor Bravo. Immagine pubblicata dal PARC e diffusa dal Computer History Museum di Palo Alto.

L'ultimo grande passo del percorso delle interfacce digitali ebbe luogo nel 1981. Nella variegata e dinamica evoluzione che il mercato dei calcolatori vide in quel breve lasso di tempo, l'ultimo grande passo verso l'interfaccia contemporanea avvenne nello stesso luogo del precedente, con la comparsa di Xerox Star⁴¹. In un mercato in cui le interfacce grafiche, a partire dalla demo di Engelbart, stavano iniziando a dominare le interfacce, lo Star rappresentò la prima adozione integrale della metafora della scrivania e un'interfaccia a icone il cui utilizzo è incentrato sul mouse per distinguersi dalla concorrenza⁴². Le interfacce a linea di comando non svanirono (né sono mai del tutto svanite dai computer contemporanei), ma vennero definitivamente nascoste all'utente comune:

⁴¹ J. Johnson et al., *The Xerox Star: A Retrospective*, in *IEEE standard* (September 1989), pp. 11-26.

⁴² Ibid.

Star's designers assumed that the target users were interested in getting their work done and not at all interested in computers. Therefore, an important design goal was to make the "computer" as invisible to users as possible. The applications included in the system were those that office professionals would supposedly need: documents, business graphics, tables, personal databases, and electronic mail. The set was fixed, always loaded, and automatically associated with data files, eliminating the need to obtain, install, and start the right application for a given task or data file. Users could focus on their work, oblivious to concepts like software, operating systems, applications, and programs.

Another important assumption was that Star's users would be casual, occasional users rather than people who spent most of their time at the machine. This assumption led to the goal of having Star be easy to learn and remember.⁴³

Il paradigma della Graphical User Interface, incentrato sull'utilizzo di icone basate su metafore (e metonimie) e un sistema di puntamento, aveva trionfato, e con esso l'estetica dell'ufficio. Da allora, l'Interfaccia Grafica (GUI) è rimasta la forma dominante delle interfacce puramente digitali. Essa porta con sé, oltre al carattere intrinsecamente gerarchico della struttura basata su documenti e cartelle, precise connotazioni spaziali e la necessità di avatarizzazione che queste implicano, un'agentività puntuale e specialistica, specifiche forme decisionali e un enorme bagaglio semiotico anche al più superficiale livello iconico (inteso anche nel più banale dei sensi: quello delle icone che popolano i software). L'ecosistema della GUI è sufficientemente astratto da ospitare innumerevoli altre interfacce dentro di sé, distinte secondo differenti criteri, dalla metafora alla forma fisica o tecnica dell'interattore materiale. Tali interfacce rimangono però accomunate dalla necessaria estetica segmentata in frammenti, delineati da cornici, univocamente identificabili tramite un qualche sistema di puntamento il cui movimento è a discrezione dell'utente, secondo un meccanismo perfettamente incarnato ormai

⁴³ Ibid.

quattrocento anni fa grazie all'improbabile collaborazione tra un giovane astronomo e il ragno che aveva tessuto la propria tela nel suo studio.

III. Classificare l'interfaccia

Torniamo dunque all'affermazione che ha avviato il presente lavoro, quella secondo cui l'interfaccia si dice in molti modi. Nell'esemplificare definizioni molto differenti di interfaccia è stato possibile osservare uno spettro costituito di ambiti di grande varietà, all'interno dei quali il termine interfaccia assume significati precisi, che possono poi venir utilizzati letteralmente o declinati metaforicamente nel discorso quotidiano. Tuttavia, l'interfaccia si dice in molti modi anche all'interno del solo ambito progettuale e digitale. Nella progettazione di interfacce, infatti, così come in un qualunque discorso sull'informatica in senso più ampio, non solo si distinguono numerosi tipi di interfaccia, ma anche un non trascurabile numero di criteri e tassonomie per determinare tali tipi. Tali criteri tassonomici sono tanti e tali da meritare a loro volta una tassonomia. Avendo già accennato a come un tipo di interfaccia possa costituire un paradigma estetico per interfacce che le sottostanno, abbiamo essenzialmente già stabilito l'esistenza di differenti livelli gerarchici e ontologici di interfaccia, che si incapsulano e combinano per dare forma a specifiche istanze, ibridi altamente specializzati e modularmente configurati per soddisfare determinate esigenze comunicative, agentive, narrative o funzionali.

È altresì opportuno evidenziare, prima di qualunque osservazione tassonomica o classificazione, che l'interfaccia utente-macchina è una tecnologia altamente diffusa, seppur altamente variata e dinamica, e come ogni tecnologia altamente diffusa è storicamente giunta a uno stato normalizzato e normalizzante. Alcuni dei principali meccanismi che hanno guidato questo processo sono stati presentati nella storia e nell'archeologia mediale dell'interfaccia nel presente lavoro. La ricerca nel settore dell'interazione utente-macchina, e in particolare nel settore delle tecnologie per l'assistenza alla disabilità, è alla costante ricerca di forme alternative di agentività che permettano di aggirare il processo di esclusione di determinate forme fisiche o culturali, ma ciò rappresenta ovviamente una strettissima minoranza del panorama tecnologico, il cui mercato necessita la diffusione di pratiche comuni

tanto nella progettazione quanto nell'utilizzo delle interfacce, di riconoscibilità dei prodotti e delle loro funzionalità e di trasferibilità delle competenze dei progettisti da un prodotto a un altro e da una società a un'altra. Dal punto di vista dell'utenza, imparare un linguaggio per farne uso una sola volta e dover poi apprenderne uno nuovo al primo cambio di dispositivo elettronico è praticamente insostenibile. Questo non ci esonera tuttavia dall'osservare cosa differenti tipi di tecnologia richiedano, tramite la loro interfaccia, all'utilizzatore, imponendo gerarchie di azione, di corpi e di linguaggi attraverso la catena di traduzione dell'interfaccia stessa. L'intreccio del concetto di interfaccia con quello di trasparenza impone la necessità di un sistema visuale normalmente funzionante, un requisito presunto anche in questo testo nell'ampio utilizzo di verbi come "osservare" e "vedere" per descrivere atti di riflessione e comprensione. Il processo normalizzante avviene anche a livello macchinico: diversi tipi di interfaccia sono definiti in quanto differenti da una norma tecnica o progettuale, una norma impostasi spontaneamente come tale, ancora una volta per motivi solitamente legati al mercato delle tecnologie (o a dinamiche di monopolio dello stesso).

Di contro, dispositivi complessi e altamente specializzati, di uso professionale, creati per rispondere ad esigenze fortemente specifiche o che facciano dell'interfaccia parte integrante del prodotto possono dare origine a interfacce notevolmente lontane dagli archetipi comuni. L'emergere di nuove tecnologie causa l'espansione del dominio di determinati tipi di interfaccia e la scomparsa di altri. L'arte digitale cerca l'emergere di interfacce scevre dalle logiche dell'ottimizzazione funzionale e la fuga dalla conformazione a pratiche stabilite e istituite. Fornire una classificazione completa e dettagliata di ogni possibile interfaccia, come avviene nella tassonomia classica, significherebbe categorizzare ogni interfaccia esistente individualmente. Le interfacce non sono però vincolate all'ereditarietà dei caratteri, e ogni progetto può dar vita, senza emergere nella scena pubblica, a una nuova interfaccia che si può identificare come una commistione unica di elementi di altre interfacce pre-esistenti o una sovrapposizione parziale di interfacce note. Riteniamo dunque più utile fornire una categorizzazione delle

interfacce istituite e utilizzate nelle pratiche della progettazione digitale, secondo criteri di distinzione di dominio ontologico (materiale, estetico, teleologico).

La distinzione da praticare è dunque quella tra tipologie di interfaccia. Le diverse tipologie hanno la loro origine nell'ambito progettuale, dove differenti priorità in diverse fasi della progettazione di un sistema ricercano differenti caratteristiche di un'interfaccia, e sono pertanto tipologie di interesse pratico. Di conseguenza, differenti gruppi e settori dell'informatica e della progettazione di interfacce finiranno per applicare differenti distinzioni. Questo implica una forte sovrapposibilità tra interfacce appartenenti alle diverse categorie, così come la coesistenza di interfacce appartenenti alla stessa categoria.

A. Interfacce hardware

La prima categoria è quella costituita da interfacce distinte secondo la forma fisica e materiale dei loro dispositivi di input e out. Al di fuori dell'ambito della ricerca, i mercati di differenti dispositivi sono stati per la maggior parte del tempo dominati da un singolo tipo di interfaccia hardware: nel caso del computer, la combinazione di un mouse, una tastiera, uno schermo e un dispositivo di riproduzione audio. È in primis questa forma materiale del calcolatore che ha consacrato la dimostrazione di Engelbart del '68 come "la madre di tutte le demo"¹. La conseguenza più immediata di un simile monopolio, ai fini della presente classificazione, è che questa combinazione di dispositivi non viene percepita come interfaccia a sé stante ma al più come parte integrante del computer nel senso più profondo, si fa trasparente al punto da divenire condizione intuitiva dell'operare con il medium, e cessa di essere a sua volta un medium individuato che connette l'utente a una macchina il cui funzionamento prescinde da essi.

In *The Interface Effect*, Alexander Galloway descrive questo effetto come parte del processo che ha trasformato una società di superfici riflettenti in una di superfici trasparenti. Il suo argomento si concentra sul confronto tra un mondo catottrico, di spettacolo, e un mondo diottrico, di controllo, secondo un effetto perfettamente applicabile ai dispositivi qui descritti:

As technology, the more a dioptric device erases the traces of its own functioning (in actually delivering the thing represented beyond), the more it succeeds in its functional mandate; yet this very achievement undercuts the ultimate goal: the more intuitive a device becomes, the more it risks falling out of media altogether, becoming as naturalized as air or as common as dirt.²

Un'analisi anche solo vagamente accurata degli effetti schermici, della storia o della semiotica degli schermi è un lavoro immenso – non immenso per un singolo

¹ Cap 1.II

² A. Galloway, *The Interface Effect*, Polity Press, Cambridge (2012), p.25.

ricercatore o testo, ma per interi campi disciplinari. Sullo schermo, lo studio del cinema ha prodotto una quantità incommensurabile di ottimi testi in grado di fornire panoramiche mediali, semiotiche, narratologiche e posturologiche. Al di là di ciò che è stato già discusso nell'archeologia dell'interfaccia grafica riguardo la funzione informativa della superficie piatta e controllabile, e del sottolineare il ruolo fondamentale che l'editor di testo Bravo e la forma testuale dell'informazione abbiano giocato nella storia dell'informatica – portando con sé la pagina come cornice e superficiale rettangolare – non possiamo che rimandare alla vastissima letteratura riguardo il nostro rapporto con le superfici schermiche per non esondare completamente dai temi trattati in questo capitolo³. Ciò che possiamo dire è che lo schermo sembra non essere così lontano dal raggiungimento dello stato di “demediazione” descritto da Galloway: due generazioni di nativi digitali stanno familiarizzando con gli schermi fin dalla più tenera età, al punto da riconfigurare attraverso i dispositivi portatili le loro capacità di apprendimento⁴ e i loro rapporti familiari⁵.

Di certo mouse e tastiera non hanno ultimato questo processo, né probabilmente lo ultimeranno mai, ma nel panorama delle interfacce hardware sono tra le poche ad aver percorso parte di questa strada. L'interfaccia costituita da mouse e tastiera non si presenta con nomi commercializzabili o sigle in lingua inglese: le altre interfacce hardware, di contro, necessitano di titoli precisi per identificarsi come distinte dall'interfaccia mouse e tastiera. Sottolineare la pervasività di mouse e tastiera, di per sé evidente, è necessario per comprendere la profonda e parimenti diffusa portata dei loro effetti.

³ I possibili riferimenti a questa breve digressione sono, come qui dichiarato, estremamente numerosi. Sceglieremo, per impatto sulla letteratura e sullo studio dell'argomento da parte dell'autore, così come per completezza disciplinare e ricchezza di contenuti, G. Bruno, *Surface: Matters of Aesthetics, Materiality and Media*, University of Chicago Press, Chicago (2014).

⁴ M. M. Neumann, D. L. Neumann, *Touch Screen Tablets and Emergent Literacy*, in *Early Child Education Journal*, 42 (2014), pp. 231-239.

⁵ H. Seo, C. S. Lee, *Emotion Matters: What Happens Between Young Children and Parents in a Touch Screen World*, in *International Journal of Communication*, 11 (2017), pp. 561-580.

Il primo di questi effetti ha a che fare con l'accessibilità. Interfacciare gli utenti a una tecnologia di importanza ormai capitale come il computer attraverso una specifica interfaccia fisica presuppone connettere a tale tecnologia, e all'universo a cui essa a sua volta interfaccia, un determinato tipo di corpi. È immediato osservare come l'utilizzo di un mouse e una tastiera implichi un corretto utilizzo di braccia dotate di mani funzionali, ma la normalizzazione è di spettro più ampio: non poter mantenere una postura seduta per motivi patologici, corporali o di altra natura previene un utilizzo normale di un'interfaccia che non ha visto, nella sua forma più diffusa, evoluzioni significative da quando fu proposta come oggetto da utilizzare alla scrivania di un ufficio da professionisti dell'informatica, discendendo a sua volta da un altro strumento di utilizzo professionale come la macchina da scrivere. Qualunque patologia causi dolore nell'uso degli arti superiori, così come qualunque situazione di ipovisione, rende molto difficoltoso l'uso di strumenti il cui funzionamento impone l'atomizzazione di gesti manuali operati ad alta frequenza determinati attraverso l'osservazione di uno schermo.

Vi è inoltre una forte barriera culturale, imposta dalla forma della tastiera, nell'atomizzare l'input in tasti: l'assunzione di modularità alfabetica del contenuto da produrre. Se questo è un problema sicuramente riconducibile alla macchina da scrivere, non è d'altra parte certo stato mitigato dalla tastiera: la travagliatissima storia dell'automazione della scrittura cinese, le cui difficoltà risalgono alla stampa in caratteri mobili (ideata in Cina prima che in Europa, ma di scarso successo proprio a causa della complessità di atomizzazione della lingua scritta) ma non sono state risolte prima di tempi recentissimi, nonostante innumerevoli ingegnosi tentativi di ricondurre la scrittura cinese, con la sua scrittura a modularità morfemica a un insieme ridotto di gesti automatizzabili⁶. Qualunque linguaggio non prettamente alfabetico, basato su pittogrammi e ideogrammi o altre forme di costruzione della lingua scritta, è destinato a incontrare le stesse barriere.

⁶ T. S. Mullaney, *The Chinese Typewriter, a History*, The MIT Press, Cambridge, MA (2018). L'autore deve ammettere con umiltà di essersi dovuto affidare alla letteratura occidentale sull'argomento, alla cui conoscenza è in origine pervenuto per racconto orale da parte di una sua conoscenza di origine cinese, per scarsissima conoscenza della lingua.

A CHINESE TYPEWRITER

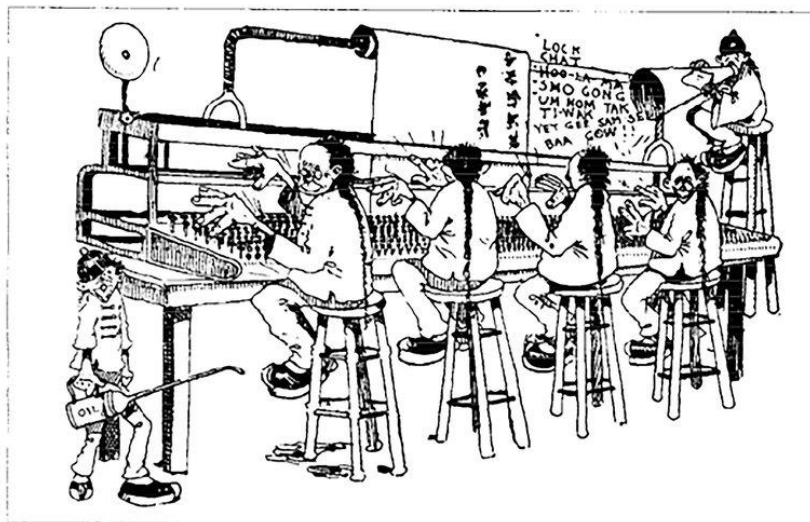


Fig. 10: Una vignetta satirica di forte connotazione razzista. Come in molti processi di normalizzazione, l'imposizione di meccanismi esclusivisti si accompagna alla denigrazione dei non normati. La vignetta apparve sul San Francisco Examiner il 22 gennaio 1900.

Certo Engelbart non aveva ideato la tastiera perché divenisse una tecnologia globale nella forma in cui aveva deciso di installarla su una manciata di macchine nel suo laboratorio e non avrebbe potuto, a meno di slanci di immenso ottimismo, prevedere il successo pervasivo che essa avrebbe incontrato. Tuttavia, prendere coscienza di simili meccanismi è fondamentale nella misura in cui, nell'osservare altre interfacce hardware, è possibile osservare la profondità dell'impatto di una forma il cui uso decennale ha plasmato l'idea stessa di cosa significhi *input* e *output*, di quali siano le operazioni possibili su di essi (vincolo di grande interesse per due concetti, come *input* e *output*, che dovrebbero rappresentare informazioni svincolate da forma) e con essi cosa un dispositivo elettronico possa o meno presentare come "contenuto". D'altro canto, l'efficacia di questa interfaccia nell'universo normato è indiscutibile e dimostrata dalla loro resistenza al cambiamento: i tentativi di trasformazione della tecnologia non sono stati pochi, da tastiere alternative di forma sperimentale a piccoli bottoni rossi, presenti nelle tastiere dei laptop, per pilotare il puntatore del

mouse, ma ad oggi nulla è riuscito a scalzare la dicotomia “mouse-tastiera rettangolare” dall’uso comune.

Interfacce tattili

Il secondo tipo di interfaccia hardware d'uso e riconoscimento comune, nonché la prima che distinguiamo da mouse e tastiera, è l'interfaccia tattile, o Touch User Interface (TUI). Nonostante il nome suggerisca un utilizzo di parti del corpo dotate di capacità aptiche, il nome si riferisce a ogni interfaccia il cui dispositivo di input sia una superficie che necessita di esercitare una forza su di essa ma non di superare di una determinata soglia di pressione per validare l'*input*, e in cui, nella maggior parte dei casi, l'area dedicata al raccoglimento dell'*input* fornito dall'utente coincide con una superficie schermica di *output*. Le due principali varianti, dal punto di vista tecnologico, si differenziano in base alla necessità di fornire l'*input* con una parte del corpo umano (e in particolare l'epidermide, dunque un contatto non mediato da stoffe o altre ostruzioni materiali) o meno, permettendo in questo secondo caso l'uso di strumenti inorganici come penne in plastica. Il primo caso è quello delle superfici capacitivie, da capacità, la grandezza elettrica rilevante in un condensatore (che prende in lingua inglese il nome di "capacitor"), sul cui funzionamento è basata la ricezione del segnale di *input*, mentre il secondo è il caso delle superfici resistive, il cui funzionamento è cioè basato sull'uso di resistenze elettriche.

Rispetto a mouse e tastiera questo tipo di interfaccia è di diffusione più recente, ma non meno esplosiva. Ad oggi l'utilizzo di interfacce tattili permea tecnologie domestiche e dispositivi pubblici, come i bancomat e le biglietterie elettroniche, e persino i monitor di computer portatili dotati di tastiera e mouse, e ovviamente gli schermi degli smartphone.

L'impatto delle interfacce tattili è difficile a definirsi o quantificarsi, sia a causa della rapidità della loro diffusione sia a causa della capillarità con cui hanno invaso la vita quotidiana di miliardi di persone. Seppur la loro prima adozione sperimentale preceda il mouse e la tastiera di Engelbart⁷ e tecnologie come gli sportelli bancari

⁷ E. A. Johnson, *Touch Display — a Novel Input/Output Device for Computers*, in *The Institution of Engineering and Technology* (originale del 1965), presso <https://digital-library.theiet.org/doi/abs/10.1049/e1%3A19650200>, ultima visita 25/10/24.

ne abbiano fatto uso fin dagli anni '80, è difficile non attribuirne la vera globalizzazione ai primi anni di questo secolo, quando l'esplosione del mercato degli smartphone e altri dispositivi portatili (come la console Nintendo DS, di cui più di 50 milioni di copie furono vendute solo in Europa⁸) ha condotto alla naturalizzazione del loro utilizzo. Simili interfacce saranno e sono già, per le generazioni più giovani, un fattore determinante nel rapporto con il tocco, con le superfici schermiche e di conseguenza nel rapporto con la realtà.

La necessità di far collimare dato e rappresentazione domina le interfacce software che accompagnano i dispositivi tattili. Questo causa l'emergere di nuove metafore e metonimie in grado di comunicare all'utente tanto un possibile senso del dato quanto le operazioni che su tale dato è possibile compiere.

Dal punto di vista tecnico, l'utilizzo di uno spazio di *input* manuale continuo, contro quello atomizzato della tastiera e dei tasti del mouse, ha ricercato la coordinazione oculo-manuale (ancora più pressante che nelle interfacce a mouse e tastiera) nella coincidenza tra spazio rappresentazionale e spazio di controllo dei dispositivi, con interessanti conseguenze per l'interfaccia digitale implementata nei dispositivi dotati di sistemi tattili.

L'accentramento del contatto tattile tra la mano dell'utente e lo schermo come interfaccia tra l'utente e la rappresentazione di dati, in genere attraverso una metafora fisica, implica una maggiore complessità dello spazio di input: i punti significativi di uno schermo tattile sono molto più numerosi e meno salienti dei tasti di una tastiera, e il loro significato è sottoposto a costante e più profonda variazione. Nessun punto di uno schermo tattile corrisponde sempre e comunque a una data lettera o simbolo. Questo rappresenta un fattore di sicurezza a favore di simili interfacce contro agenti dannosi, poiché fornisce una maggiore resilienza al tracciamento e all'interpretazione dell'*input* dell'utente da parte di terzi⁹. Dall'altra, la responsabilità di significare l'input viene affidata all'interfaccia formale sottostante o a interfacce fisiche utilizzate simultaneamente a quella tattile, come

⁸ <https://www.statista.com/statistics/1101884/unit-sales-nintendo-ds-region/>, ultima visita 25/10/24.

⁹ L. Cai, H. Chen, *TouchLogger: Inferring Keystrokes On Touch Screen From Smartphone Motion*, in *6th USENIX Workshop on Hot Topics in Security*, San Francisco (2011), conference paper.

ad esempio il tasto Home di diverse generazioni di iPhone, la cui funzione universale era quella di mettere in pausa qualunque operazione interrompibile e riportare il dispositivo alla schermata iniziale.

Le interfacce tattili hanno guadagnato progressivo interesse nella ricerca negli ultimi vent'anni. Le potenzialità creative e plastiche della superficie in quanto tale (dopotutto, il requisito di tattilità restringe molto poco il campo delle superfici possibili) sono state di forte stimolo per il mondo accademico¹⁰. La ricerca sull'interfaccia tattile dal punto di vista hardware implica sperimentare con la superficie stessa. A titolo esemplificativo, l'autore ha avuto modo, durante il suo soggiorno presso l'ENAC di Tolosa, di conoscere il lavoro corrente di alcuni progettisti francesi su interfacce tattili pieghevoli da impiegare nella cabina di pilotaggio degli aerei di linea, la cui forma permetta ai piloti di riconoscerne i segmenti e agire su diverse parti dell'interfaccia senza dovervi posare lo sguardo¹¹. Si tratta di una superficie plastica in grado di piegarsi come una fisarmonica in segmenti di pari lunghezza lungo il suo asse maggiore, che la mano può navigare e controllare indipendentemente, ad esempio piegandone alcune e lasciandone altre più piatte, secondo le proprie preferenze. Tramite questo approccio, la separazione dei quattro principali compiti del pilota (pilotaggio, navigazione, comunicazione e gestione dei sistemi¹²) viene parzialmente riflessa in una separazione delle modalità attentive, separando lo sguardo dalla gestione dei sistemi e dal pilotaggio e riducendo così il lavoro di integrazione sensoriale necessario al pilota. La superficie non è però né puramente tattile né del tutto priva di schermo. Si tratta innanzitutto di un display, che mostra informazione testuale e alcune elementari icone, per evitare che la mappatura delle diverse funzioni spazializzate debba essere affidata alla pura memoria. La superficie non presenta alcun segno tattile diverso dalle sue

¹⁰ M. Hafez, M. Benali Khoudja, *Tactile Interfaces: Technologies, Applications and Challenges*, in *REM2005 Conference Proceedings* (2005), pp. 267-272.

¹¹ A. Cantu et al., *Does Folding Improve the Usability of Interactive Surfaces in Future Airliner Cockpits? An Evaluation Under Turbulent Conditions and Varying Cognitive Load*, in *IMH '21 Conference Proceedings* (2021), online.

¹² P. C. Schutte, A. C. Trujillo, *Flight Crew Task Management in Non-normal Situations*, in *The Human Factors and Ergonomics Society 1996 Annual Meeting Conference Proceedings*, Vol. 40 (1996), pp. 244-248.

pieghe che possa fungere da segno o icona di ciò che quella sezione controlla. Inoltre, è possibile controllare l'informazione visualizzata direttamente tramite la plasticità della superficie: comprimere o distendere l'intera interfaccia in avanti modifica la scala dei parametri visualizzati sul display tattile, riavvicinando il suo ruolo a quello di uno schermo tattile.

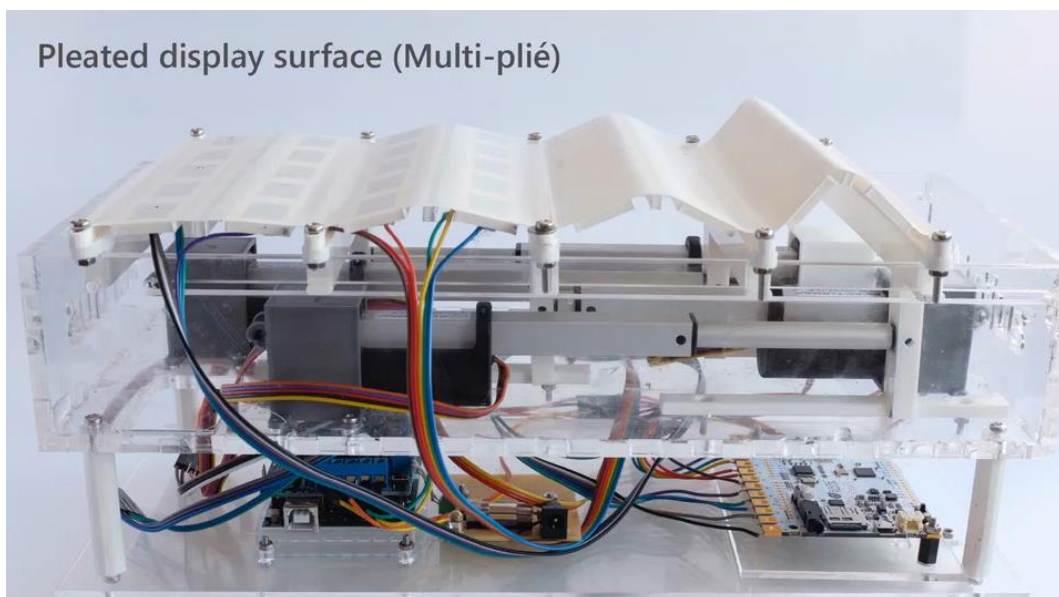


Fig. 11: Superficie tattile pieghevole e circuito sottostante. Immagine estratta dal filmato di presentazione del brevetto.

Al di fuori del più comune uso degli schermi tattili e dell'uso sperimentale di superfici plastiche, le interfacce tattili presentano varianti di differente responsività, in particolare tramite pressione e vibrazioni. In una review del 2020, Basdogan e colleghi hanno finemente classificato i diversi tipi di feedback tattile categorizzati come vibrotattili, elettrostatici o a ultrasuoni¹³. La review non si concentra su un utilizzo specifico delle interfacce tattili, ma considera ad ampio spettro differenti utilizzi, siano essi applicati a tele-tocco, videogiochi, supporto a persone ipovedenti, acquisti online, visualizzazione dei dati, apprezzamento artistico, istruzione o interfacce utente tattili funzionali (come quelle degli schermi tattili siti

¹³ C. Basdogan et al., *A Review of Surface Haptics: Enabling Tactile Effects on Touch Surfaces*, in *IEEE Transactions on Haptics*, Vol. 13, No. 3 (2020), pp. 450-470.

all'interno del cruscotto di un'automobile). Gli autori hanno d'altra parte deciso di escludere strategie tecniche meno comuni, come quelle basate sull'attuazione elettromagnetica, misurazioni di pressione tramite fluidi e interazioni aptiche cinestetiche (queste ultime, non prevedendo il contatto fisico diretto tra l'utente e la superficie, sono escluse dalle interfacce tattili anche nella classificazione proposta da questo testo). Dal loro lavoro emerge, in un'analisi delle tecnologie e della fisica che ad esse pertiene, come la maggior parte delle interfacce tattili svolga la funzione di comunicazione verso un utente tramite attuatori preposti a generare forze o vibrazioni sull'asse normale o sul piano tangenziale alla superficie tattile, tramite vibrazione continua o impulsiva nel primo caso, o tramite vibrazione, frizione (ad esempio, tramite generatori a ultrasuoni¹⁴) o forze nette nel secondo. Nell'analizzare le tecnologie dal punto di vista della risposta umana agli stimoli appena descritti, Basdogan e colleghi si appoggiano alla teoria della percezione tattile basata sulla percezione a quattro canali¹⁵, e in particolare al funzionamento dei canali meccanorecettori "Fast-Adaptive". Ciò che sottolineano è come la sensazione vibrotattile tenda ad emergere come una commistione della sensazione fornita dal materiale della superficie e di quella causata dal comportamento dell'interfaccia, in una situazione simile a quella descritta da Jeon e Choi come "Realtà Aumentata Aptica"¹⁶. L'approccio dominante è simulativo. Un esempio di applicazione a pieno campo di un simile approccio si può trovare nell'utilizzo dell'interfaccia tattile di *Death Stranding*, gioco di Hideo Kojima pubblicato nel 2019¹⁷. In un tentativo di esplorazione del costante ciclo di azione e reazione tra utente e videogioco come specificità mediale del campo videoludico, il gioco

¹⁴ L'applicazione della tecnologia a ultrasuoni alle interfacce tattili fu proposta in T. Watanabe, S. Fukui, *A Method for Controlling Tactile Sensation of Surface Roughness Using Ultrasonic Vibration*, in *IEEE International Conference on Robotics and Automation Proceedings*, Vol. 1 (1995), pp. 1134-1139.

¹⁵ J. S. Bolanowski et al., *Four Channels Mediate the Mechanical Aspects of Touch*, in *The Journal of the Acoustical Society of America*, Vol. 84 (1988), pp. 1680-1694.

¹⁶ S. Jeon, S. Choi, *Haptic Augmented Reality: Taxonomy and an Example of Stiffness Modulation*, in *Presence*, Vol. 18, No. 5 (2009), pp. 387-408.

¹⁷ Hideo Kojima, *Death Stranding*, Play Station 4, Kojima Productions (2019). Hideo Kojima è padre della saga di Metal Gear, nota per il suo, pionieristico, ampio ed efficace utilizzo di stimoli vibrotattili tramite controller.

sviluppato da Kojima atomizza l'azione del camminare rendendo significativo ogni singolo passo del protagonista. Nel trasportare il suo pesante carico, il personaggio guidato dal giocatore (e interpretato da Norman Reedus) oscilla e barcolla costantemente nel suo deambulare, e vibrazioni di diversa intensità, durata e forma comunicano alle mani dell'utente come il corpo del suo avatar stia navigando lo spazio. Un simile utilizzo dell'interfaccia tattile ha una forte capacità di ingaggio dell'utenza, creando una simulazione multisensoriale e (se ben realizzata) fortemente sinestetica, che permette di integrare stimoli molto diversi tra loro con forte naturalezza. Dal punto di vista tecnico, tuttavia, l'utilizzo costante di stimoli vibrotattili significa il costante utilizzo di un motore meccanico e dunque un forte dispendio di energia per dispositivi autonomi di piccole dimensioni¹⁸.

Infine, Basdogan e colleghi hanno offerto una valutazione delle soluzioni tecnologiche analizzate dal punto di vista della progettazione di interfacce digitali, principalmente secondo la comune metrica della Legge di Fitts¹⁹. Il secondo aspetto su cui gli autori della review si soffermano è il forte legame tra le interfacce tattili e diversi tipi di interfacce a bio-sensori, il cui uso è favorito dalla già presente necessità di contatto fisico tra la pelle dell'utente e la superficie del dispositivo tecnico. Si noti però che, a differenza dei sensori tattili ad alta densità di punti atti a percepire il tocco intenzionale dell'utente, i bio-sensori sono in gran parte ricettori di segnali non intenzionati (come il battito cardiaco o la risposta galvanica cutanea) che l'utente non può decidere di non fornire. Le interfacce biologiche, e in particolare le interfacce mioelettriche, saranno discusse in seguito nel presente capitolo.

¹⁸ Per un approfondimento quantitativo si veda I. P. Lyan et al., *On the Issue of Energy Consumption of Vibration Technological Machines*, in *IOP Conference Series: Materials Science and Engineering 747 Proceedings* (2020), online.

¹⁹ La formula, che predice la prestazione di un utente nel raggiungere un determinato punto d'arrivo in base alla sua distanza e alla sua larghezza, è spesso utilizzata per valutare l'intuitività di un'interfaccia confrontando dei test su utenti con il risultato previsto. L'autore omonimo è quello stesso Paul Fitts il cui rapporto all'aeronautica sulla necessità di un'interfaccia HUD è stato discusso nel capitolo sull'archeologia dell'interfaccia grafica.

Gli autori²⁰ hanno evidenziato come le interfacce tattili si prestino con grande efficacia ad alcune delle operazioni considerate basilari nell'utilizzo di interfacce digitali, e in particolare nel puntamento, ma carenti in altre, come l'utilizzo di bottoni e widget storicamente appartenenti alle interfacce grafiche, ma che non possono restituire all'utente un segnale fisico come avviene nel caso di un mouse. Per questo motivo, si osserva come comandi a scorrimento e a bottone, che impongono all'utente l'utilizzo di un solo dito, tendano a essere meno efficaci di controlli simulanti manopole e altri dispositivi tecnici fisici che necessitino di più dita, in particolare nel caso di interfacce tattili a frizione. L'effetto è rinforzato dall'impossibilità di localizzare la risposta del dispositivo a una singola area della superficie tattile.

Osserveremo ora un ultimo aspetto fortemente presente nel testo, ma lungi dall'essergli esclusivo nella letteratura di settore, utilizzandolo dunque come accidentale esempio. Nel descrivere l'operare le interfacce tattili, gli autori utilizzano "target acquisition" e "pointing" come perfetti sinonimi. Lo studio dell'archeologia delle interfacce grafiche operato in precedenza ci permette di affermare che ciò era vero fin dai tempi di Gascoigne, ma l'operazione di puntamento – che costituisce una pratica ben differente da quella di acquisizione di un bersaglio e ancor più di dichiarazione di intenzione - è stata attuata da strumenti via via più ridotti e precisi nella storia dell'interfaccia. L'idea che situare le proprie intenzioni coincida con il puntare qualcosa è divenuta fortemente normata dal punto di vista corporeo: un dito, o una sua avatarizzazione dotata di punta, è nella maggior parte dei casi l'unico modo accettabile per acquisire un bersaglio²¹. Su questo principio si fondano molte interfacce hardware, e le interfacce tattili ne sono un esempio particolarmente evidente. È probabilmente superfluo evidenziare come l'esplosiva diffusione a macchia d'olio delle interfacce tattili, in particolare tramite

²⁰ Ibid.

²¹ Esistono notevoli eccezioni, come ad esempio far uso della punta del naso per operare lo schermo di uno smartphone in caso le mani siano occupate. Tale soluzione, tuttavia, tenderebbe a sfociare nel ridicolo in molti contesti se applicata frequentemente, e richiede quantomeno l'uso di una mano funzionale per mantenere una postura a sua volta normata.

telefoni cellulari, abbia generato una nuova ondata di esclusione fisica e culturale nei confronti di coloro che non sono dotati di mani o dita pienamente funzionali o della vista, un meccanismo di esclusione di più difficile compensazione rispetto a quello operato da un'interfaccia basata su mouse e tastiera.

Interfacce cinetiche

Il terzo tipo di interfaccia hardware, e il primo privo di contatto fisico tra il dispositivo e il corpo, è l'interfaccia cinetica o Kinetic User Interface (KUI). Per interfaccia cinetica si intende un'interfaccia il cui *input* sia costituito da coordinate spaziali di oggetti o corpi materiali, con particolare attenzione alle loro variazioni. Le prime modalità di *input* sono dunque il *motion tracking* e la *location awareness*²². Bruegger e Hirsbrunner descrivono le KUI come sistemi composti di entità, o oggetti cinetici (possibilmente viventi), e osservatori, o agenti, la cui finalità è quella di istituire ecosistemi tecnologici pervasivi e non invadenti (*unobtrusive*)²³.

Il significato di *unobtrusive* può declinarsi diversamente in diverse istanze, dalla non-necessità di aderenza dell'oggetto tecnico al corpo dell'utente, all'adozione di sistemi di "input impliciti" tramite telecamere, ossia di lettura da parte di un sistema pervasivo della gestualità dell'utente quando essa non è finalizzata a comunicare con il dispositivo, per inferirne contestualmente le intenzioni²⁴. Simili interfacce sono fortemente connesse al campo della computazione affettiva e dell'interazione naturale.

Simili interfacce costituiscono nel presente testo un punto di particolare interesse: le interfacce dotate di un qualche livello di trasparenza, come di certo sono le interfacce cinetiche, sono state definite come il punto di contatto tra due entità di cui almeno una è in grado di osservare l'altra. Ma anche se non è mai stata posta alcuna restrizione formale su quale sia l'entità osservante e quale l'entità osservata, e alla presenza di un canale di scambio di *input* e *output* l'osservazione può apparire bidirezionale, questo è di certo il primo caso di interfaccia che sembra fare della

²² P. Bruegger, B. Hirsbrunner, *Kinetic User Interface: Interaction through Motion for Pervasive Computing Systems*, in C. Stephanidis (ed.), *Universal Access in Human-Computer Interaction*, in *HCI International Conference Proceedings*, San Diego (2009), pp. 297-306.

²³ Ibid.

²⁴ V. Pallotta, *Unobtrusive Interaction with Mobile and Ubiquitous Computing Systems Through Kinetic User Interfaces*, in M. M. Cruz-Cunha, F. Moreira (eds.), in *Handbook of Research on Mobility and Computing: Evolving Technologies and Ubiquitous Impacts*, IGI Global (2011), pp. 689-702.

macchina la sola parte osservante nella relazione tra uomo e computer. È dunque necessario operare una distinzione precisa.

Nel discutere le interfacce cinetiche, ci occuperemo qui di interfacce il cui utilizzo, quantomeno nella forma della loro installazione, è operato consapevolmente dall'utente che ne sarà poi osservato e che beneficerà direttamente della loro attività, osservando a sua volta i risultati del proprio agire. Tale distinzione è particolarmente necessaria nella presente fase di entusiasmo tecnologico per l'intelligenza artificiale, di cui alcune istanze (e in particolare quelle legate alla visione computazionale) permettono di automatizzare la sorveglianza a circuito chiuso, il riconoscimento facciale e l'osservazione di un gran numero di individui. Le implicazioni tecniche, sociali, etiche e biopolitiche delle applicazioni di massa dell'intelligenza artificiale sono ovviamente immense ed esulano dai temi di questo lavoro. Poiché perfino la loro appartenenza alla categoria delle interfacce cinetiche, così come la definizione della loro effettiva utenza, può essere ampiamente messa in discussione, il discorso in essere si limiterà alle tecnologie il cui uso è quello appena descritto.

L'interfaccia cinetica occupa un posto differente dalle precedenti nella catena di traduzione dall'utente alla macchina. Una maggior libertà gestuale dell'utente significa una maggior varietà di carico interpretativo per la macchina. Il primo aspetto di questo maggior carico è la necessità di una precisa lettura di movimenti a distanza. Sul piano hardware, esistono diverse soluzioni per ottenere una maggior precisione in lettura:

- 1) Imporre la posizione spaziale dell'utente in relazione alla posizione spaziale degli agenti osservanti
- 2) Aumentare il numero di agenti e variare le loro posizioni nello spazio
- 3) Utilizzare dei marcatori, sensorizzati o facilmente tracciabili, da applicare all'entità osservata

Il primo approccio, di più semplice realizzazione, è apparso per la prima volta in un'interfaccia pervasiva nel 1979 presso l'MIT, come parte del progetto "Put-That-

There”²⁵. Il sistema combinava ampi gesti delle braccia e comandi vocali segmentati per interagire con un grande schermo all’interno di una stanza. Nei video dimostrativi del progetto è possibile osservare uno dei ricercatori seduto su una poltrona al centro dell’ambiente, posizionato frontalmente al monitor. Mentre muove il braccio teso a indicare diversi punti dello schermo, un crosshair si sposta a indicare quale area il suo braccio stia indicando. Dopodiché, impartendo comandi come “create a yellow triangle there” lo schermo visualizza la figura richiesta nell’area indicata dal puntatore al termine del comando. È anche possibile impartire istruzioni spazialmente relative alle figure presenti: è ad esempio possibile spostare una figura “below that” indicandone un’altra²⁶.

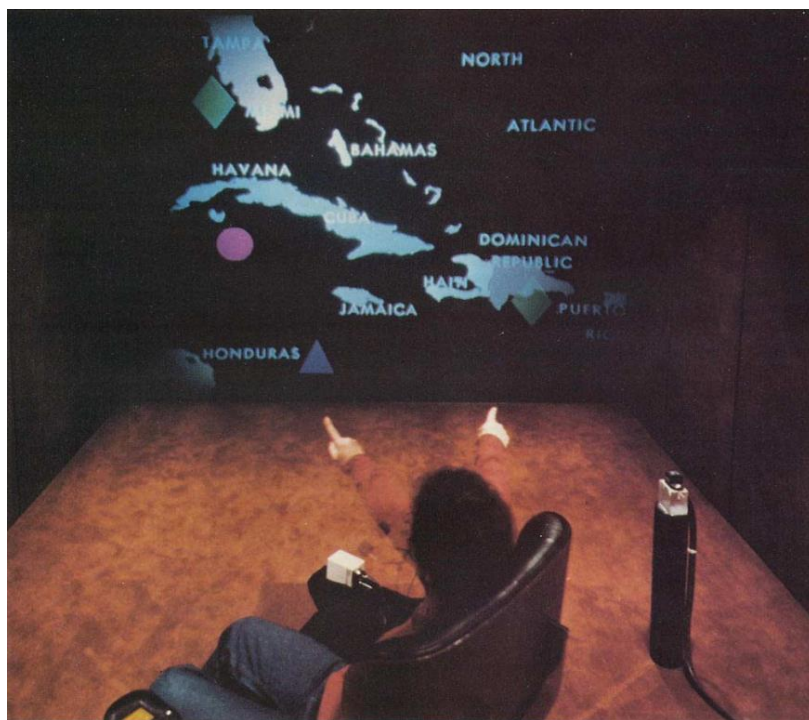


Fig. 12: Il prototipo di Put-That-There. Immagine estratta dalla pubblicazione di Bolt.

²⁵ R. A. Bolt, “Put-That-There”: *Voice and Gesture at the Graphics Interface*, report to the Cybernetics Technology Division of the Defense Advanced Research Projects Agency, MIT (1980).

²⁶ Put That There (Original), presso <https://www.youtube.com/watch?v=RyBEUyEtxQo>, ultima visita 31/10/2024.

Dal punto di vista tecnico, il sistema di interfaccia cinetica si affidava a dei generatori e a dei sensori di campo magnetico costante per “normalizzare” l’ambiente, leggendo le variazioni di tale campo per determinare l’angolo tridimensionale del braccio dell’utente. Si trattava di una delle tecnologie della Media Room, un ambiente chiave nella ricerca sulla possibile creazione di Spatial Data-Management Systems. La Media Room avrebbe dovuto contenere altri sistemi differenti, tra cui dei joystick, due televisori dotati di superfici tattili, microfoni e diverse fonti audio spazializzate. Come nel caso del memex di Bush e dell’Interfaccia Grafica, a dominare l’idea progettuale dello Spatial Data-Management System e la gestualità delle sue interazioni è ancora una volta la scrivania:

The specific rationale for spatially indexing data derives from our everyday experience of retrieving items, say, from our desktop: the phone to the right and above the blotter; the appointment calendar in the lower right; the "in-box" nearby the ashtray at the lower left, and so forth. Retrieval is natural and automatic for these items, with even an apparently "messy" desk having a spatial logic well-known to its creator and user, the knowledge of where this and that item are located being encoded conjointly in mental and motor models of the layout of the desktop. The world of information in SDMS, dubbed "Dataland," appears in its entirety upon one of the color TV monitors near-by the user chair. A small transparent rectangular overlay, a "you-are-here" marker, can be moved and positioned about Dataland by the user's managing of the chair's right-hand joystick (or by direct touch on the TV screen, if desired). That sub-portion of the Dataland surface indicated by the "you-are-here" rectangle is portrayed with increased detail on the large screen, effectively a magnifying window onto Dataland. The left-hand joystick on the user chair enables the user to zoom-in upon information to get a closer look at any of a number of multimedia data-types (e.g., maps, electronic "books," videodisc episodes), and perhaps to peruse them with the aid of an associated touch-sensitive "Keymap" which comes up on the Other TV monitor by the user chair.²⁷

²⁷ R. A. Bolt, *“Put-That-There”*: *Voice and Gesture at the Graphics Interface*, report to the Cybernetics Technology Division of the Defense Advanced Research Projects Agency, MIT (1980).

Sappiamo oggi che il progetto che diede origine a Put-That-There non è entrato nelle nostre case. Tuttavia, l'idea di utilizzare un sensore in posizione fissa rispetto all'utente per tracciarne la gestualità ha riscosso un buon grado di successo. Nel settore dell'industria videoludica il sensore Kinect, basato su questa tecnologia, fu rilasciato inizialmente per la console xBox 360 nel 2010 e divenne parte integrante della sua erede, la console xBox One, vendendo effettivamente milioni di copie. Il dispositivo si è inoltre dimostrato portante in differenti tentativi di creazione di interfacce per l'accessibilità nel mondo accademico. Il caso più noto di interfaccia cinetica di supporto a una persona affetta della disabilità è con ogni probabilità il sistema ACAT, creato da IBM per Stephen Hawking e il cui software, adattato per essere utilizzato con qualunque telecamera e monitor, è stato reso pubblicamente disponibile da IBM sulla piattaforma github²⁸. Il sistema faceva uso di una telecamera appositamente installata sulla sedia a rotelle di Stephen Hawking e di un monitor che permettesse allo scienziato di seguire i movimenti di un indicatore a scorrimento lungo una tastiera o attraverso lo schermo tramite un puntatore, a cui impartire comandi di selezione o annullamento tramite il movimento delle sopracciglia e di uno zigomo²⁹.

Il secondo approccio è comune quando emerge la necessità di misurare il movimento in quanto tale. Ciò può avvenire nel già citato caso dell'interazione naturale, ma anche in contesti in cui la finalità del sistema è lo studio o l'acquisizione del movimento in ogni suo aspetto in un ambiente controllato. Un esempio efficace di questa tecnologia si può ritrovare ancora una volta nell'industria videoludica, e in forme differenti. I dispositivi indossabili di realtà virtuale commerciali più noti hanno adottato questa strategia: i caschi della serie

²⁸ <https://github.com/intel/acadat>, ultima visita 31/10/2024.

²⁹ Le fonti per queste informazioni sono puramente esperienziali. Per il funzionamento del sistema, l'autore ha speso una quantità non trascurabile di tempo a impratichirsi (non senza soddisfazioni) nell'uso dell'ACAT tramite un computer portatile Lenovo ThinkPad, mentre per la struttura della sedia a rotelle l'autore si affida all'osservazione dal vivo della lectio magistralis, tenuta dal celebre fisico nell'anno 2006, in occasione della quale Hawking ricevette la cittadinanza onoraria e le chiavi della città di Padova.

Oculus Quest di Meta fanno uso di sistemi multicamera per tracciare movimento nello spazio, eventualmente integrati da accelerometri, giroscopi e magnetometri MEMS³⁰ per raffinare i dati acquisiti in tempo reale.

I caschi Oculus Quest 2 impiegano un sistema di quattro telecamere a sei gradi di libertà³¹ montate sul casco stesso per tracciare i movimenti dell'ambiente circostante, utilizzati per calcolare gli spostamenti del casco e quindi della testa dell'utente. Integrare il sistema nel casco è necessario, trattandosi di un dispositivo autonomo dotato del proprio processore e indipendente da altri dispositivi per il proprio funzionamento, caratteristiche che permettono libertà di movimento nell'ambiente. Le telecamere sono inoltre utilizzate per tracciare la posizione e i movimenti delle mani dell'utente, che può controllare gestualmente il sistema secondo i limiti dell'interfaccia del software in uso. Altri dispositivi indossabili simili utilizzano sistemi più o meno equivalenti, a volte potenziati tramite telecamere esterne e a volte meno forniti ma meno dispendiosi a livello energetico e computazionale. Il tracciamento manuale è anche alla base del controllo del sistema HoloLens 2 di Microsoft, il cui impianto sensoristico è molto simile a quello di Quest 2: un gruppo di telecamere che traccia l'ambiente per tenere traccia dei movimenti del capo e del movimento manuale³².

Un altro importante uso di simili interfacce si riscontra nell'ambito biomedicale, nello studio del movimento umano. In simili casi, per motivi di efficientamento, il sistema tende a essere utilizzato insieme ad altri differenti strumenti di misura per l'acquisizione e l'integrazione di dati diversi, ma il lavoro svolto dall'interfaccia a telecamere multiple le è esclusivo. Un esempio di un simile sistema è il sistema

³⁰ Micro Electro-Mechanical Systems.

³¹ L'espressione "gradi di libertà", sinonimo della sigla DoF (Degrees of Freedom) nei documenti tecnici, si riferisce nel caso di dispositivi cinetici al numero di variabili spaziali misurabili dal dispositivo, e distingue in genere sistemi a 6 DoF, in grado di misurare sia lo spostamento lungo i 3 diversi assi spaziali sia la rotazione intorno agli stessi, da sistemi a 3 DoF, in grado di misurare solo lo spostamento o la rotazione.

³² Il funzionamento è simile a quello di altri dispositivi AR. Come nei casi precedenti, la scelta è stata operata sulla base di una maggiore esperienza d'uso rispetto ad altri dispositivi simili. La documentazione è consultabile presso <https://www.microsoft.com/en-us/hololens/hardware#document-experiences>, ultima visita 08/11/2024.

Smart Dx³³ di BTS Engineering, azienda di produzione di tecnologie biomedicali. Il nucleo del sistema consiste in un gruppo di sedici telecamere in grado di tracciare con grande precisione e frequenza di acquisizione il movimento nello spazio di un corpo o un oggetto obiettivo. Il gruppo, insieme a una piattaforma di forza, costituisce un sistema finalizzato allo studio degli equilibri anatomici e dinamici di un atleta. La nuova forma di Smart Dx, Smart Dx EVO, monta fino a 256 (sedici al quadrato) telecamere per massimizzare la precisione del tracciamento³⁴. Il tracciamento operato dalle telecamere può essere, in caso di necessità, potenziato con un sistema di *markers* sferici di colore bianco posizionati lungo l'oggetto o il punto del corpo di cui tenere traccia, similmente a come avviene sui set cinematografici tramite le tute per la cattura del movimento.

I sistemi multicamera incontrano qui dunque il terzo approccio all'interfaccia cinetica, ossia l'implementazione di una parte dell'interfaccia esterna ai sensori e indossata dall'entità tracciabile. Le modalità possibili per questo tipo di interfaccia sono tante quante le possibili entità da tracciare, e vanno da indicatori materiali privi di alcuna elettronica ma di facile tracciabilità, come nel caso degli appena citati *markers*, a controllori in grado di impartire diversi tipi di comando insieme a, e a volte dipendenti da, la loro posizione e il loro moto nello spazio, come nel caso del telecomando della console Nintendo Wii o di diversi controller per dispositivi di Realtà Virtuale, come il Play Station VR di Sony o lo stesso Meta Quest 2, che tracciano i loro controller per poterli utilizzare come dispositivi di puntamento. Le tecniche di tracciamento possono anche variare con la tecnologia: i tre esempi appena citati si appoggiano, rispettivamente, su tecniche di visione automatica pura, su emissione e ricezione di segnali infrarossi combinate a trasmissione di comandi e sulla combinazione dei due approcci combinazione.

È d'altra parte da evidenziare come tutti e tre gli approcci descritti non possano essere indipendenti tra loro, specialmente dal punto di vista progettuale. Tracciare

³³ Il sistema Smart Dx è ormai fuori commercio, ed è stato sostituito dalla sua versione successiva, Smart Dx EVO. Come avvenuto più volte in questo capitolo, la scelta del dispositivo è stata compiuta secondo un criterio di maggiore esperienza diretta con i dispositivi stessi.

³⁴ <https://www.btsbioengineering.com/it/products/smart-dx-evo/>, ultima visita 31/10/2024.

un corpo umano o un oggetto nello spazio richiede la capacità di riconoscere punti di salienza da utilizzare come marcatori, anche in assenza di essi: senza riconoscere lo zigomo, il sistema ACAT precedentemente descritto non potrebbe funzionare. Le interfacce cinetiche così integrate offrono però grande precisione: è a un sistema VR dotato di controller tracciabili, arricchito di pedaliera per il controllo della metà inferiore del corpo, che la NASA ha affidato i comandi dei Robonauts, i suoi astronauti remoti³⁵.

È in ultimo necessario citare il paradigma dell'interazione tangibile³⁶ (Tangible Interaction), di cui parleremo più approfonditamente in seguito. Il principio centrale dell'interazione tangibile è l'utilizzo di oggetti tracciabili, manipolati o spostati nello spazio. Nonostante si tratti di una filosofia progettuale orientata alla manipolazione dei dispositivi di *input*, e dunque dal punto di vista agentivo più affine a un mouse che a un alfabeto di input gestuale di fronte a un sensore, esiste grande varietà nei sistemi di tracciamento degli oggetti manipolabili, che spesso svolgono il ruolo di *marker* sensorizzati come quelli discussi poc'anzi. Dal punto di vista *hardware* rimangono interfacce a cavallo tra due delle categorie descritte, e non realmente determinate da un tipo specifico di dispositivo tecnico, per cui ricadranno in una tassonomia successiva. Il loro inserimento nel presente paragrafo è finalizzato ad evidenziare come i dispositivi tracciabili possano variare in forma e funzionamento, pur rimanendo sistemi di interfaccia orientati a una navigazione libera dello spazio, e non progettate per una scrivania come nel caso di mouse e

³⁵ R. Ambrose et al., *Mobile Manipulation using NASA's Robonauts*, in *IEEE International Conference on Robotics and Automation Proceedings* (2004), pp. 2104-2109.

³⁶ B. Ullmer, H. Ishii, *Emerging Frameworks for Tangible User Interfaces*, in *IBM Systems Journal* (2000), pp. 915-931.

tastiere. L'attenzione all'esplorazione spaziale è evidenziata con particolare efficacia nelle carte di lavoro³⁷ dell'interazione tangibile proposte da Eichhorn³⁸.

³⁷ L'utilizzo di carte di supporto alla progettazione, tematizzate secondo i valori di una determinata filosofia progettuale e arricchite di domande utili a orientare il lavoro del progettista, è comune nell'ambito della progettazione di sistemi e interfacce. Le transfacce di Letondal et al. citate precedentemente, ad esempio, hanno dato origine a un relativo mazzo di dieci carte. Un altro esempio notevole è il testo *The Art of Game Design* di Jesse Schell, che in ogni sua edizione è arricchito da un'applicazione telefonica fornita di decine di carte, o "lenti", attraverso cui "guardare" al proprio progetto.

³⁸ E. Eichhorn, *The Tangible Interaction Framework Cards – Introduction*, risorsa progettuale reperibile sul sito personale di Elisabeth Eichhorn presso <http://www.ehornecker.de/>, ultima visita 5/11/24.

Interfacce a tracciamento oculare

Particolarmente vicine alle interfacce cinetiche dal punto di vista tecnico e concettuale sono le interfacce a tracciamento oculare. Il principio di funzionamento non richiede molte spiegazioni: si tratta di sistemi in grado di tracciare gli spostamenti intenzionali dell'occhio e utilizzarli come segnale di *input*.

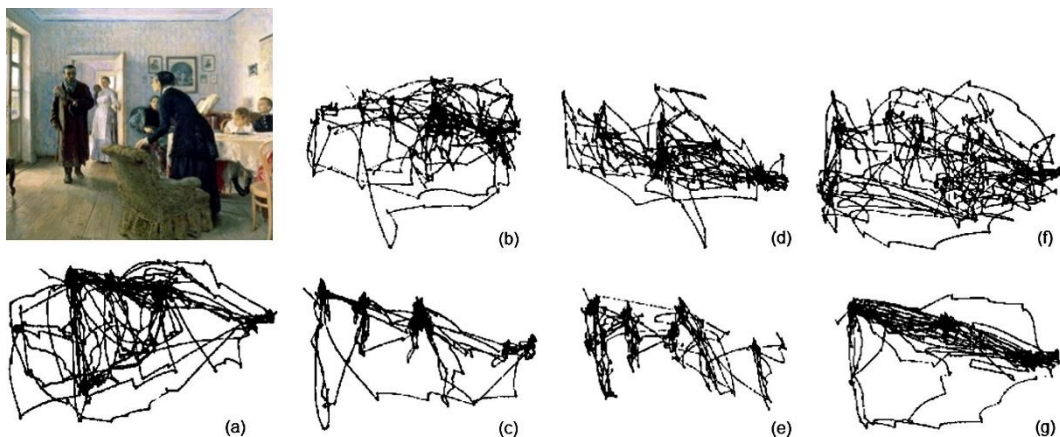


Fig. 13: Pattern oculari registrati da Yarbus in corrispondenza di diverse operazioni da svolgere sull'immagine.

Dal punto di vista pratico, la realizzazione e l'utilizzo di simili interfacce è meno immediato. La relazione tra sguardo e operazioni da svolgere è assai complessa, come verificò sperimentalmente Yarbus negli anni '60³⁹. I movimenti oculari involontari sono costantemente variati da scatti microsaccadici, e gli strumenti di misura in grado di tenerne traccia possono essere ingombranti e costosi. Per questo motivo, spesso simili interfacce vengono accoppiate con sistemi di tracciamento del capo, che fornisce una direzione generale dello sguardo da rifinirsi con il tracciamento oculare per le misure più precise. Ottimo esempio di una simile implementazione è il sistema Radi-Eye di Sidenmark e colleghi⁴⁰. Nel sistema progettato dai ricercatori britannici, la coordinazione tra capo e occhi pilota un'interfaccia digitale di forma radiale strutturata come una gerarchia di menù

³⁹ A. L. Yarbus, *Eye Movement and Vision*, Springer Science+Business Media, New York (1967).

⁴⁰ L. Sidenmark et al., *Radi-Eye - Hands-Free Radial Interfaces for 3D Interaction using Gaze-Activated Head-Crossing*, in *CHI '21 Conference Proceedings*, Yokohama (2021), online.

anulari concentrici, che mantiene sull'anello centrale le opzioni principali e d'uso più comune e visualizza i sottomenù una volta che si rendono necessari al proseguimento delle operazioni dell'utente. Il tracciamento di capo e occhi è stato operato tramite un casco di Realtà Virtuale HTC Vive integrato con un sensore oculare Tobii. Il già citato sistema indossabile Hololens 2 sfrutta la stessa combinazione di segnali per il tracciamento oculare.

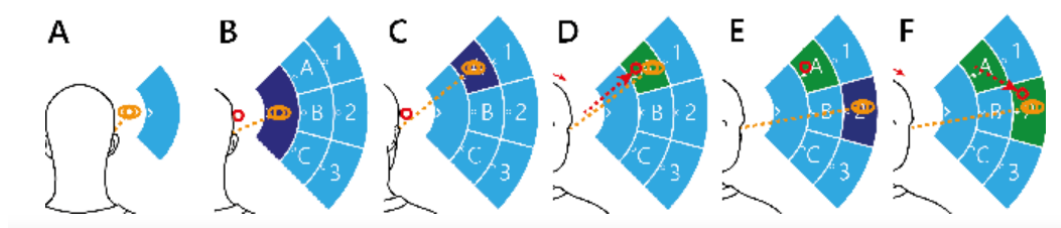


Fig. 14: Schema di una sequenza di selezioni tramite Radi-Eye. Immagine estratta dalla pubblicazione.

Radi-eye delega all'interfaccia digitale e alla sua struttura, e in particolare al suo progressivo decentramento man mano che l'utente si concentra su operazioni gerarchicamente inferiori, la risoluzione del problema della precisione del tracciamento. La sua interfaccia digitale opera inoltre tramite *crossing*, ossia permettendo all'utente di selezionare un'opzione attraversandola, un'operazione più semplice da tracciare rispetto a una fissazione precisa. Gli autori hanno progettato il loro sistema su un preciso presupposto:

Gaze is fast and effortless in directing attention to objects while head movement is more precise for control. However, both modalities lack an intrinsic mechanism for selection and expose Midas Touch issues as they are 'always on'⁴¹.

In un altro sistema dotato di tracciamento oculare, chiamato GazeTap, anche Hatscher e colleghi⁴² identificano il Tocco di Mida come il principale limite di

⁴¹ Ibid.

⁴² B. Hatscher et al., *GazeTap: Towards Hands-Free Interaction in the Operating Room*, in *ICTMI'17 Conference Proceedings*, Glasgow (2017), pp. 243-251.

questo tipo di sistema. L'effetto è di diffusa notorietà nell'ambiente della progettazione di interfacce: si tratta della caratteristica di determinate modalità di *input* di non poter essere "disattivate", e fornire dunque segnali indesiderati alla macchina che legge ogni variazione come un segnale significativo e intenzionale. Per questo motivo, è molto comune che le interfacce a tracciamento oculare e altri sistemi esposti all'effetto Tocco di Mida vengano integrate con altre modalità di *input* atte a significare i comandi forniti ocularmente o a identificarli come tali, dichiarando alla macchina quale intervallo o istante di tempo considerare saliente nell'acquisizione dell'*input*. Sistemi di interfaccia dotati di comandi pertinenti a differenti sfere sensoriali o agentive distinte, e in particolare quelle che sfruttano differenti modalità di *input* e *output* in modi e tempi separati come la sala operatoria GazeTap di Hatscher, che fa uso di un pavimento tattile, un sistema di tracciamento oculare, due laptop e un monitor di grandi dimensioni, prendono il nome di interfacce multimodali. Poiché il loro carattere fondamentale è quello dell'intersezione di forme precisamente delimitate di interfaccia e nient'altro, non si tratta di interfacce classificabili secondo le loro caratteristiche *hardware* ma solo secondo la presenza di multiple e differenti interfacce *hardware*, e saranno più correttamente classificate in seguito.

Interfacce organiche

Il tipo di interfacce trattato in questo paragrafo è di più complessa denominazione, e andrà descritto con la cura appropriata. Si tratta infatti di un tipo di interfaccia hardware il cui rapporto con il corpo e l'ambiente non è descritto dai precedenti, ossia quello delle interfacce che ricevono *input* direttamente dai corpi e dai luoghi. Si tratta di interfacce di grande interesse per il settore protesico, così come per il settore della ricerca, alla data di oggi poco presenti nel mercato delle tecnologie quotidiane. In questa categoria tassonomica sono incluse tutte le tecnologie che interagiscono tramite segnali elettrici o chimici biologicamente generati.

Va innanzitutto delineato un confine, ossia quello tra interfaccia uomo-macchina e sensore. Il presente lavoro si è occupato sinora di interfacce uomo-macchina ad uso intenzionale. È immediato osservare come alcuni dei sistemi già descritti nei precedenti paragrafi, e in particolare quello sul tracciamento e l'interfaccia cinetica, pertengano al settore della sensoristica remota. Si potrebbe argomentare che un'interfaccia hardware è poco altro che un sistema di sensori e un qualche gruppo di dispositivi che comunichi a un utente una risposta a come l'utente stimoli tali sensori. Vi è tuttavia una fondamentale differenza tra un sistema che traccia un movimento corporeo e un elettro-cardiogramma: non si può decidere di controllare coscientemente un sistema tramite il proprio battito cardiaco⁴³. Descrivendo dunque interfacce controllabili, la restrizione più intuitiva è quella che limita la categoria ai dispositivi connessi al sistema nervoso: le interfacce cerebrali. Queste non sono però le uniche pertinenti.

Il primo ambito applicativo a cui rivolgersi per l'utilizzo di simili tecnologie è quello medico e protesico, e in particolare l'utilizzo di arti sostitutivi controllabili tramite interfacce organiche. L'interfaccia che pilota gli arti protesici può essere controllata secondo tre fondamentali modalità:

- 1) Tramite sensori mioelettrici cutanei

⁴³ L'autore è consapevole della possibilità, quantunque poco probabile, di venir contraddetto tramite futuro controesempio da qualche individuo eccezionale in grado di manipolare il proprio battito tramite eccellenti forme di autocontrollo emotivo e psicologico. Il presente testo non si rivolge a quell'individuo.

- 2) Tramite sensori cerebrali
- 3) Tramite sensori intra-corporei

Il sensore mioelettrico ottiene i propri *input* dai segnali elettrici dei neuroni che controllano il tessuto muscolare. Il principio fondamentale di questa tecnologia è quello secondo cui in tal modo è possibile registrare sia i comandi intenzionali al muscolo sia tutto ciò che i neuroni presenti nei muscoli generano spontaneamente, come i gesti riflessi non consci.

Sistemi di controllo di questo tipo sono comuni negli arti superiori protesici, come nel caso dei True Limb di Unlimited Tomorrow⁴⁴ o di Ability Hand di Psyonic⁴⁵. Un sistema molto denso di connessioni mioelettriche legge dall'esterno del braccio i segnali elettrici tramite i quali i muscoli del braccio verrebbero normalmente controllati, e li trasmette al sistema di controllo del braccio bionico. Simili interfacce, il cui ambiente operativo è composto da una grande varietà e complessità di segnali, fanno recentemente uso di algoritmi di apprendimento per raffinare la risposta all'input e adattarsi all'utente.

Il caso delle interfacce cerebrali (Brain-Computer Interfaces) è, ad oggi, associato a tecnologie più sperimentali e la cui sperimentazione risulta inevitabilmente controversa, a causa della fragilità del cervello umano. Esistono però già sul mercato dispositivi non invasivi che fanno uso di segnali elettrici raccolti sul capo dell'utente, come i prodotti MindWave di NeuroSky⁴⁶. Tra i diversi tipi di *input* che i prodotti MindWave possono utilizzare, quello degno di nota per il presente paragrafo è l'elettroencefalogramma personale. Tramite l'applicazione di due elettrodi, uno al lobo auricolare (che funge da riferimento elettrico) e uno sulla fronte, è possibile registrare e utilizzare come segnale di ingresso la corrente alternata del lobo prefrontale, e in particolare la sua intensità. La semplicità del dispositivo può però essere il suo limite: un'interfaccia composta di un solo valore,

⁴⁴ <https://www.unlimitedtomorrow.com/how-are-prosthetic-arms-controlled/>, ultima visita 03/12/24.

⁴⁵ *App Setup Guide*, Psyonic, reperibile online al sito di Psyonic presso <https://www.psyonic.io/users>, ultima visita 5/11/24.

⁴⁶ Ancora una volta, il dispositivo è stato scelto per familiarità d'uso. La documentazione relativa si può reperire online presso <https://store.neurosky.com/pages/mindwave>, ultima visita 5/11/24.

quantunque variabile, non permette una traduzione efficace dell'intenzione dell'utente al sistema se quell'intenzione si fa complessa. Per questo, sistemi di interfaccia cerebrale più costosi e raffinati (ma pur sempre indossati esternamente) tendono a necessitare di complessi sistemi di elettrodi sul capo: ne sono esempio i popolari dispositivi della serie EPOC di EMOTIV⁴⁷, dotati di un elevato numero di elettrodi sul capo la cui complessità permette, ad esempio, di controllare un arto artificiale⁴⁸ o un computer dotato di interfaccia grafica classica, con un puntatore a freccia e la possibilità di ricevere *input* tipici di una tastiera⁴⁹. Anche in questi casi, come per il mioelettrico, si tratta per lo più di interfacce per sistemi protesici o di sistemi protesici come interfacce, con la notevole eccezione dell'ambito della robotica⁵⁰. Il principale limite di simili interfacce è biologico, ed è situato nella cute dell'utente. La pelle impone un'interfaccia organica tra gli elettrodi e i segnali elettrici nervosi, e la sudorazione può alterare la conduttività della superficie causando errori di lettura (su queste variazioni di conduttività si basa, ad esempio, la misurazione della risposta galvanica cutanea, o GSR).

Ciò è particolarmente vero nel caso delle interfacce mioelettriche, che necessitano di operare con segnali che provengono dal tessuto muscolare. Diverse persone presentano comportamenti cutanei e caratteristiche fisiche diverse, e per adattarsi a questa varietà di istanze compagnie come Psyonic e Unlimited Tomorrow ricorrono all'apprendimento tramite intelligenza artificiale per raffinare le risposte dei loro dispositivi una volta che l'utente ha iniziato a farne uso.

⁴⁷ Documentazione reperibile al sito di EMOTIV, presso <https://www.emotiv.com/products/epoc-x>, ultima visita 5/11/24.

⁴⁸ R. Bousseta et al., *EEG Based Brain Computer Interface for Controlling a Robot Arm Movement Through Thought*, in *IRBM*, Volume 39, No. 2 (2018), pp. 129-135.

⁴⁹ B. Sumak et al., *An Empirical Evaluation of a Hands-free Computer Interaction for Users with Motor Disabilities*, in *Journal of Biomedical Informatics*, 96 (2019), online.

⁵⁰ M. Bryan et al., *An Adaptive Brain-Computer Interface for Humanoid Robot Control*, in *IEEE-RAS International Conference on Humanoid Robots proceedings* (2011), pp. 199-204.



Fig. 14: EPOC X di Emotiv. Immagine promozionale ricavata sul sito di Emotiv.

Superare fisicamente il limite cutaneo significa, invece, invadere l'organismo dell'utente, in maniere che genericamente implicano una maggiore difficoltà di rimozione del dispositivo. Se ciò è controverso nel caso delle tecnologie intracraniche, un'alternativa di grande interesse è offerta dai sistemi a sensori intra-ossei. Accedere direttamente ai nervi, ad esempio tramite installazione di sensori nervosi, permette di minimizzare il rumore cutaneo e ottenere segnali più precisi anche nel caso di *input* a basso volume. È questo il caso del lavoro del professor Ortiz-Catalan⁵¹. Anche in questo caso, gli arti bionici vengono controllati direttamente dal paziente come se si trattasse del loro arto biologico, dopo un periodo di addestramento del paziente tramite un sistema di Realtà Aumentata e delle sessioni di addestramento del sistema tramite apprendimento automatico⁵². Rispetto a interfacce cutanee, un sistema come quelli ricercati dalla squadra di Ortiz-Catalan richiede l'installazione chirurgica di diversi sensori intra-corporei

⁵¹ M. Ortiz-Catalan et al., *An Osseointegrated Human-Machine Gateway for Long-term Sensory Feedback and Motor Control of Artificial Limbs*, in *Science Translational Medicine*, 6 (2014), online.

⁵² M. Ortiz-Catalan et al., *A Highly Integrated Bionic Hand with Neural Control and Feedback for Use in Daily Life*, in *Science Robotics*, 8 (2023), online. Negli ultimi due anni, il lavoro sul fronte protesico di Ortiz-Catalan si è intensificato e arricchito a causa del suo impegno sul territorio ucraino. Molte delle informazioni qui riportate possono essere riscontrate o arricchite tramite le numerose pubblicazioni simili che il suo gruppo ha prodotto durante il conflitto.

(uno per il nervo mediano, uno per il nervo radiale e uno per il nervo ulneare), ma permette di agganciare il braccio bionico all'ulna e al radio del paziente attraverso l'uso di viti in titanio, utilizzandone la rotazione fisica come ulteriore sistema di controllo.

Interfacce a pulsantiera

Le interfacce a pulsantiera sono, come le interfacce a tastiera, eredità del mondo pre-digitale. Seppur basate sulla discretizzazione del comando di pressione digitale (nel senso letterale del termine, “relativo al dito”), se ne differenziano per la loro forma, che dipende dal dispositivo che devono comandare. La pulsantiera è evidentemente uno strumento fortemente manuale, e come la tastiera accentra l’importanza della postura e della posizione della mano: la maggior parte delle pulsantiere, e la loro principale differenza dalle pulsantiere di minoranza, è nella loro impugnabilità.

Le varianti più comuni e di facile riconoscimento, nonostante possano istanzarsi in forme estremamente diverse, ricadono nella datata categoria del telecomando. Ad accomunarle è però proprio la fondamentale distanza dall’interfaccia a tastiera: lungi dall’essere un dispositivo universale, telecomandi, controller e plance specializzate sono strumenti altamente comunicativi, che si fanno carico dell’insegnare all’utente a navigare lo spazio agentivo del dispositivo che controllano. Nella catena semiotica delle interfacce, dunque, svolgono un lavoro di normalizzazione più esplicito e dichiarativo rispetto a interfacce più trasversali, poiché nessun dispositivo controllato da uno specifico controllore a pulsantiera ha raggiunto l’universalità e la trasversalità del personal computer, e nessun tipo di controllore a pulsantiera è mai divenuto trasversale come la tastiera. Se per ogni personal computer si utilizza una tastiera, genericamente QWERTY, nemmeno la diffusione secolare delle televisioni ha mai prodotto un telecomando universale. Ogni dispositivo viene venduto col suo telecomando, dotato della sua pulsantiera, del suo codice semiotico e di un set di funzioni precisamente mappato sul telecomando e sulla sua legenda. Al contrario, la variazione sulle tastiere è marginale.

La forma più trasversale dell’interfaccia a pulsantiera, dopo il telecomando, è quella del controller: dal controllo di droni a quello delle console videoludiche, dispositivi di *input* da impugnare con ambedue le mani e dotate sia di pulsanti sia di leve analogiche per fornire indicazioni direzionali si incontrano in ambiti molto diversi.

Ma anche nel caso dei controller, differenti brand, esigenze e ambiti tecnologici hanno generato negli anni innumerevoli esemplari differenti, che istanziano diverse necessità di specificazione o universalizzazione dell'*input*, in maniera antipodale a ciò che avviene nei telecomandi.

Un particolare effetto di questo fenomeno è nella ricerca e nell'istituzione di simboli vuoti. Quando il controller diviene dispositivo di *input* trasversale come lo è una tastiera, esso deve affrontare l'impossibilità (che tipicamente ha forma di non-necessità) di dotarsi di un intero apparato alfanumerico di pulsanti. Ciò è tanto più vero quanto maggiore è la velocità operativa richiesta all'utente dal dispositivo. Il pulsante stesso è protagonista di questo svuotamento semiotico, poiché una volta svuotato di un significato imposto non gli resta altro senso che non siano l'appartenenza al sistema e le sue caratteristiche materiali, e diviene così più vicino a puro mezzo, al puro simbolo. Nelle sale arcade, sulle plance dei cabinati videoludici (così come sui joystick CX40 prodotto da Atari⁵³), i pulsanti non presentavano neppure simboli identificativi: il loro colore e la loro posizione erano tutto ciò che permetteva di distinguerli, riducendoli al loro minimo materiale. Nei decenni della storia delle console, invece, i diversi brand e i loro diversi dispositivi hanno assegnato, per facilità comunicativa e mnemonica, lettere o simboli puri a diversi tasti, sempre biunivocamente determinati dalla loro posizione su un controller che, nel caso delle console, è invece impugnato. Così, nessun significato universale pertiene al tasto Quadrato o Triangolo su un DualShock Sony che non sia, rispettivamente, quello di pulsante sinistro o superiore nella metà destra del DualShock, nonostante differenti periodi storici durante le diverse generazioni di console Sony abbiano visto una tendenza a universalizzare l'uno o l'altro come tasto delegato alla funzione Undo o Indietro. La pulsantiera del controller Sony è il più eloquente tra gli esempi di segni puri, utilizzati poi da diversi software per costruire il linguaggio della loro specifica agentività. È altresì possibile far uso di controller

⁵³ M. J. P. Wolf, *Before the Crash. Early Video Game History*, Wayne State University Press, Detroit (2012).

così strutturati su un personal computer: diversi software di interfaccia⁵⁴ da controller a tastiera permettono la traduzione dei segnali dell'uno in specifici tasti dell'altra (ad esempio, mappando il pulsante X sul tasto Invio).

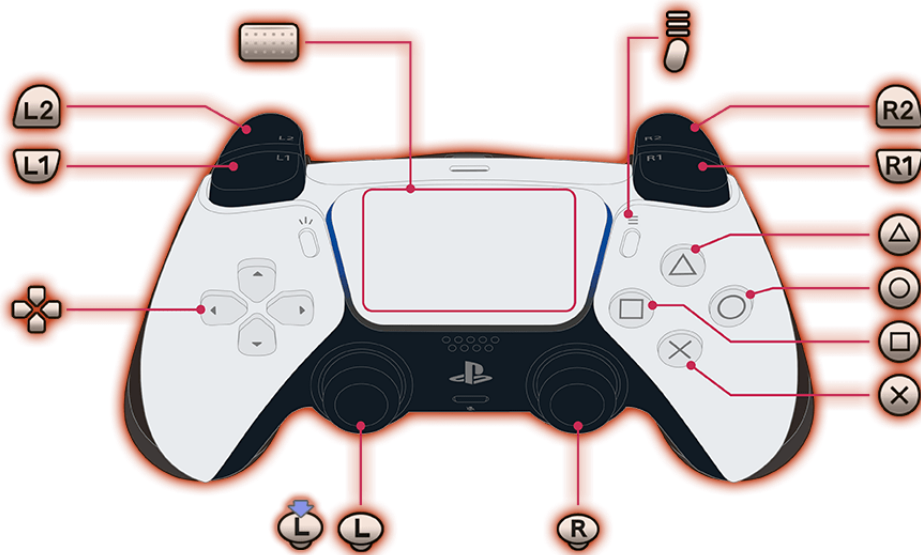


Fig. 15: La pulsantiera di un controller Dualshock per PlayStation 5.

Al di fuori dello speciale ambiente delle console videoludiche, tuttavia, l'utilizzo di pulsantiere universali è di scarsa attrattiva. L'utilizzo di un telecomando universale non cablato presuppone protocolli universali di comunicazione senza fili e riporterebbe sui singoli dispositivi la necessità di mappare le attivazioni del telecomando sulle opzioni d'uso di un televisore o di una lampada, implicando la necessità e l'inattività di diverse quantità e gruppi di pulsanti in base al dispositivo pilotato, e generando inevitabile confusione. Così ogni dispositivo domestico viene dotato della sua interfaccia, le azioni possibili coincidenti con i comandi impartibili e la sua legenda. Emerge dunque il problema opposto: architettare simbologie e

⁵⁴ Nel caso del DualShock Sony, si veda l'esempio di DS4Windows, reperibile presso <https://ds4-windows.com/>, ultima visita 7/11/2024.

interfacce fisiche in grado di comunicare correttamente all'utente le sue possibilità d'azione.

La mappatura (e con essa la costrizione) delle possibilità agentive dell'utente di fronte a uno specifico dispositivo è un processo iterato e complesso. Come già avvenuto in questo lavoro, per presentare un caso chiaro di adattamento a specifiche necessità, estrarremo dal design di sistemi per individui non normati, che offre un vasto arsenale di esempi virtuosi. Nello studio del sistema di controllo CandyFly, Garcia e Brock hanno dedicato numerosi workshop in un lavoro di quattro anni alla prototipazione e al test in vivo di una pulsantiera a plancia per il controllo ricreativo di droni destinati a un pubblico di persone affette da differenti forme di disabilità, sia fisiche che cognitive⁵⁵. Dopo una fase iniziale di test con tastiere e controller di diversi tipi, i ricercatori hanno implementato diverse plance con pulsanti di grandi dimensioni e a forma di freccia direzionale. Diverse plance permettevano un maggiore o minore controllo delle direzioni di volo. Dalle discussioni dei ricercatori con i terapeuti dei partecipanti affetti da disabilità cognitive, il fattore più importante per un'interfaccia di controllo efficace è risultata essere l'evidenza del rapporto di causalità tra l'azione dell'utente che pilota il drone e il movimento del drone stesso. Per questo motivo, sistemi di controllo cinetico o tangibile sono stati esclusi dalla sperimentazione, ma hanno permesso agli utenti e ai loro accompagnatori di apportare aggiustamenti in tempo reale al funzionamento dei comandi. Hanno anche osservato come, in alcuni casi, escludere interamente un'asse spaziale dalle operazioni possibili abbia migliorato il controllo del drone e la sua sicurezza, minimizzando il rischio di incidenti, senza che a ciò corrispondesse un calo della soddisfazione riportata dai piloti.

Per quanto la pulsantiera sia genericamente associata a una pressione digitale o manuale, esistono eccezioni. È innanzitutto possibile istituire pulsantiere podali.

⁵⁵ J. Garcia, A. M. Brock, *CandyFly: Bringing Fun to Drone Pilots with Disabilities Through Adapted and Adaptable Interactions*, in *CHI '22 Conference Proceedings*, New Orleans (2022), online.

Nella già citata interfaccia della sala operatoria a mani libere⁵⁶, il chirurgo può controllare l'*output* visuale dei monitor attraverso un pavimento tattile. Il sensore può percepire la pressione del piede sul posto o sulle posizioni adiacenti alla posizione iniziale (avanti, indietro, destra e sinistra), leggendo sia lo spostamento sia il numero di pressioni operate, e integra i comandi così impartiti con un sensore di tracciamento oculare in maniera multimodale. Il sistema è basato sul lavoro di Çöltekin e colleghi⁵⁷ sull'interfaccia multimodale oculo-podale. Il lavoro del gruppo di Çöltekin faceva utilizzo di una piccola piattaforma di forza e una telecamera nell'utilizzo di un computer da scrivania, utilizzando differenti gesti combinati dei piedi (come sollevarli entrambi o simulare una camminata per ingrandire o rimpicciolire un'immagine). Il sistema di Hatscher invece fa un uso più granulare del gesto podale, mantenendo un vocabolario gestuale più semplice alla ricerca di una "mappatura naturale del gesto del piede" ("Natural Mapping of Foot Gestures"⁵⁸). Non esiteremo tuttavia a considerare anche il lavoro di Çöltekin parte integrante della categoria di interfacce qui descritta: nonostante la loro forma più complessa e coordinata, i gesti necessari a pilotare il loro sistema rimangono privi di un senso intrinseco nel contesto dell'interazione utente-macchina (si potrebbe argomentare che lo siano più di quelli di Hatscher e colleghi proprio grazie alla loro ricerca di un gesto più naturale) e di natura impulsiva, inteso non nel significato di gesti dettati dalla pulsione ma di gesti a singolo impulso.

Infine, va evidenziato come un sensore in grado di registrare un singolo gesto – e in particolare un cambio di pressione o di distanza – possa essere progettato per, essenzialmente, ogni parte mobile del corpo, e in particolare per ogni area del corpo che possa essere coscientemente spostata. Se ne trova un esempio perfetto nella ricerca di Röddinger e colleghi⁵⁹, che implementa un "pulsante" attivato tramite

⁵⁶ B. Hatscher et al., *GazeTap: Towards Hands-Free Interaction in the Operating Room*, in ICTMI'17 conference proceedings, Glasgow (2017), pp. 243-251.

⁵⁷ A. Çöltekin et al., *Gaze and Feet as Additional Input Modalities for Interacting with Geospatial Interfaces*, in *ISPRS Annals of the Photogrammetry Remote Sensing and Spatial Information Sciences*, III-2 (2016), pp. 113-120.

⁵⁸ Ibid.

⁵⁹ T. Röddinger et al., *EarRumble: Discreet Hands - and Eyes - Free Input by Voluntary Tensor Tympani Muscle Contraction*, in *CHI '21 Conference Proceedings*, Yokohama (2021), online.

contrazione del timpano. Stando ai dati dei ricercatori, statisticamente, il 56.8% dei potenziali lettori di questo testo avrà bisogno di un ulteriore chiarimento riguardo la modalità di *input*. Esiste infatti in parte della popolazione umana – secondo una ricerca statistica operata dagli autori su 192 individui, circa il 43.2% – in grado di contrarre volontariamente i muscoli del timpano con diversi livelli di complessità e abilità. La forma più basilare di questa capacità è la contrazione muscolare dell’intera area, che causa la percezione di un costante rombo in grado di attutire e ovattare la maggior parte dei suoni provenienti dall’ambiente esterno, e in particolare quelli a bassa frequenza – una sorta di grezzo filtro passa-alto anatomico. È inoltre possibile, per alcuni, variare impulsivamente la pressione del timpano, di fatto “stappandosi le orecchie” efficacemente senza l’ausilio di gesti ulteriori, come l’otturazione del naso o la deglutizione⁶⁰. Il lavoro dei ricercatori aveva scopo esplorativo: il sistema EarRumble è stato utilizzando come fonte di *input* unico, mappandone l’output su “click” di diversa durata. Nondimeno, il gesto monodimensionale di contrazione del tensore timpanico potrebbe permettere la costruzione di un vocabolario, seppur limitato, di gesti discreti. Inoltre, a meno di situazioni particolari, il sistema presenta l’eccezionale carattere di immunità dall’effetto Tocco di Mida nonostante sia controllato da un organo costantemente attivo.

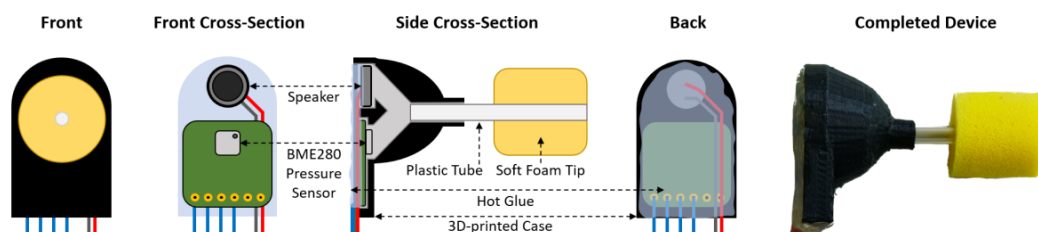


Fig. 16: Schema della struttura di EarRumble. Immagine estratta dalla pubblicazione.

⁶⁰ Deve essere qui annotato che l’unica abilità descritta nel lavoro di Röddinger e colleghi è quella di contrarre il muscolo tensore timpanico e causare il rombo. Per le restanti informazioni la fonte è diretta, e l’autore conosce il fenomeno in prima persona.

Ciò che appare da sistemi come GazeTap e EarRumble è che essi finiscono inevitabilmente per vincolare la parte del corpo a cui l'interfaccia si rivolge. Una pedaliera implica l'impossibilità di spostarsi senza perdere il controllo dell'interfaccia, e un dispositivo inserito nell'orecchio a misurare la pressione del canale auricolare implica l'impossibilità di muovere rapidamente il capo⁶¹. Ciò è nondimeno vero anche nel caso delle mani, ma esse rimangono ancora una volta una norma corporea forte, la cui disponibilità è un assunto a priori. Le due interfacce qui citate sono esempi di ricerca nell'ambito delle interfacce "hands-free", la cui finalità è quella di sostituire le mani per poterle reimpegnare in altro⁶² o per compensare la mancanza delle stesse.

Va ricordato che le interfacce a pulsantiera possono essere applicate (seppur non a finalità di liberazione manuale) anche a corpi non umani, integrando altri esseri viventi come parti di una catena di interfacce di livello superiore a quella organica a loro applicata. Divenuto celeberrimo grazie a un articolo del *The Economist* del 2021⁶³, il sistema di controllo acquifero *Gruba Kaska* di Varsavia è basato su un'interfaccia organica per vongole. I molluschi, di estrema sensibilità ad alcuni elementi tossici all'interno dell'acqua, rispondono a un aumento dei livelli di tali elementi chiudendo il loro guscio. Le otto vongole del sistema di sicurezza di Varsavia sono state dotate di interruttori che si attivano alla chiusura dei bivalvi, segnalando agli operatori e al sistema le condizioni di rischio. Differenti corpi necessitano però, come questo caso evidenzia, di interfacce *ad hoc* per essere correttamente tracciati, e costituiscono casi più eccezionali che regolari. Per quanto si tratti di sistemi estremamente affascinanti e di grande interesse scientifico ed ecologico, che in forme non intrusive ed eticamente solide potrebbero rappresentare un'importante frontiera per una maggiore integrazione di piante e animali nelle

⁶¹ Si noti che il sistema EarRumble fa uso di una spugna per rimanere all'interno del canale auricolare in maniera meno intrusiva possibile, ed è pertanto suscettibile a movimenti bruschi.

⁶² Nei paper citati in questi paragrafi, non è raro l'utilizzo del termine "manipolazione" per descrivere i gesti compiuti con altre parti del corpo o i loro effetti su parti dell'interfaccia digitale.

⁶³ *How Clams help Keep Polish Water Clean*, presso <https://www.economist.com/europe/2021/01/21/how-clams-help-keep-polish-water-clean>, ultima visita 8/11/2024.

infrastrutture e nelle città, l'assenza di un utente umano ad interfacciarsi intenzionalmente per impartirvi comandi le pone al di fuori dell'ambito di approfondimento del presente testo.

Interfacce vocali

Dialogare con la macchina è, ancora una volta un'operazione che precede il personal computer e le interfacce individuali, e ciò è principalmente dovuto al lavoro del laboratorio Bell Labs in New Jersey. Storicamente, le macchine hanno imparato a parlare prima di imparare ad ascoltare. Nell'articolo di Vannevar Bush del 1945⁶⁴, ampiamente discusso nel capitolo sulla storia e l'archeologia dell'interfaccia grafica, viene citato un dispositivo chiamato Voder, esposto nel 1939, in grado di produrre suoni riconoscibili come parole umane ricevendo comandi da una tastiera operata da una ragazza, senza che nella macchina venisse immesso alcun suono. Solamente undici anni dopo, nel Novembre del 1950, fa la sua comparsa un articolo di Jean Dreyfus-Graf intitolato *Sonograph and Sound Mechanics*⁶⁵. L'autore descrive la sua invenzione come segue:

I constructed the word sonograph from the Latin "sonus," "sound," and from the Greek "graphé," which means an "action" of engraving or of writing. It designates a new class of electracoustical instruments. Such an instrument comprises essentially an electronic part (impulse generator) converting sounds into simplified groups of electric impulses, and a mechanical part, which transforms these impulses into characteristic actions like graphic symbols or remote control. When this mechanical part includes a drawing oscillograph or a typing machine, we talk about a steno- or a typo-sonograph respectively. The electrical part may select predominantly speech-characteristics of sounds; then we speak about a phonetic sonograph (or phoneto- graph). This kind of instrument may translate speech sounds into a natural stenography or into the conventional alphabetical letters.

La macchina riconosceva pattern sonori molto semplici, senza attribuirgli significato, della durata media di una sillaba. Due anni dopo, nel 1952, ancora una volta nei Bell Labs, fece la sua comparsa AUDREY (AUtomatic Digits

⁶⁴ V. Bush, *As We May Think*, in *The Atlantic Monthly* (Luglio 1945).

⁶⁵ J. Dreyfus-Graf, *Sonograph and Sound Mechanics*, in *The Journal of the Acoustic Society of America*, Vol. 22 (1950), pp. 731-739.

REcognition machine), e per la prima volta una macchina iniziò a distinguere il linguaggio umano⁶⁶. Il sistema AUDREY poteva riconoscere solo i nomi delle singole cifre, ma era in grado di riconoscere le variazioni di frequenza del suono al netto delle differenze inflessionali e vocali dei differenti interlocutori. L'approccio differenziale permetteva di ignorare le differenze tra voci acute e voci gravi. Questa è, con ogni probabilità, la caratteristica più fondamentale di ogni interfaccia vocale dal punto di vista fisico: la capacità di leggere l'*input* con buona approssimazione. Dal punto di vista puramente hardware, un'interfaccia vocale strettamente intesa non è che un microfono associato a una qualunque forma di *output*, possibilmente ma non necessariamente di natura acustica. Per questo motivo, si tratta di una delle interfacce maggiormente afflitte dal problema del Tocco di Mida. Settant'anni dopo la comparsa del sonografo e di AUDREY, e con la diffusione di agenti vocali come Siri e Cortana e l'esplosione del mercato dell'intelligenza artificiale, le interfacce vocali possono ascoltare qualunque forma di *input* verbale ed elaborare strutture ben più complesse e contestuali di una singola cifra in lingua inglese. "Qualunque cosa" implica, naturalmente, sia ciò che è rivolto all'interfaccia, sia ciò che non lo è, con conseguenze che possono andare dal sorprendente all'eticamente dubbio. Murad e colleghi attribuiscono le cause delle difficoltà di usabilità delle interfacce vocali all'assenza di buone pratiche di design diffuse e insegnate fino al livello universitario, dove da molto tempo le discipline del Design di Interfacce, Design di Esperienza Utente e Design di Interazione si concentrano sull'interfaccia grafica⁶⁷. Ciò è più che mai rilevante nell'era dell'IoT controllato da dispositivi come Amazon Alexa e Google Home, la cui installazione impone sugli utenti un orecchio in costante ascolto nell'ambiente domestico. Un caso studio analizzato da Murad durante una lezione sull'argomento (il caso del video "Italian grandmother learning to use google home"⁶⁸, che ha ricevuto a Novembre 2024 più di sei milioni di

⁶⁶ R. Pieraccini, *The Voice in the Machine. Building Computers That Understand Human Speech*, The MIT Press, Cambridge MA (2012).

⁶⁷ C. Murad et al., *Designing Voice Interfaces: Back to the (Curriculum) Basics*, in *CHI '20 Conference on Human Factors in Computing Systems Conference Proceedings* (2020), online.

⁶⁸ Italian grandmother learns to use Google Home, <https://www.youtube.com/watch?v=e2R0NSKtVA0> (2018), ultima visita 12/11/2024.

visualizzazioni su YouTube), evidenzia come la prima difficoltà d'uso di simili interfacce emerga dal comando di attivazione dell'interfaccia vocale, l'esclamazione "Hey Google" o "Okay Google" necessaria ad avviare la ricezione di comando, di difficile comprensione per l'anziana utente del dispositivo. L'attivazione è la soluzione impiegata da Google, come da Amazon, al problema del Tocco di Mida: una parola chiave che permetta di distinguere l'intervallo temporale in cui tutto ciò che viene detto è da considerarsi rivolto alla macchina dal normale tempo di "non-ascolto".

Se dal punto di vista ingegneristico si tratta di una soluzione al problema di separazione del segnale dal rumore, da un punto di vista dell'esperienza significa "denaturalizzare" la conversazione con la macchina. Il fascino delle interfacce vocali sta nella loro capacità di prestarsi a interazioni "naturali", che non interrompano il flusso d'azione e pensiero dell'utente mentre chiede distrattamente qualcosa alla macchina e ne viene accontentato. L'attivatore spezza quest'illusione, riportando sempre l'utente alla consapevolezza di star impartendo un comando meccanico, un preciso "gesto verbale" attraverso un'interfaccia che si opacizza nel processo e che, come ogni interfaccia fortemente gestuale, vincola l'utente al proprio vocabolario di segni e il suo corpo ai ritmi e alle tecniche del dispositivo. D'altra parte, la non separazione dei momenti di tempo implica l'ascolto costante. Senza alcuna necessità di scomodare le complesse implicazioni biopolitiche dell'impossibilità di interrompere la generazione di dati privati a guadagno di un ascoltatore anonimo in un paradigma post-capitalista di costante compravendita di dati, ci si può rivolgere al più innocente caso studio del videogioco NBA 2015⁶⁹, e in particolare alla sua versione per la console Xbox One che, come già precedentemente riportato, era dotata di un sensore Kinect nel suo kit base. Ciò che non è stato precedentemente accennato è il fatto che quel sensore Kinect era anche dotato di un microfono aperto, che permetteva di interagire vocalmente con la console. Il titolo ufficiale dell'associazione cestistica americana decise di farne pieno uso in una maniera che colse molti utenti alla sprovvista: durante la partita, a

⁶⁹ Visual Concepts, *NBA 2K15*, 2K Sports (2014).

valle di un'imprecazione del giocatore, convinto di essere isolato sul suo divano e non ascoltato da nessuno, il gioco poteva attribuire alla sua squadra un fallo tecnico nella partita virtuale come se il giocatore avesse imprecato su un effettivo campo di gioco.

A cercare di sfruttare la costante apertura del microfono del Kinect erano inoltre i giocatori stessi: non era raro, nel primo periodo dopo il rilascio della console, "incontrare" in rete molti utenti il cui nickname coincideva con un preciso comando vocale, come "XboxShutDown". Questo faceva sì che, nel tentativo di comunicare tramite una chat vocale con quei giocatori o nominandoli per commentarne l'operato, la console si spegnesse o eseguisse altre operazioni contro la volontà dell'utente.

Se il tentativo di alcuni utenti di prendersi gioco dei loro compagni di squadra meno attenti tendeva a non riscontrare troppo successo, esso può non di meno ricondurci alla complessità problema della distinzione del rumore dal segnale in presenza dell'effetto Tocco di Mida, evidenziando l'esistenza di casi in cui essi sono assolutamente indistinguibili per l'interfaccia vocale, e ciò è tanto più vero quanto più la conversazione con la macchina si avvicina a una conversazione reale. Ciò che la macchina non può inferire, a meno di non possedere altri dati, è il contesto di tale interazione. Per questo motivo, l'interfaccia vocale è comune nei sistemi multimodali, di cui abbiamo già presentato diversi esempi.

Il secondo, e più evidentemente esclusivo, principale problema delle interfacce vocali è più banale ed evidente: la distinzione del segnale dal rumore acustico. La soluzione dipende dal – e spesso consiste nel – posizionamento adeguato dell'interfaccia nello spazio in relazione all'utente. Difficilmente sarà necessario imporre filtri particolarmente complessi nell'interfaccia se essa è destinata a un quieto ambiente domestico. Differente sarà l'approccio nel progettare un sistema di comandi vocali, per quanto chiari e semplici, da utilizzare all'interno di un elicottero in volo⁷⁰. La gestione di questo secondo problema è però interamente

⁷⁰ L'argomento, di notevole complessità, è stato affrontato dall'autore in alcune interessantissime conversazioni private con Stéphane Conversy, docente di Interazione Uomo-Macchina dell'ENAC di Tolosa. Si tratta, ancora una volta, di sistemi multimodali la cui finalità è utilizzare la voce per

legata al contesto, ed è difficilmente risolvibile sul piano dell'interfaccia *hardware*. L'elaborazione del segnale dopo la sua ricezione può affidarsi a molti strumenti, dal filtraggio al riconoscimento di segnali tramite reti neurali, ma è un compito delegato a un livello successivo della catena di traduzione del messaggio e dunque a progettisti di differenti parti del sistema.

È in ultimo opportuno catalogare tra le interfacce vocali i sistemi che fanno uso di *output* puramente sonori. Nonostante le soluzioni acustiche siano molte e principalmente non di carattere vocale, queste difficilmente riescono a costituire interfacce vere e proprie. La causa è ben spiegata da Donald Norman, nel suo storico testo *The Design of Everyday Things*:

Feedback must also be informative. Many companies try to save money by using inexpensive lights or sound generators for feedback. These simple light flashes or beeps are usually more annoying than useful. They tell us that something has happened, but convey very little information about what has happened, and then nothing about what we should do about it. When the signal is auditory, in many cases we cannot even be certain which device has created the sound.⁷¹

La loro immane diffusione è dovuta, secondo Norman, al bassissimo prezzo di “beepers” e luci monocromatiche rispetto a qualunque altro componente di feedback. Non si tratta, evidentemente, di interfacce rilevanti per l'analisi in corso. Più attinenti sono invece le interfacce a lettura, comunemente denominate “text-to-speech”. Simili interfacce sono di particolare utilità per persone affette da cecità, persone ipovedenti e persone affette da dislessia. In altri casi, sono desiderabili dove sistemi di feedback di carattere più visuale non siano possibili, fungendo da interfaccia liberatoria per modalità gerarchicamente superiori, com'è già stato osservato per le interfacce “hands-free”.

permettere alle mani di disimpegnarsi da alcune delle operazioni necessarie per poterne svolgere di ulteriori.

⁷¹ D. Norman, *The Design of Everyday Things*, Basic Books (New York 2013).

Particolarmente significativo è il caso di alcuni dispositivi per non-vedenti. All'autore è capitato, per motivi personali, di trovarsi diverse volte a trascorrere del tempo con l'utente non-vedente di un telefono cellulare che fa uso di un'interfaccia di questo tipo. Il dispositivo, dotato di una tastiera a dodici bottoni del tipo comunemente diffuso prima dell'arrivo degli smartphones arricchiti dall'aggiunta dei numeri in alfabeto Braile, si presenta tanto efficace nella ricezione dell'*input* quanto tragicamente inefficace nel fornire *output*. Il telefono è dotato di uno schermo, indirizzato principalmente a chi assiste l'utente nella risoluzione di task complessi, che funge da oggetto di lettura per l'interfaccia vocale. La voce del dispositivo opera infatti enunciando le varie componenti salienti e le diverse opzioni presenti sullo schermo mentre l'utente le scorre tramite i pulsanti meccanici. Si tratta sostanzialmente di un leggero adattamento dell'interfaccia grafica, completamente visuale nel suo paradigma. In altre parole, è la trasformazione di uno schermo in una sequenza di descrittori vocali. Ciò è particolarmente evidente quando la disposizione spaziale degli elementi grafici non offre un criterio evidente per l'ordine di lettura. Lasciare che la disposizione spaziale funga da struttura di presentazione dell'interfaccia nonostante il suo significato per l'utente non-vedente sia sostanzialmente nullo appare quantomeno controintuitivo, ma non è altro che un ulteriore argomento a favore della tesi di Murad. Nonostante l'indipendenza nella realizzazione dell'individuo sia un valore culturalmente costruito, di matrice americana e fortemente abilista, come ha eccellentemente esposto Garland Thomson in *Extraordinary Bodies*⁷², il telefono cellulare è un dispositivo costituito per incarnare quel valore di indipendenza, più individuale del personal computer e fortemente privato, e ciò può causare prevedibilmente elevati livelli di frustrazione in un utente che debba costantemente rivolgersi ad altri per poterne fare un uso efficace, anche quando l'interfaccia sembra a primo acchito voler sostituire una persona da cui dipendere tramite l'uso della voce sintetica. Lo schermo, con il suo

⁷² R. Garland Thomson, *Extraordinary Bodies: Figuring Physical Disability in American Culture and Literature*, Columbia University Press, New York (1997).

montaggio spaziale, costituisce una norma talmente potente da invadere un reame sensorio che non gli pertiene.

La ricerca di forme di interfaccia vocale non oculocentriche, la cui struttura ed esistenza non dipenda direttamente dalle forme dell'interfaccia visuale classica, è un tema di grande necessità e forte presenza in letteratura, specialmente in tempi recenti. L'avvento dell'Intelligenza Artificiale, che permette interazioni complesse e fortemente didascaliche nell'autonomia dello scambio vocale, potrebbe certo costituire parte della soluzione. Tuttavia, il radicamento del protagonismo grafico e visuale è profondo, e dovrà essere indirizzato se si desidera permettere davvero agli esseri umani di comunicare e ricevere senso da un dispositivo attraverso la sola voce.

B. Interfacce formali

Superato il livello sensoristico, quello più propriamente “trasduttorio”, nella catena di traduzione delle interfacce, si giunge a un secondo livello ontologico dell’interfaccia. Gli effetti di interfaccia presenti a questo livello si rivolgono, sul piano dell’utente, a un’entità costituita dall’umano e dai sistemi di input materiali su cui agisce, e ne direzionano l’agire in maniera più dinamica, modellando l’informazione binaria immagazzinata nella macchina secondo diverse forme, spesso fortemente metaforiche o metonimiche, a partire dal “paradigma WYSIWYG” in poi. Per questo, si tratta del livello che prende spesso il nome di “interfaccia digitale”. È il livello di interfaccia più eloquente, ed emerge dalla composizione degli *output* dei singoli elementi del livello hardware. La separazione da quest’ultimo è inevitabilmente solo parziale: essi si codeterminano nella genesi e nell’evoluzione, adattandosi l’uno all’altro. Così, anche questo livello è dominato da forti norme, quelle dell’interfaccia grafica. La logica dell’interfaccia formale è quella che determina in che modo qualunque oggetto culturale appare ai nostri sensi attraverso un dispositivo digitale. L’universalità di questo processo è sotto ai nostri occhi ogni giorno, e proprio nel suo essere così persistente e pervasivo è la potenza del suo effetto. Manovich ne evidenziava la portata nelle conseguenze della diffusione del computer, prima che divenisse tascabile:

Negli anni Novanta, con la sempre maggior diffusione di Internet, il ruolo del computer si trasformò da tecnologia specifica (calcolatore, processore di simboli, manipolatore di immagini, etc.) a filtro per l’intera cultura, ovvero a forma di mediazione per tutti i tipi di produzione artistica e culturale. Quando, tutt’a un tratto, la finestra di un browser sostituì lo schermo televisivo e cinematografico, la parete di una galleria d’arte, la biblioteca e il libro, la nuova situazione si manifestò in tutta la sua portata. Tutta la cultura, del passato e del presente, veniva ormai filtrata dal computer, con la sua particolare interfaccia uomo-macchina.¹

¹ L. Manovich, *Il linguaggio dei nuovi media*, Merlini, R. (trad.), Edizioni Olivares, Milano (2022), p.90.

L'effetto filtrante qui descritto è duplice. Il computer, in ogni sua forma, filtra ogni oggetto culturale che attraversa la sua lunga catena di interfacce imponendogli le sue gerarchie, le sue metafore e i suoi parametri di giudizio, la sua indiscriminata efficienza e i valori che ne determinano la definizione. Ma filtrare non significa solo modificare: significa includere qualcosa ed escludere qualcos'altro, lasciar passare solo ciò che attraversa il filtro. Così, quando l'interfaccia diviene filtro di "tutta la cultura", riforma la cultura escludendone ciò che non può passare attraverso un'interfaccia. Essa colonizza ogni area culturale verso cui può espandersi, e marginalizza ciò che non può includere. Il cinema e lo schermo televisivo, il muro della galleria d'arte, la biblioteca e il libro sono oggetti scopici che si presentano di fronte a noi, con quasi tutta la loro informazione sul piano frontale, ed eventualmente (come nel caso del cinema o dei non citati monologhi e oggetti musicali) su quello uditivo.

Ciò non pertiene solamente al computer: abbiamo già osservato quanto vasta sia la matrice dei principali dispositivi da cui, per vie parallele e intrecciate, esso si è venuto a creare. Ciò non vale solamente per l'interfaccia di tipo più materiale, ma anche per la genesi dei segni, dei meccanismi e della dinamica di ciò che appare e di ciò che può essere operato nella sfera dell'informazione tramite il metamedium computer, che presenta un esempio eccezionalmente evidente del precetto McLuhaniano secondo cui il contenuto di un medium è sempre un altro medium², e ne eredita e trasforma i modi in cui esso plasma le relazioni tra individui, spazi e tempi.

Così come molte interfacce hardware non si definivano se non tramite la loro differenza rispetto a mouse e tastiera, analogamente molte interfacce formali si distinguono in virtù della loro relazione con l'interfaccia grafica. L'interfaccia grafica non è però un'interfaccia definita secondo le sue inequivocabili specifiche: è un modo di "fare" interfaccia, è la scelta di utilizzare simboli, metafore e metonimie visuali e di proiettare l'utente tramite un puntatore, il tutto a prescindere

² M. McLuhan, *Gli Strumenti del Comunicare*, Capriolo, E. (trad.), Garzanti, Milano (1967).

dalla forma dei simboli o dei puntatori stessi. Per questo, buona parte della sezione dedicata alle interfacce formali sarà costituita di interfacce determinate dal loro regime visuale e dalle logiche del loro montaggio, principalmente ma non esclusivamente spaziale, e dunque dal modo in cui dispongono (o non dispongono) l'informazione.

L'interfaccia formale più comune nei personal computer è quella che discende dal mondo materiale dell'ufficio. La sua forma è basata sull'utilizzo degli strumenti WIMP (Windows, Icons, Menus, Pointers), che sono al centro del suo funzionamento. Le finestre fungono da cornici e vetrine, le icone e i menù sintetizzano e veicolano l'informazione e le possibilità agentive. L'interfaccia WIMP classica soddisfa la propria fame di icone con metafore di carta e cartone: cartelle, documenti, pagine dall'angolo piegato per indicarne la salienza, post-it adesivi a indicare il genere e la destinazione del materiale cartaceo con i loro codici cromatici, loghi aziendali a differenziare diversi prodotti software e schede di lavoro, organizzate in finestre che possono essere spostate come oggetti su una scrivania. È il punto d'arrivo della ricerca della *working surface*, la verticalizzazione della scrivania e la sua fusione con lo schermo. Il suo protagonista è il puntatore, avatarizzazione della mano dell'utente che se ne fa incarnazione esplorativa. Si tratta di un *avatāra* vero e proprio, un'emanazione di un'entità distinta che ne enfatizza un determinato aspetto, esistendo su un piano differente pur senza alterare gli altri aspetti dell'entità originale³. Il puntatore può indicare qualcosa su un menù, afferrare, trascinare e lasciar andare, spostare una scheda di fronte a un'altra e riordinarle, scorrere un documento, premere bottoni ed evidenziare, grazie al supporto di miriadi di segnalatori che rispondono istantaneamente ai suoi tocchi essenziali. Esso comanda e il resto dell'interfaccia

³ Un eccellente esempio del concetto, di origine indiana, si trova nel settimo capitolo del *Devi Mahatmya*, in cui la dea Kali ha origine dalla fronte della dea della guerra Kaushiki, di cui rappresenta gli aspetti di ferocia e distruzione. Poiché l'avatarizzazione implica un'estensione e una specificazione ma non una re-istanziamento o un'alterazione, si è scelto il termine "emanazione", più adatto di "incarnazione", sia per il fenomeno proprio della mitologia indiana sia per il processo di estensione della mano nel mouse. L'episodio descritto in questa nota si trova nel *Devi Mahatmya*, B. Uday Nath (trad.), VII, 6.

risponde, come avveniva sin dai tempi del telescopio di Gascoigne. Sulla superficie numerata dello schermo esso produce le coordinate, e le produce secondo le dinamiche di una mano avatarizzata. Il puntatore manipola. Le espressioni “manipolazione di menù” e “manipolazione di oggetti 3D”, ampiamente presente nella letteratura della progettazione di interfacce, esprimono la chiralità del puntatore con la forza dell’ovvio, anche quando esso non è un mouse né alcun dispositivo di *input* ad uso manuale⁴.

L’interfaccia basata sull’iconografia della scrivania ha anche un carattere di ibridazione tra il tecnico e il privato, che lo schermo del personal computer incarna come luogo. È un ambiente costituito su misura di un singolo essere umano, per il quale ogni cosa è reperibile e consultabile in ogni dato momento, un piccolo *milieu* fortemente accentrato e ordinato da una sola autorità. Attraverso l’interfaccia grafica la cultura, “tutta la cultura”, non è solo estremamente, infinitamente riproducibile, impugnabile e manovrabile, nella forma più estrema di un processo di Benjaminiana memoria⁵, ma anche attivabile e disattivabile, appare e scompare al comando di un singolo dito, divenendo ancora più privata e insieme costantemente disponibile e, molto letteralmente, alla mano.

Nel pieno spirito di questa manipolabilità, l’interfaccia grafica classica è fortemente sezionabile nelle sue logiche formali. L’acronimo WIMP indica un insieme di strumenti complementari e modulari, ma anche dotati di una loro indipendenza funzionale, storica e visuale, che permette l’emergere di forme parziali di interfaccia grafica, o sottointerfacce, basate sull’utilizzo di alcuni di quegli attrezzi piuttosto che di altri. L’interfaccia di qualunque editor di testo o immagini è evidentemente una sottointerfaccia che rientra nel paradigma WIMP, ma non fa necessariamente uso di finestre. Le interfacce a menù (*menu-driven interfaces*) non sono altro che sottointerfacce formali dell’interfaccia grafica la cui finalità è quella

⁴ Per un eccellente esempio, si veda l’uso di termini relativi alla mano in P. Monteiro et al., *Hands-free Interaction in Immersive Virtual Reality: A Systematic Review*, in *IEEE Transactions on Visualization and Computer Graphics*, Vol. 27 (2021), pp. 2702-2713.

⁵ Ci si riferisce qui all’epocale saggio di W. Benjamin, *L’opera d’arte nell’epoca della sua riproducibilità tecnica*, Feltrinelli, Trebaseleghe (2021).

di segmentare l'agentività dell'utente e ridurre la necessità di esplorazione, in modo che il suo processo di lavoro tramite l'interfaccia sia mappato da una gerarchia precisa e possibilmente intuitiva, oltre ad annullare la necessità di un puntatore e dunque, potenzialmente, di un dispositivo dotato di un sistema di pilotaggio hardware per puntatori, come il mouse. Un altro esempio di finalità analoga è costituito dalle interfacce a modulo (*form-based interfaces*) o le interfacce a questionario, sottointerfacce formali sotto ogni punto di vista, preposte a dirigere ed efficientare la raccolta di informazioni da parte dell'utente che, a causa della precisa mappatura tra strumenti complementari dell'interfaccia grafica e forme dell'altrettanto complementari di agentività dell'utente, sezionano ciò che gli viene offerto per ottenere determinati comportamenti in precisi spazi e tempi. Il piano digitale delle interfacce è per definizione modulare e di facilissima trasformabilità e componibilità, infinitamente più del livello hardware precedentemente discusso, e questo è particolarmente valido nel caso dell'interfaccia grafica. Ciò è all'origine del suo successo, così come della forza e della vastità dei processi normalizzanti, filtranti e trasformativi che essa ha messo in atto nel mondo contemporaneo. È dunque naturale che, nuovamente, le diverse categorie di interfaccia formale che si relazionano all'interfaccia grafica secondo le loro differenze siano esplorate in relazione a tale norma, per evidenziarne gli aspetti dominanti e i limiti, mettendo così in luce quelli della norma stessa.

Interfacce grafiche tattili

Di uso ormai universale, le interfacce formali associate a dispositivi tattili sono strettamente legate alla logica metaforica, gerarchica e di massima disponibilità dell'interfaccia grafica del personal computer, con una superiore attenzione alla capacità di recidere tale disponibilità per salvaguardare il carattere privato dell'informazione che espongono. Fortemente influenzate (e in buona parte distribuite) dai grandi produttori di cellulari e sistemi operativi per dispositivi tascabili, come Alphabet/Google e Apple, sono caratterizzate da un montaggio più temporale e meno spaziale rispetto a quello dell'interfaccia grafica classica, poiché lo spazio schermico è una risorsa di cui i dispositivi tattili spesso dispongono in misura ben minore rispetto a un computer. La logica che domina tali dispositivi è infatti quella del contenimento spaziale di un elevatissimo numero di funzioni. A essi fanno eccezione terminali altamente specializzati, come lo sportello del bancomat e il menù tattile di ristoranti particolarmente tecnofili, la cui univocità funzionale fa però sì che l'iconizzazione di software che veicolano differenti attività non sia più necessaria. Tali interfacce si limitano quindi al solo uso di sottointerfacce a menù. L'interfaccia grafica tattile è più precisamente, come l'interfaccia grafica classica, un modo di fare interfaccia che permette la costituzione di innumerevoli architetture visuo-tattili tramite la sua modularità. Non si tratta però di un'interfaccia WIMP, in quanto totalmente priva di puntatore.

La presenza di una qualsivoglia forma di avatarizzazione della mano è evidentemente superflua alla presenza della mano come operatore diretto. Ciò avvicina però il piano digitale dell'interfaccia formale a quello più materiale dell'organo (o dello strumento impugnato dall'organo, che pur sempre *organon* rimane). La conseguenza di questo avvicinamento è che ogni icona, ogni pulsante, ogni voce di menù e componente dell'interfaccia deve dichiarare più eloquentemente il suo stato di "disponibilità per la mano", il suo prestarsi all'*affordance*, distinguendosi come ciò che si presta a un determinato gesto e ottenendo così uno speciale posto nella percezione e nella memoria di chi ne fa uso. James Gibson, padre del concetto di *affordance* come lo intendiamo oggi, forniva

una descrizione di questo processo nel paragrafo “Learning the affordances of objects” nel testo sui sensi come sistema percettivo:

When the constant properties of constant objects are perceived (the shape, size, color, texture, composition, motion, animation, and position relative to other objects), the observer can go on to detect their *affordances*. I have coined this word as a substitute for *values*, a term which carries an old burden of philosophical meaning. I mean simply what things furnish, for good or ill. What they *afford* the observer, after all, depends on their properties. The simplest affordances, as food, for example, or as a predatory enemy, may well be detected without learning by the young of some animals, but in general learning is all-important for this kind of perception. The child learns what things are manipulable and how they can be manipulated, what things are hurtful, what things are edible, what things can be put together with other things or put inside other things - and so on without limit. He also learns what objects can be used as the means to obtain a goal, or to make other desirable objects, or to make people do what he wants them to do.⁶

Di certo questo meccanismo non è appannaggio esclusivo delle interfacce tattili: ricondurre i singoli elementi dell'interfaccia grafica a oggetti noti è il significato stesso dell'espressione *What You See Is What You Get*, la parziale illusione di coincidenza delle *affordances* proposte dagli oggetti presenti sulla scrivania e quelle proposte sulla superficie schermica. Inoltre, il concetto di *affordance* è considerato di fondamentale importanza nella letteratura della progettazione di interfacce, dove è stimata come miglior forma tra i modi di comunicare all'utenza la finalità e l'uso di una determinata componente visuale o meccanica. Il termine è di uso così comune da essere considerato un vero e proprio strumento della disciplina, al punto in cui il significato ne viene spesso travisato e alterato, rendendo l'*affordance* una caratteristica universale degli elementi dell'interfaccia che il progettista deve saper imporre, invece di una relazione che si instaura

⁶ J. Gibson, *Senses as Perceptual Systems*, George Allen & Unwin, London (1966), p. 285.

contestualmente tra oggetto e soggetto, in particolare nel soggetto dotato di mani⁷. Nel regime WIMP, però, il parallelo tra elementi di interfacce e oggetti manipolabili è fondamentalmente metaforico e supportato dal linguaggio più che da un effettivo atto di manipolazione: nessuna mano, né organica né avatarizzata, apre una cartella digitale, il doppio click di un mouse su un'icona a forma di cartella si limita a far comparire una finestra. Nel contesto tattile la metafora si fa più forte, in particolare attraverso l'uso dell'animazione degli elementi grafici. Le schede si muovono guidate dal (e guidando il) moto delle dita, numerose icone a pulsante, a scorrimento e a interruttori simulano uno spostamento di materia coerente al gesto manuale, gli elementi grafici più voluminosi dell'interfaccia si fingono più tridimensionali attraverso l'uso di ombre e simulazioni prospettiche, e quelli più contenuti si animano muovendosi nello spazio schermico o cambiando dimensione in risposta agli stimoli tattili. Nel 2014, Google ha lanciato Material, il sistema che esprime la filosofia progettuale del Material Design⁸. I punti cardine del sistema sono un ampio uso di effetti di profondità, animazioni e un montaggio spaziale altamente dinamico. I fondamenti dell'interfaccia grafica tattile dei dispositivi Apple non sono distanti: famosi per la trasversale coerenza, mantenuta tramite l'imposizione di rigide regole di interfaccia per qualunque applicazione desideri apparire sul loro App Store, gli elementi grafici di iPhones e iPads sono sempre precisamente delineati nello spazio, i loro interruttori mobili vibrano e rispondono con leggerissimo ritardo al tocco, ed elementi funzionalmente simili tra loro (come la funzione di uscita o "indietro") si trovano sempre nelle stesse aree dello schermo, producendo un effetto di permanenza e coerenza spaziale che facilita la costituzione di *affordances* e dota di senso frazioni di spazio.

Come già discusso nella controparte hardware delle interfacce grafiche tattili, la totale centralità del sistema visuo-tattile in interfacce così trasversali può rendere

⁷ Questa particolare incomprendimento è enfaticamente denunciata da Bill Buxton, storico designer di Microsoft, in B. Buxton, *Sketching User Experiences: Getting the Design Right and the Right Design*, Morgan Kaufmann Pub, Burlington (2011).

⁸ La documentazione relativa a Material è disponibile sul sito ufficiale del sistema, presso <https://m3.material.io/>, ultima visita 20/11/24.

difficile la ricerca di soluzioni alternative, poiché non si dà nell'uso di simili dispositivi l'utilizzo delle mani senza lo sguardo e viceversa, se non nei rari casi in cui determinate parti dell'interfaccia siano state perfettamente memorizzate o quelli in cui, per un numero limitato di tocchi, altre parti del corpo possano sostituire le dita. L'adesione a regole di montaggio spaziale proprie di un regime tattile, in cui gli oggetti sono posizionati in relazione allo spazio più che l'uno in relazione all'altro, ostacolano l'adattamento dell'interfaccia in forme differenti. A meno di un montaggio spaziale acustico, che richiede uno spazio reale sufficientemente ampio o sistemi acustici molto costosi (come cuffie molto avanzate o impianti sonori distribuiti), la logica visuo-tattile non può essere mappata efficacemente nel reame del suono, e il suo protagonismo rallenta lo studio e l'esplorazione di interfacce in regimi differenti.

Il montaggio non è peraltro limitato allo spazio schermico: nella sua forma più forte, la produzione di interfacce tattili si spinge al limite dello scheuomorfismo (la riproduzione di immagini o meccanismi pertinenti a oggetti o tecnologie differenti da quella in uso) divenendo di fatto un montaggio di interfacce sullo stesso piano ontologico. L'informazione da comunicare e manipolare viene posta su "oggetti digitali" che simulano il funzionamento meccanico di interfacce non digitali, come rotelle e manopole virtuali. Il valore di simili scheuomorfismi è tale che alcuni rivenditori di dispositivi touch pubblicizzano i loro prodotti accentuando proprio il loro fortissimo scheuomorfismo. È questo il caso di reMarkable⁹, che fa della somiglianza alla carta il principale fattore di merito dei propri tablet¹⁰, la cui interfaccia cerca di riprodurre il comportamento materiale di un foglio sul piano visuale e acustico (tralasciando, per necessità, un aspetto stereotipicamente fondamentale nel dibattito tra carta e digitale, ossia quello olfattivo). Allo stesso modo, Material di Google contiene selettori d'orario digitali a forma di orologi analogici e selettori di elementi a carosello, da manipolare "direttamente" per causarne la rotazione. In particolare, le soluzioni come quest'ultima sono popolari

⁹ <https://remarkable.com/>, ultima visita 20/11/2024.

¹⁰ Questo è in particolare il caso della serie reMarkable Paper, presentati con lo slogan "The future of paper is here".

nella progettazione di interfacce la cui forma si rivolge alla mano, sia nel regime tattile sia in quello più prettamente gestuale della Realtà Virtuale o Aumentata: dal palco rotante di Dachselt¹¹ fino ai menù rotanti delle moderne piattaforme immersive, passando per i caroselli 3D generalizzati di Wang e colleghi¹², il fascino dei selettori a rotazione è testimone di come la metafora dell'inerzia e dell'attrito possano veicolare un senso di manualità di grande efficacia. Le interfacce grafiche tattili costituiscono un potentissimo esempio di invisibilizzazione dei processi della macchina: ogni istruzione, ogni elemento, ogni meccanismo si copre di una maschera da semplice risposta meccanica, intuitiva e prevedibile per minimizzare il carico cognitivo dell'utente, componendo un sistema di regole coerenti e facilmente verificabili costituite delle cause e degli effetti più semplici ed elementari possibili.

¹¹ Il sistema IMPLANATORIUM è presentato in R. Dachselt, *The Challenge to Build Flexible User Interface Components for Non-immersive 3D Environments*, in H.-J. Bullinger, J. Ziegler (eds.), *HCI (2)* (1999), pp. 1055-1059.

¹² S. Wang et al., *Designing a Generalized 3D Carousel View*, in *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, New York (2005), pp. 2017-2020.

Interfacce a linea di comando

All'estremo opposto troviamo interfacce di maggiore anzianità, le interfacce a linea di comando, che proprio alla loro anzianità devono la loro estetica e la loro riconoscibilità. Il loro caratteristico sfondo nero a scritte bianche precede gli schermi digitale, e si origina dagli schermi al fosforo, per i quali illuminare solo la porzione di schermo corrispondente alle lettere era insieme economicamente più efficiente e meno prone al rischio di bruciature.

Le interfacce a linea di comando sono nate per rivolgersi a un'utenza fortemente dedicata. La loro forma è estremamente criptica per i non iniziati. Dal punto di vista visuale, si tratta di interfacce essenziali quanto lo è lo spazio d'azione che offrono: esse forniscono lo spazio per inserire un comando e darlo in pasto alla macchina. I comandi possono essere volti all'esecuzione di singole azioni o di gruppi di azioni, e in questo secondo caso non è raro che l'interfaccia prenda il nome di *interfaccia batch*.

Il reale spazio d'azione è, in questo caso, nascosto nel linguaggio. Utilizziamo qui il termine "nascosto" non per esprimere un qualche carattere secretista o occulto dei linguaggi di programmazione *bash*, ma per evidenziare il fatto che simili possibilità d'azione offerte dal linguaggio debbano essere attivamente cercate dall'utente. L'interfaccia non le propone all'utente stesso, a meno di non saperli cercare tramite l'interfaccia stessa (facendo ad esempio uso delle variazioni locali di comandi d'aiuto, come *-help*). Anche tale comando, come tutti gli altri, dovrà però essere noto all'utente, per merito dell'interfaccia o di fonti esterne. Sono interfacce che non si offrono all'esplorazione casuale, poiché non è questo il loro scopo: sono strumenti ad uso di chi conosce il loro linguaggio, che fungono da vere e proprie superfici di traduzione da una precisa lingua di inflessibile grammatica e immutabile vocabolario alla lingua della macchina. L'evoluzione delle interfacce, e in particolare la genesi dell'interfaccia grafica, si può interpretare come un progressivo allontanamento dall'ergodicità dell'interfaccia a linea di comando, rendendo l'interfaccia digitale sempre più eloquente, arricchendola di ampi linguaggi di progressiva leggibilità atti a nascondere, fino alla totale sparizione, la

complessità della macchina. Il termine ergodicità è qui utilizzato nell'accezione intesa da Elspeth J. Aarseth in *Cybertext: Perspectives on Ergodic Literature*¹³, nel quale l'espressione "testo ergodico" (dal greco ἔργον, *ergon*, lavoro e ὁδός, *hodos*, percorso) indica qualunque tipo di testo che richieda un lavoro non triviale al lettore per essere interamente fruito. Se le obiezioni al concetto riportate dallo stesso Aarseth evidenziano come nessun testo sia realmente triviale, poiché nessun lettore è passivo nei confronti di un testo (si può facilmente affermare, con Gibson, che cognitivamente non lo è alcuna forma di percezione¹⁴), non esiteremo ad affermare che il caso delle interfacce a linea di comando si pone su un differente livello di ergodicità rispetto a qualunque interfaccia metaforica e intuitiva, poiché richiede al suo fruitore di compiere lavoro extratestuale. La conoscenza del linguaggio necessario a interfacciarsi con la macchina tramite la linea di comando proviene spesso infatti da studio esterno, autonomo o didattico, o da ricerca su altri testi di forma differente.

Era evidentemente questo il caso fin dall'origine delle interfacce a linea di comando: il computer mainframe IBM 7090, rilasciato nel 1959, era accompagnato da un manuale d'uso¹⁵ che riportava molto di ciò che, ad oggi, costituisce il programma d'esame di un corso universitario in Architettura degli Elaboratori¹⁶: informazioni relative all'architettura a blocchi della macchina e al suo funzionamento, alla conversione delle istruzioni impartibili in azioni della macchina e a tutti gli aspetti del funzionamento della stessa, la cui comprensione era necessaria all'utilizzo della sua interfaccia.

Le interfacce a linea di comando contemporanee hanno volontariamente mantenuto una loro ergodicità, e riscuotono ancora grande successo tra gli "iniziati", i

¹³ E. J. Aarseth, *Cybertext: Perspectives on Ergodic Literature*, The John Hopkins University Press, Baltimore (1997).

¹⁴ J. Gibson, *Senses as perceptual systems*, George Allen & Unwin, London (1966).

¹⁵ *Reference Manual, IBM 7090 Data Processing System*, in *International Business Machines Corporation* (Minor Revisions ed. 1962).

¹⁶ A titolo di esempio, si possono osservare numerosissimi paralleli tra i capitoli del manuale dell'IBM 7090 e gli argomenti trattati in S. Congiu, *Architettura degli elaboratori. Organizzazione dell'hardware e programmazione in linguaggio assembly*, Patron editore, Granarolo dell'Emilia, BO (2012).

professionisti, i cultori dell'informatica e tutti coloro che desiderino poter agire sulla propria macchina in maniera più diretta, a costo di doverne studiare i meccanismi e imparare i linguaggi, liberandosi di un'intercapedine metaforica il cui ridotto carico ergodico implica la riduzione della dimensione dello spazio agentivo offerto dal linguaggio macchina, per accedere al quale alcuni sono disposti a compiere tutto il lavoro extratestuale necessario.

Interfacce a linguaggio naturale

Distaccandoci ora dal paradigma visuo-tattile, dovremo confrontarci con il tipo di interfaccia di più tarda diffusione, quella basata sul linguaggio naturale. Nonostante il termine richiami pratiche estremamente recenti come il *Natural Language Processing* di grande rilievo nel settore dell'Intelligenza Artificiale, si può affermare che questo tipo di interfacce precedano l'era dell'esplosione dell'AI. Non si tratta però di un'affermazione semplice: l'aggettivo "naturale" è assai insidioso, su numerosi fronti. Nel discutere interfacce vocali, per "naturale" si intende in genere ciò che si contrappone a un linguaggio "imposto dall'uso della macchina". Anche volendosi cautamente mantenere su significati più elementari del termine, limitandosi al dizionario di Oxford, la voce *natural language* è definita come "a language that has developed in a natural way by people using it to communicate, rather than an invented language or computer code"¹⁷, che non risolve realmente il dubbio sul significato di "naturale", ed espone invece il termine "linguaggio inventato". Non si può trascurare però la parte centrale della definizione, quella che permette di discriminare linguaggi naturali da linguaggi più innaturali secondo un criterio teleologico: sarebbe naturale un linguaggio evoluto per comunicare tra esseri umani. Se questo non risolve realmente il problema, poiché non ci fornisce informazioni esaustive sulla flessibilità e l'ampiezza di tali linguaggi (è difficile sostenere, ad esempio, che il linguaggio naturale utilizzato nel commentare eventi mondani durante un pranzo di Natale presenti la stessa formularità del linguaggio utilizzato in un'aula di processo), ci dice quantomeno che tale linguaggio non può aver avuto origine dalla macchina stessa e che almeno due utenti ne abbiano dovuto far uso prima che esso divenisse parte di un sistema informatico.

Poiché in questo paragrafo stiamo distinguendo differenti interfacce secondo un criterio formale, sarà però opportuno adottare una definizione più circoscritta di questa: affermeremo dunque che un linguaggio naturale per l'interazione con la macchina, in questo campo, è un linguaggio che si costituisce di proposizioni

¹⁷ <https://www.oxfordlearnersdictionaries.com/definition/english/natural-language>, ultima visita 25/11/2024.

compiute di senso autonomo, dotate di soggetto (eventualmente implicito) e predicato non appartenenti a priori a uno specifico contesto, con un qualche margine non nullo di tolleranza nelle scelte lessicali permesse all'utente senza variare il risultato delle sue azioni. Questa definizione è fondamentalmente finalizzata a escludere dalla categoria delle interfacce a linguaggio naturale le forme basate sul preciso gesto vocale, molto più elementari e meno comuni, come ne sono state descritte nella sezione dedicata alle interfacce vocali.

Evidenzieremo qui che una simile definizione è lungi dall'essere perfetta e soprattutto dall'essere valida attraverso tempi differenti, né questo testo ha la pretesa di risolvere una questione complessa come l'eterno, resilientissimo dibattito tra ciò che è naturale e ciò che è culturale nell'ambito della tecnica, né quello sull'esistenza di una simile distinzione. Ci è non di meno di grande utilità, poiché ci permette di identificare la forma di interfacce qui descritta coerentemente con il discorso contemporaneo.

L'idea di conversare con la macchina è longeva quasi quanto il personal computer, e se ne ritrovano tracce pratiche fin dall'inizio degli anni '70, nella dissertazione dottorale di Terry Winograd¹⁸ (il cui nome è stato reso famoso dagli schemi di Winograd, espressioni linguistiche di eccezionale difficoltà di comprensione per una macchina priva di esperienza materiale del mondo¹⁹, un problema che sembra aver ricevuto soluzione solo nel 2019²⁰). La dissertazione presentava il sistema SHRDLU, una sequenza di lettere che non costituisce acronimo, in grado di descrivere e modificare uno spazio virtuale abitato da alcuni poliedri e forme tridimensionali, oltre a rispondere a domande poste dall'utente in inglese riguardo tale spazio. Il suo funzionamento è riproposto online da *blocks word*²¹. Da allora, la strada delle interfacce a linguaggio naturale ha camminato sulle gambe di

¹⁸ T. Winograd, *Procedures as a Representation for Data in a Computer Program for Understanding Natural Language*, The Massachusetts Institute of Technology (1971).

¹⁹ H. J. Lovesque et al., *The Winograd Schema Challenge*, in *Thirteen International Conference on Principles of Knowledge Representation and Reasoning proceedings*, AAAI (2012), pp. 552-561.

²⁰ V. Kocijan et al., *The defeat of the Winograd Schema Challenge*, in *Artificial Intelligence*, Vol. 325 (2023), online.

²¹ <https://patrickvanbergen.com/blocks-world/>, ultima visita 25/11/2024.

sporadici sistemi e bot, fino a convergere negli smartphone. Il primo caso di interfaccia a linguaggio naturale diffuso nelle tasche di una porzione significativa della popolazione è quello di Siri, diffusa da Apple nei telefoni iPhone all'interno del sistema operativo iOS 5²². L'esempio fu seguito da Microsoft con Cortana, la presentificazione vocale dell'assistente artificiale del protagonista della serie di videogiochi Halo, proprietà di Microsoft stessa, le cui funzioni non divergono da quelle di Siri. Si tratta però di interfacce adottate a sistemi già molto abbondanti in interfacce tattili e grafiche, il cui uso è principalmente circoscritto a situazioni di impossibilità d'uso manuale.

I più diffusi sistemi il cui principale utilizzo è un'interfaccia a linguaggio naturale, Amazon Alexa e Google Home, fanno ampio uso delle tecnologie di Intelligenza Artificiale. La principale implicazione del fondamento di queste interfacce sulle moderne tecnologie di IA è che, a differenza delle interfacce descritte finora in questo testo, un'interfaccia basata sull'uso del linguaggio naturale da parte di un IA non può essere interamente progettata, ma deve essere in parte addestrata secondo un processo le cui dimensioni e dinamiche non sono, ad oggi, del tutto prevedibili da un progettista umano²³. Il cosiddetto problema dell'*explainability*²⁴, ossia la ricerca di metodi in grado di esporre per intero a un osservatore umano i criteri e i meccanismi che regolano i comportamenti delle Intelligenze Artificiali contemporanee, è estremamente attuale e di grande interesse politico, economico e accademico. Il fronte dell'elaborazione del linguaggio naturale (testuale o scritto), ad oggi veicolato tramite i Trasformatori Generativi Pre-addestrati (Generative Pre-trained Transformers, GPT) come ChatGPT, che hanno visto la loro esplosione dopo la pubblicazione del già storico *Attention Is All You Need* di Vaswani e colleghi²⁵ nel 2017, sta ricevendo molta attenzione e, dopo i primi anni di ricerca,

²² La documentazione relativa, diffusa online nel 2012, è archiviata presso <https://web.archive.org/web/20120910202035/http://www.apple.com/ios/>, ultima visita 25/11/2024.

²³ R. Guidotti et al., *A Survey of Methods for Explaining Black Box Models*, in *ACM Computer Surveys*, Vol. 51 (2018), online.

²⁴ Ibid.

²⁵ A. Vaswani et al., *Attention Is All You Need*, in *31st Conference on Neural Information Processing System* (2017), pp. 6000-6010.

alcune soluzioni all'inspiegabilità dei meccanismi interni degli agenti conversativi. Un esempio concreto è rappresentato dal lavoro di Benfenati e colleghi²⁶, che trasformando lo spazio parametrico dei Trasformatori tramite la geometria di Riemann hanno generato un metodo per mappare delle precise classi di equivalenza nello spazio di *input* della rete, delimitando quali termini siano tra loro equivalenti in un dato contesto, come all'interno di uno stesso prompt. Un'eventuale conoscenza di simili classi di equivalenza, o almeno un'intuizione della loro esistenza e struttura, permetterebbe di certo all'utente non specializzato un uso più efficiente e consapevole delle interfacce a linguaggio naturale, dei loro limiti nella distinzione di accezioni e distinzioni terminologiche contestuali e dei meccanismi di incomprensione quando essi si verificano, migliorandone l'usabilità e aumentando il margine di controllo dell'individuo, potenziandone l'agentività.

²⁶ A. Benfenati et al., *Using Transformers Perception by Exploring Input Manifolds*, in processo di revisione, reperibile presso <https://arxiv.org/abs/2410.06019> il 26/11/2024.

Interfacce geospaziali

Le interfacce geospaziali costituiscono un tipo particolare di interfacce di grande presenza nell'ecosistema tecnologico contemporaneo. La categoria comprende tutte le interfacce progettate per richiedere, ottenere e manipolare informazioni secondo la loro relazione con uno spazio geografico, inclusa l'eventuale evoluzione di tali informazioni nel tempo, visualizzandole su una mappa. Simili interfacce appaiono su navigatori automobilistici, Google Maps e la famiglia di software che prende il nome di Sistemi Informativi Geografici (Geographic Information System, GIS). Si tratta dunque di interfacce che non si interpongono solo tra l'utenza e un dato quantitativo di informazione, ma sempre necessariamente tra l'utenza e uno spazio fisico con cui l'utenza può interfacciarsi in maniere differenti. La loro caratteristica di interfacce atte a filtrare l'informazione relativa a una porzione di realtà fisica per alterarne la percezione, potenziandola o filtrandola, le qualifica come *de facto* interfacce di Realtà Aumentata, nonostante simili interfacce siano imposte su una proiezione della realtà e non una visualizzazione della stessa. I casi in cui la categoria di Realtà Aumentata è stata usata per descrivere applicativi che sovrappongono informazione a una mappa, infatti, non sono pochi²⁷, anche se l'etichetta è in genere attribuita più facilmente a software che fanno coincidere la posizione geografica del dispositivo nello spazio fisico con la corrispondente posizione di un avatar nella rappresentazione digitale di quello spazio.

Simili interfacce possono trasformare il rapporto con lo spazio in molti modi. Le interfacce geospaziali personali che sfruttano la posizione di chi ne fa uso necessitano di apparecchiature adeguate a tracciare tale posizione, come sistemi GPS e magnetometri, e mappano lo spazio in grafi costituiti di destinazioni e percorsi che corrispondono a nodi e connettori, pesati secondo il criterio del tempo o del consumo di carburante.

²⁷ Si pensi, ad esempio, ai videogiochi basati sulla struttura di *Ingress* di Niantic, che fa ampio utilizzo di Google Maps, tra cui il celeberrimo *Pokémon Go* sviluppato successivamente dalla stessa casa di produzione.



Fig. 17: L'interfaccia di Pokémon Go, che riporta la posizione dell'utente sulla mappa tramite il suo avatar.

Cartwright e colleghi, nel 2001, hanno modellato il funzionamento delle tecnologie geospaziali come il rapporto tra mondo fisico, Sistemi Informativi Geografici e spazio personale, inteso come insieme delle percezioni, mediati da attuatori e sensori²⁸. Simili interfacce allungano la catena di traduzione più di altre categorie qui descritte: interfacciano il soggetto a un sistema di icone, attuatori e informazioni che lo interfacciano a loro volta alla mappa, che è a sua volta un'interfaccia prettamente visuale. La questione della mappa, della proiezione e della visione non umana dall'alto è di portata filosofica immensa (si pensi solamente al lavoro di Ernst Kapp²⁹ o di Giuliana Bruno³⁰), e non è questa la sede adatta ad affrontarla, ma dobbiamo evidenziare che simili interfacce sono pienamente investite dalle

²⁸ W. Cartwright et al., *Geospatial Information Visualization User Interface Issues*, in *Cartography and Geographic Information System*, Vol. 28 (2001), pp. 45-60.

²⁹ Ci riferiamo qui al concetto di proiezione elaborato nel primo testo di filosofia della tecnica, E. Kapp, *Elements of a Philosophy of Technology*, Kirkwood, J. W., Weatherby, L. (eds.), University of Minnesota Press, Minneapolis (2018), prima ed. 1877.

³⁰ Nella prolifica e importante produzione di Giuliana Bruno sul tema della superficie e della mappa, citeremo qui per rilevanza G. Bruno, *Atlante delle emozioni. In viaggio tra arte, architettura e cinema*, Mondadori, Milano (2006).

dinamiche proprie della tecnica della mediazione cartografica, trattandosi a tutti gli effetti di stratificazioni di interfacce al di sopra di una mappa.

La peculiarità delle interfacce geospaziali a tracciamento è che lo spazio che proiettano viene ripopolato da luoghi, attività e punti di interesse che ri-appaiono e consentono di presentarsi alla percezione dell'utente, rendendo di contro invisibili tutti i luoghi che non partecipano a quel livello di interfaccia, e che un simile processo è procedurale. Il loro linguaggio è caratterizzato dal protagonismo dell'icona, che funge da pulsante e portale di informazione. La loro logica è quella della compressione di informazione, poiché la quantità di informazione non è determinata, ma lo spazio che essa può occupare lo è pre-imposto.

Di contro, le interfacce geospaziali non traccianti hanno finalità meno generalizzate e non sono soggette a intrusioni informative dinamiche: riportano i dati che sono state predisposte a presentare e non permettono a fonti esterne di arricchire la proiezione. Sono interfacce che fanno ampio uso di codici cromatici per veicolare confronti tra aree delimitate, come il numero di vittime del SARS-CoV 2 in diverse nazioni o l'aumento o la diminuzione dei voti ricevuti da una determinata area politica in diverse zone di uno stato, e di elementi euclidei fondamentali, come punti e segmenti, per localizzare luoghi o delimitare spazi in un'interfaccia di struttura cartesiana. Le interfacce dei Sistemi Informativi Geografici professionali, in particolare, permettono anche di inferire (e poi visualizzare) diversi tipi di informazioni dalle immagini geografiche a partire da immagini raster, ottenute da fonti esterne come Google Earth Engine³¹. L'approccio incentrato sull'immagine raster, ossia sulla rappresentazione puntuale dell'immagine che assegna alle coordinate di ogni pixel il suo valore (contrapposta alla rappresentazione vettoriale), è fortemente sinergica a una rappresentazione cartesiana dell'informazione sulla superficie cartografica, e permette grande varietà di analisi e rappresentazioni tra loro con semplici strumenti matematici. Questa loro elasticità le rende comuni anche in altri tipi di analisi dell'immagine, come nel caso dello

³¹ J. A. Cardille et al., *Cloud-Based Remote Sensing with Google Earth Engine*, Springer (2023).

studio dei tessuti ossei³². Simili strumenti di analisi hanno un enorme impatto umano, poiché si prestano alla differenziazione socioeconomica di differenti aree sulla base dei dati estratti. Un banale esempio è quello della determinazione della pericolosità di una zona, che determina parametri come il costo delle polizze assicurative di chi vi abita³³.

Non si può sopravvalutare il fatto che processo di filtraggio dell'informazione operato dalle interfacce geospaziali sia per lo più operato su spazi fisici, e che dunque ogni traduzione della realtà materiale in dato e ogni aggregazione di dati, ogni criterio (inevitabilmente assiologico) di selezione di informazioni salienti ed esclusione di altri, ogni aspetto della rappresentazione degli stessi fino alla semiotica dei colori disposti su una mappa ha il potere di riconfigurare spazi e comunità.

Con l'emersione di nuove trasformazioni geopolitiche, non è azzardato credere che in un futuro non lontano la simbologia delle interfacce geospaziali, costituita ad oggi principalmente da codici cromatici (come le gradazioni di colore per la qualità dell'aria o la diffusione di precipitazioni o i percorsi evidenziati in blu nei software di navigazione) e indicatori puntuali dovranno necessitare di nuove forme visuali per arricchire il loro linguaggio. D'altra parte, come è già stato accennato per altri tipi di interfaccia in questo testo, si tratta di un tipo di interfaccia di difficilissima trasformazione nel reame del non visibile, il più resistente tra le categorie finora proposte, con le conseguenze normalizzanti, sociali e tecniche che questa resistenza comporta.

³² D. C. Rose et al., *Technical note: The Use of Geographical Information Systems Software for the Spatial Analysis of Bone Microstructure*, in *American Journal of Physical Anthropology*, Vol. 148 (2012), pp. 648-654.

³³ C. Garrad, *Geoprocessing with Python*, Manning, eBook (2016).

C. Interfacce teleologiche

L'ultimo livello ontologico che ci interessa analizzare è quello di più difficile localizzazione, e che presenta per questo la massima varietà. Si tratta di una famiglia di interfacce definite secondo la loro finalità, in particolare nell'influenzare direttamente il comportamento dell'utente, e che prescinde, almeno in parte, dal dispositivo materiale e dalla configurazione degli elementi digitali utilizzati nella sua realizzazione. Alcuni esempi di questa categoria, già citati nel descrivere interfacce d'altro tipo, sono le interfacce naturali, le interfacce di Realtà Virtuale o Aumentata e le interfacce a interazione tangibile. La discussione di simili interfacce impone la venuta al pettine di un nodo a lungo rimandato: la questione dell'interazione. Il termine è sostanzialmente inevitabile in qualunque studio, discussione o progettazione di interfacce digitali, a cui è profondamente connesso. È innegabile che la focalizzazione della progettazione sull'interazione e non "sull'interfaccia in sé", spesso erroneamente intesa come nient'altro che un insieme di oggetti digitali su uno schermo, rappresenti un punto di svolta nella storia della progettazione di tecnologie digitali¹. Tuttavia, ciò non significa che l'interazione non sia a sua volta stata travisata. L'autore, a riguardo, non ha molto da aggiungere a ciò che Lev Manovich sostiene nel già citato *The Language of New Media*: abbiamo già affermato che l'effetto che emerge dalla relazione tra soggetto e dispositivo, la cui stessa divisione netta potrebbe apparire problematica, altro non è che l'interfaccia stessa, e il concetto di interattività ne risulta ridondante. La relazione con un computer o uno smartphone presume l'agire dell'utente ed è generalmente procedurale e partecipativa, avvicinandosi già drasticamente con queste due caratteristiche alla definizione di interattività proposta da Janet Murray², una delle più intuitive e non strettamente vincolate alla digitalità. Si potrebbe inoltre osservare come l'interazione con simili dispositivi sia sempre finalizzata

¹ B. Moggridge, *Designing Interaction*, MIT Press, Cambridge MA (2007).

² J. H. Murray, *Inventing the Medium. Principles of Interaction Design as Cultural Practice*, MIT Press, Cambridge MA (2011). Le definizioni presenti nel libro sono state riportate in un glossario online di facile consultazione presso <https://inventingthemedium.com/glossary/>, ultima visita 27/11/2024.

all'attuazione di altre relazioni, come nel caso dello scambio di messaggi, di una telefonata o della creazione di un contenuto da far fruire a qualcun altro, di fatto interagendo con l'altro al di là del medium. La catena di interazioni inanellate che si va così costruendo corrisponde al concetto di catena di interfacce finora utilizzato in questo testo. La fondamentale differenza nel parlare di interazione invece che di interfaccia è nella diminuzione del rischio di concentrarsi su un oggetto digitale o fisico che viene investito del titolo di interfaccia a sé stante, riportando nel discorso sulla relazione la persona che fa uso di quello strumento.

Parlare di interazione³ significa utilizzare un linguaggio equivalente a quello sull'interfaccia, ma più resistente all'errore che consiste nel confondere la progettazione di interfaccia con la composizione di immagine o con la progettazione dell'architettura di un elaboratore. Non si può negare il valore pratico di questa resistenza. Ciò non di meno, abbiamo sostenuto che in uno studio dell'interfaccia in quanto tale essa debba essere definita come un effetto che emerge in presenza e all'agire dell'essere umano. Abbiamo dunque ritenuto più utile limitare l'uso del termine "interazione" per concentrarci sull'interfaccia. Discutere in termini di interfaccia ci permette anche di non scordare che le tecnologie qui discusse sono caratterizzate da un rapporto complesso con la trasparenza e che, se da una parete si presentano come strumenti per svolgere un compito "immediatamente" nel senso temporale, proprio nel presentarsi si frappongono con estrema forza nei rapporti che mediano e richiedono un ingaggio attivo. Ciò sarebbe stato assai più complesso da esprimere in un discorso sull'interazione. Ma poiché lo studio in essere ha anche carattere tassonomico, e poiché non ci arroghiamo il diritto di rinominare ciò che è generalmente discusso tramite un ampio uso del termine "interazione", ne faremo in questa sezione tutto l'uso necessario, a valle della presente disambiguazione metodologica.

Se davvero esiste una categoria di interfacce che si definiscono secondo le modalità interattive che i progettisti desiderano ottenere, questi desideri devono definirsi

³ Si noti che ci si riferisce qui sempre all'interazione con dispositivi digitali, e non si desidera scomodare il campo della medialità tutta e di ogni forma dell'interazione umana.

come differenti da una norma. In questo caso, più che nei precedenti, la norma dettata dai dispositivi e le interfacce più comuni è talmente trasformativa e modulare da aver perso una definizione fondamentale. Non si tratta solo di un preciso equilibrio visuo-audio-tattile che situa gli organi di senso e di moto in una gerarchia, ma anche del rinforzo di una norma posturale, di precisi rapporti prensili con la materia, di un modo preciso di intendere la luce, il suono e la presenza: un concetto, quest'ultimo, che ha incontrato un inaspettato interesse filosofico proprio dopo che la sua controparte, la "telepresenza", è stata coniata da Marvin Minsky proprio per comprendere i meccanismi generati dagli strumenti telematici di lavoro remoto⁴.

Le componenti del comportamento umano trasformate dalla tecnologia sono innumerevoli, soprattutto quando tale tecnologia è fortemente pervasiva. Per questo, le interfacce teleologiche tendono a concentrarsi solamente su determinati aspetti ed eventualmente indirizzarne altri come mezzi per raggiungere il loro obiettivo, lavorando su diversi livelli ontologici a profondità molto diverse tra loro e offrendosi a un numero altamente variabile di applicazioni. Nei casi più estremi, simili interfacce possono rappresentare veri e propri modi di pensare l'interfaccia stessa nel suo senso più ampio, alterando completamente il rapporto di strumentalità tra utente e macchina, riconfigurando i tempi, gli spazi e le modalità delle tecnologie digitali e spostando significativamente molto del carico semiotico verso una delle due parti. In casi meno estremi, si tratta di ampi metalinguaggi che non abbandonano la priorità grafica o gli strumenti hardware delle interfacce classiche, ma ne ristrutturano radicalmente gli elementi per trasformare gli strumenti digitali e le loro finalità.

La norma a partire da cui queste interfacce saranno catalogate è quella che riteniamo di poter attribuire ai computer ad uso personale da Engelbart in poi, una finalità allineata a quella di Gascoigne, di Fitts e di Bush: la proiezione dell'informazione in una forma manipolabile. Questa norma finalistica è centrale nella circoscrizione di innumerevoli interfacce al reame del visuo-tattile, ed è così efficacemente

⁴ M. Minsky, *Telepresence*.

espressa dal personal computer e dal telefono cellulare da essere divenuta sinonimo dell'uso di dispositivi digitali. Con la saggezza del senno di poi, potremmo anche affermare che la fusione tra i due dispositivi nello smartphone era probabilmente inevitabile, poiché le loro evoluzioni si muovevano in parallelo alla ricerca dello stesso obiettivo. La forza del rapporto tra proiezione, mano e strumento precede, peraltro, le tecnologie digitali, ed è alla base del discorso fondativo della filosofia della tecnica come disciplina. Ciò non di meno, l'ecosistema delle interfacce teleologiche non normate offre sia casi studio che sono stati capaci di svincolarsi dal regime proiettivo e manipolativo, sia casi studio che l'hanno abbracciato in maniera completamente differente, ristrutturando la relazione tra materia e informazione digitale.

Interfacce tangibili

L'interfaccia tangibile costituisce una strategia esplorativa del modo di interfacciarsi con gli strumenti digitali, fortemente connessa all'esistenza di un tipo preciso di interfaccia tattile. Nel 2006, Hornecker e Buur ne fornivano una comprensiva definizione:

Tangible User Interfaces (TUIs) and Tangible Interaction are terms increasingly gaining currency within HCI. This field of research relies on tangibility and full-body interaction and gives computational resources and data material form. Embedding computing in the everyday environment and supporting intuitive use, it shares goals with other novel approaches to HCI. Variations of this approach have been pursued over the last two decades as 'graspable user interfaces', 'tangible user interfaces', 'tangible interaction', or physical-digital interactions and digitally-augmented physical spaces.

While in traditional desktop computing the screen is merely a window through which we reach into a digital world, with tangible interfaces we act within and touch the interface itself. Designing tangible interfaces requires not only designing the digital but also the physical, and their interrelations within hybrid ensembles, as well as designing new types of interaction that can be characterized as full-body, haptic, and spatial - new challenges for design and HCI. As building upon users' experience of interacting with the real world lowers the threshold for activity, the embodiment of interaction objects alleviates the 'access bottleneck' of the keyboard, and interaction with these systems is easily observable, they lend themselves to the support of face-to-face social interaction.

This is reflected in a considerable number of systems aimed at cooperative scenarios.⁵

Dal punto di vista della realizzazione, i sistemi di interfaccia tangibile si concentrano sull'utilizzo di artefatti manipolabili la cui posizione e orientamento nello spazio è tracciata e utilizzata per interagire con l'informazione. Come suggerito da Hornecker e Buur, questo qualifica le interfacce tangibili come potenziale strumento di Realtà Aumentata, in grado di arricchire di senso e

⁵ E. Hornecker, J. Buur, *Getting a Grip on Tangible Interaction: A Framework on Physical Space and Social Interaction*, in *CHI 2006: Designing for Tangible Interaction proceedings*, Montréal (2006), pp. 437-446.

informazione oggetti materiali in tempo reale tramite la simultanea evoluzione di gemelli digitali corrispondenti.

Nel 2000, Ullmer e Ishill⁶ hanno pubblicato un insieme di caratteristiche fondamentali delle interfacce tangibili e delle relazioni tra il fisico e il digitale che queste generano, caratteristiche destinate a rimanere sostanzialmente invariate fino ad oggi:

- 1) Physical representations (*rep-p*) are computationally coupled to underlying digital information (*model*).

The central characteristic of tangible interfaces lies in the coupling of physical representations to underlying digital information and computational models. The Urp example illustrates a range of such couplings, including the binding of graphical geometries (data) to the building objects, computational simulations (operations) to the wind tool, and property modifiers (attributes) to the material wand.

- 2) Physical representations embody mechanisms for interactive control (*control*).

The physical representations of TUIs also function as interactive physical controls. The physical movement and rotation of these artifacts, their insertion or attachment to each other, and other manipulations of these physical representations serve as tangible interfaces' primary means for control.

- 3) Physical representations are perceptually coupled to actively mediated digital representations (*rep-d*).

Tangible interfaces rely upon a balance between physical and digital representations. While embodied physical elements play a central, defining role in the representation and control of TUIs, digital representations – especially, graphics and audio – often present much of the dynamic information processed by the underlying computational system. Where the above three characteristics refer directly to our MCRpd model, a fourth TUI characteristic is also Significant.

- 4) The physical state of interface artifacts partially embodies the digital state of the system.

⁶ B. Ullmer, H. Ishill, *Emerging Frameworks for Tangible Interfaces*, in *IBM System Journals* (February 2000) pp. 915-931.

[...] the MCRpd model does not specify whether a TUI's physical representations are composed of one or many physical artifacts. In practice, tangible interfaces are generally built from *systems* of physical artifacts. Taken together, these collections of objects have several important properties. As physical elements, TUI artifacts are *persistent* – they cannot be spontaneously called into or banished from existence. Where a GUI window can be destroyed or duplicated at the touch of a button, the same is not true of Urp's building objects. Also, TUI artifacts frequently may be “read” both by people and computers by their *physical state*, with their physical configurations tightly coupled to the digital state of the systems they represent.

Tutto ciò che gli autori adducono a titolo semplificativo nella spiegazione delle caratteristiche da loro esposte è riferito a software di Urban Planning (URP), in cui è possibile manipolare rappresentazioni digitali di oggetti architettonici in maniera non dissimile da quella descritta da Engelbart nel '62. Nel sistema di interazione tangibile, tuttavia, a permettere il controllo della rappresentazione digitale di un'architettura è un modellino fisico corrispondente, posato su un tavolo, che può essere ruotato e manipolato.

La necessità di coincidenza rappresentativa tra l'*affordance* offerta da un oggetto fisico e quella offerta dalla sua controparte digitale raggiunge il suo apice nelle interfacce tangibili, eppure queste non sono necessariamente veicolate da una superficie tattile o da un'interfaccia grafica tattile: la manipolazione di un artefatto fisico e la risposta del dispositivo possono essere attuate tramite un'interfaccia cinetica, come ad esempio una telecamera, sensori di movimento e un casco Oculus Rift. È questo il caso di Inner Garden, un'interfaccia sperimentale proposta da Sol Roo e colleghi⁷. Nel sistema ideato dai ricercatori francesi, chi indossa il casco di Realtà Virtuale manipola una vera e propria *sandbox*, una porzione di sabbia all'interno di una scatola, per creare un mondo in miniatura che viene visualizzato in tempo reale all'interno del casco e proiettato sulla sabbia stessa.

⁷ J. Sol Roo et al., *Inner Garden: Connecting Inner States to a Mixed Reality Sandbox for Mindfulness*, in *CHI '17 International Conference of Human-Computer Interaction Proceedings* (2017), pp. 1459-1470.

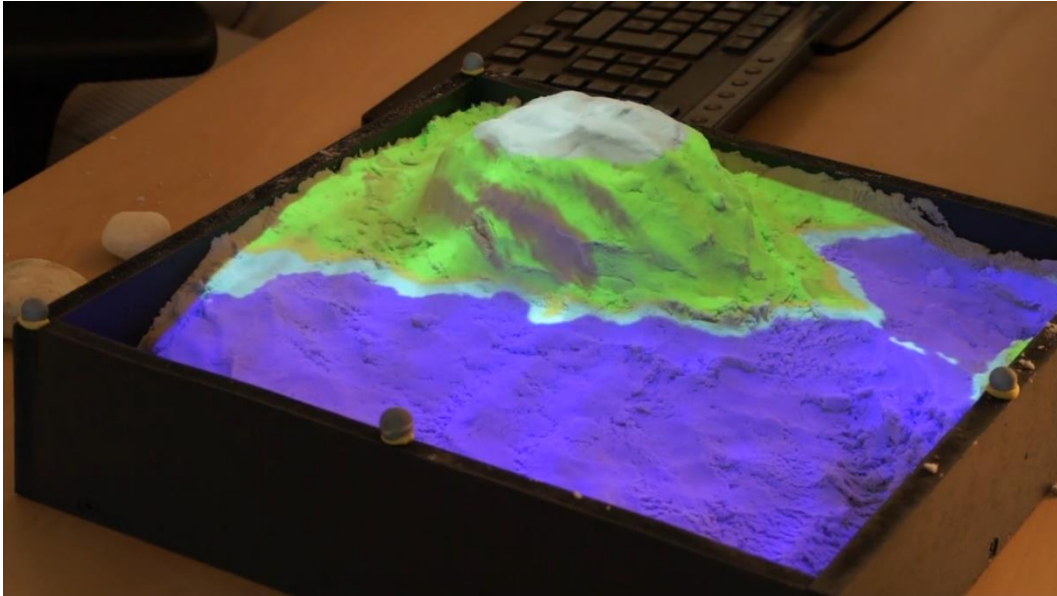


Fig. 18: Inner Garden. La generazione procedurale della proiezione dipende sia dalla modellazione della sabbia sia da segnali elettroencefalici. Immagine estratta dal video di presentazione del progetto, Vimeo.

La segmentazione della superficie osservata dalla telecamera permette di riconoscere pattern all'interno della sabbia, come diverse variazioni di altezza e forma, e generare corrispondentemente diversi elementi geologici e morfologici digitali. Un altro caso simile, accomunato dall'utilizzo di dispositivi indossabili, è quello di *The Hologram in My Hand* di Bach e colleghi⁸, che tramite un paio di occhiali Microsoft HoloLens di prima generazione proiettava visualizzazioni tridimensionali di dati nella forma di spazi cartesiani a tre assi su un apposito cartoncino impugnato dall'utente, che potevano essere ruotati e spostati liberamente per osservare le relazioni tra i dati da più punti di vista, aumentando efficacemente da due a tre le dimensioni (e dunque le variabili) naturalmente visualizzabili tramite gesti della mano pienamente intuitivi.

⁸ B. Bach et al., *The Hologram in My Hand: How Effective is Interactive Exploration of 3D Visualizations in Immersive Tangible Augmented Reality?*, in *IEEE Transactions on Visualization and Computer Graphics* (2018), pp. 457-467.

Interfacce multimodali

La categoria delle interfacce multimodali è forse quella che maggiormente si definisce nell'evidenziare la natura normata del discorso sulle interfacce. Il concetto di multimodalità non è certo esclusivo allo studio dei sistemi digitali e dell'interazione, ma riguarda ogni settore dell'agire e del comunicare umano:

As a phenomenon of communication, multimodality defines the combination of different semiotic resources, or modes, in texts and communicative events, such as still and moving image, speech, writing, layout, gesture and/or proxemics.⁹

Non ci può sfuggire che la quasi totalità dei sistemi di interfaccia descritti in questo testo presentano carattere di multimodalità, sia nell'*input* che nell'*output*. Gli esseri umani ingaggiano il mondo in maniera multimodale, a partire dai loro sistemi di percezione¹⁰. Nonostante ciò, all'interno del presente capitolo abbiamo identificato alcuni sistemi come specificamente multimodali. Se nella definizione appena riportata, infatti, la distinzione tra modalità coincide con la distinzione tra diversi *media*, una simile separazione non è del tutto intuitiva parlando delle modalità dell'azione umana. Va innanzitutto premesso, a scampo di potenziali incomprensioni, che non esiste una reale e univoca distinzione tra queste modalità. Il problema riguarda le convenzioni adottate nel relazionare modalità come simili o come categoricamente distinte, e quando modalità differenti possano offrire una sinergia olistica tra di loro.

Il discorso in essere riguarda, infatti, lo studio di interfacce in cui la multimodalità produce tipi di agentività che non sono riconducibili alla sola somma delle modalità coinvolte, ma in cui esse si coordinano e completano generando spazi d'azione nuovi. Se così non fosse, sarebbe possibile argomentare che un mouse è un'interfaccia multimodale: si utilizza un movimento continuo del braccio per

⁹ E. Adami, *Multimodality*, in O. Garcia et al. (eds.), *The Oxford Handbook of Language and Society*, Oxford University Press, eBook (2016), pp. 451-472.

¹⁰ J. Gibson, *Senses as Perceptual Systems*.

spostare la posizione del puntatore e uno o più movimenti impulsivi delle dita per segnalare azioni differenti. Ma affermare ciò sarebbe come affermare che ogni uso della mano deve essere separato dall'uso del braccio, come se i due non fossero in alcun modo correlati. Una simile distinzione è totalmente infruttuosa e fuorviante. Analogamente lo sarebbe accorpare modalità che media precedenti ci hanno insegnato ad accoppiare, come nel caso della combinazione audio-visiva del cinema e della televisione, che si ripropone in innumerevoli forme delle interfacce digitali. Potremmo infatti affermare, seguendo l'esempio di Gibson, che non sono quelle modalità attentive che storicamente abbiamo denominato "i cinque sensi", eventualmente integrati dalle diverse forme della propriocezione e dell'interocezione, a distinguere le diverse modalità: l'atto di toccare qualcosa avviene sempre in attesa del suono che accompagna il tocco, e spesso dell'osservazione della propria mano nel momento del tocco che prepara i percettori tattili all'incontro con la materia. Un saltuario suono monotonale che accompagna l'animazione di una notifica non è più multimodale della nostra consueta esperienza del mondo, che è già di per sé intrinsecamente multimodale. Nessuna interfaccia è esente dalla natura multimodale della nostra percezione e del nostro agire.

Le interfacce multimodali rappresentano, piuttosto, quei sistemi che permettono di coordinare diverse modalità di azione diretta nel comunicare alla macchina, coerentemente alla definizione di Adami. Storicamente, il primo esempio significativo di una simile interfaccia è già stato discusso nella categoria delle interfacce cinetiche: il sistema Put That There di Bolt, che comunicava alla macchina indicando con il braccio una posizione e informandola del comando tramite il linguaggio. In questo sistema, nessuna delle due parti del comando (il gesto direzionale e il comando vocale) è completa o significativa di per sé. La sola indicazione non è sufficiente per generare senso compiuto, né lo è il comando vocale. Il sistema Put That There è certo un esempio peculiare, poiché era progettato proprio per generare una coordinazione sinergica tra due modalità pienamente indipendenti tramite l'uso del dimostrativo "that", che doveva ereditare il suo senso da una specificazione differente.

Le interfacce multimodali si definiscono fruttuosamente in due modi: in contrapposizione alle interfacce unimodali e in contrapposizione alle interfacce che combinano modalità la cui combinazione è data per consolidata nelle tecniche del quotidiano. Se la prima definizione incarna un problema storico delle scienze tecniche e psicologiche, ossia la contrapposizione tra lo studio puntuale, in laboratorio, di porzioni isolate di fenomeni e la caoticità e inseparabilità di tali parti nell'esperienza quotidiana del mondo, la seconda le dignifica come esplorazioni al di fuori di una norma. Già nel 2003 Xiao e colleghi osservavano come le interfacce multimodali fossero spesso preferite da persone con corpi non normati e abilità differenti¹¹, conclusione di cui il sistema ACAT di Stephen Hawking rappresenta un estremo esempio.

Proprio in quanto definite come al di fuori di una norma, simili interfacce non si prestano a una normalizzazione loro, neppure a livello di dettami progettuali. Un encomiabile tentativo in questo senso è stato operato da Reeves e colleghi¹², il cui lavoro si focalizza su “la massimizzazione delle abilità fisiche e cognitive umane”, un'espressione che si presta a diverse interpretazioni problematiche. Se la maggior parte delle norme progettuali sono mirate all'evitare l'insorgere di errori e di sovraccarico cognitivo e di sforzo di memoria di lavoro, la proposta finisce inevitabilmente per delineare sistemi normalizzati (“Design for broadest range of users and contexts of use”). Il precetto, apparentemente intuitivo, del progettare interfacce per la maggior quantità di utenti possibili è stato discusso da Harper in *Is There Design-For-All?*¹³, che al carattere normante dei tentativi di inclusione ad ampio spettro ha contrapposto l'approccio “Design-For-One”, che concentra lo sforzo progettuale solo su corpi e categorie ben definiti, partendo da specifiche abilità di volta in volta.

¹¹ B. Xiao et al., *Modeling Multimodal Integration Patterns and Performance in Seniors: Toward Adaptive Processing of Individual Differences*, in *ACM International Conference on Multimodal Interfaces Proceedings*, Vancouver (2003), pp. 265–272.

¹² L.M. Reeves et al., *Guidelines for Multimodal User Interface Design*, in *Commun. ACM*, Vol. 47 (2004), pp. 57-59.

¹³ S. Harper, *Is There Design-For-All?*, in *Universal Access in the Information Society* (2007), pp. 111-113.

La multimodalità rappresenta verosimilmente un indirizzo esplorativo che non ha risoluzione ultima, e in questo sta la forza delle interfacce multimodali: sono oggetti di ricerca più di quanto non siano semplicemente strumenti tecnici, le cui forme e definizioni incoraggiano la sperimentazione di modi di agire e relazionarsi allo spazio e all'informazione che sfidino le forme normate delle interfacce digitali. Come evidenziato in una review sull'interazione multimodale di Turk¹⁴, se le forme di interazione considerate unimodali sono genericamente ben definite a priori, la ricerca delle loro combinazioni funzionali nel proporre interfacce robuste e comunicative può avvenire solo sul piano empirico tramite sperimentazione sul campo.

¹⁴ M. Turk, *Multimodal Interaction: A Review*, in *Pattern Recognition Letters*, Vol. 36 (2014), pp. 189-195.

Interfacce naturali

La categoria delle interfacce naturali, o interfacce basate sull'interazione naturale, si configura come la più estrema dal punto di vista del carico macchinico nel processo di traduzione dell'intenzione in operazione, e così facendo a rendere invisibile se stessa, parte integrata dell'ambiente nella maniera in cui Galloway descrive il processo di fuoriuscita del medium dalla medialità. Concetto di esplosiva popolarità negli anni 2000, quello dell'interazione naturale emerge dalla percezione di uno sbilanciamento tra la complessità delle macchine ad uso quotidiano e i compiti che gli utenti desiderano svolgere con esse, di gran lunga più semplici¹⁵.

Natural interaction is defined in terms of experience: people naturally communicate through gestures, expressions, movements, and discover the world by looking around and manipulating physical stuff; the key assumption here is that they should be allowed to interact with technology as they are used to interact with the real world in everyday life, as evolution and education taught them to do.¹⁶

Come nel caso delle interfacce multimodali, si tratta di una presa di posizione nei confronti della tecnologia e di determinate scelte progettuali, più che di una definizione di carattere tecnico o pratico. E come nel caso delle interfacce a linguaggio naturale, il termine "naturale", con la sua problematicità, assume un protagonismo già discusso nella sezione dedicata alle interfacce formali. Ad analoga questione proporremo qui analoga soluzione contestuale: la distinzione che proponiamo è netta ma lungi dal portare con sé la pretesa di risolvere questioni sulla definizione di "naturale" o di qualunque categoria si desideri contrapporgli.

In questa categorizzazione, per interfacce a interazione naturale intenderemo tutte quelle interfacce che fanno uso di gesti, forme, formule e linguaggi che precedono o eccedono il contesto dell'interazione uomo-macchina. Si tratta dunque di interfacce che si fanno carico della conoscenza del valore semiotico, meccanico o

¹⁵ A. Valli, *The Design of Natural Interaction in Multimedia Tools and Applications*, Vol. 38, 3 (2008), pp. 295-305.

¹⁶ Ibid.

biologico di diversi comportamenti umani, con la finalità di permettere all'utente di non dover apprendere alcuna nuova modalità di azione per confrontarsi con la macchina, ma interfacciando l'utente alla macchina come farebbe con una situazione o un soggetto non macchinico. La profondità a cui simili interfacce possono spingersi è molto variabile: come nel caso delle interfacce multimodali, diverse interfacce descritte nelle sezioni precedenti sottostanno alla finalità delle interfacce naturali almeno in parte. Un esempio banale è quello costituito dalle interfacce a linguaggio naturale, con accento sul carattere orale di tali interfacce. Se le interfacce a linguaggio naturale possono proporsi in forma scritta, perché un'interfaccia possa definirsi totalmente a interazione naturale anche la tastiera e il monitor dovranno decadere, come nel caso dei robot umanoidi proposti da Bärmann e colleghi¹⁷ nell'Ottobre del 2024. Se ad oggi il linguaggio è una forma estremamente popolare di interazione naturale grazie alla diffusione dei GPT, in passato la teoria dell'interfaccia dell'interazione naturale ha dedicato molta attenzione anche al gesto tattile. La rappresentazione dell'inerzia negli elementi grafici delle interfacce tattili è in tutto e per tutto un tentativo di naturalizzazione dell'interazione. Nel 2008, l'anno in cui Valli pubblicava la definizione di interazione naturale qui discussa, lui e il suo team sviluppavano *SensitiveTable*, uno schermo orizzontale integrato in un tavolo che offriva la possibilità di agire tattilmente su un'interfaccia grafica fortemente materiale allo scopo di rendere più intuitiva l'interazione con gli elementi dell'interfaccia¹⁸.

Il tema dell'interazione naturale tattile non si è esaurito con i dispositivi a schermo capacitivo: nel 2019, Kyriakou e Hermon¹⁹ hanno proposto un sistema di Realtà Aumentata e interazione naturale per risolvere il problema tattile in uno dei settori in cui esso è più annoso, quello museale.

¹⁷ L. Bärmann et al., *Incremental Learning of Humanoid Robot Behavior from Natural Interaction and Large Language Models*, in *Frontiers in Robotics and AI, Robot Learning and Evolution*, Vol. 11 (2024), online.

¹⁸ A. Valli et al., *Natural Interaction SensitiveTable*, in *CHI '08 Extended Abstracts on Human Factors in Computing Systems* (2008), pp. 2315-2318.

¹⁹ P. Kyriakou, S. Hermon, *Can I touch this? Using Natural Interaction in a Museum Augmented Reality System*, in *Digital Applications in Archaeology and Cultural Heritage*, Vol. 12 (2019), online.



Fig. 19: SensitiveTable. Immagine estratta dal video della conferenza di presentazione del progetto.

Per aggirare l'antico dettame del “guardare ma non toccare”, il loro sistema propone una forma di interazione realistica con alcuni artefatti archeologici presso il Mediterranean Science Festival di Cipro.

La frontiera più interessante dell'interazione naturale, però, è con ogni probabilità quella connessa alla computazione affettiva: la capacità macchinica di riconoscere lo stato d'animo di un individuo (non necessariamente utente, nel caso della computazione ubiqua) e comportarsi di conseguenza. Il termine fu diffuso nel 2000 da Rosalind Picard²⁰, e individua una serie di tecniche e algoritmi atti a riconoscere le singole emozioni o lo stato affettivo espressi da un individuo tramite caratteri riconoscibili²¹. Il primo dei due fondamentali approcci possibili è quello che considera l'esistenza di un insieme discreto e finito di emozioni distinguibili, la cui storia parte da Descartes²² per arrivare agli storici esperimenti di Paul Ekman e

²⁰ R. W. Picard, *Affective Computing*, The MIT Press, Cambridge MA (2000).

²¹ R. W. Picard et al., *Toward Machine Emotional Intelligence: Analysis of Affective Physiological State*, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23 (2001), pp. 1175-1191.

²² R. Descartes, *Le passioni dell'anima*, Bompiani, Milano (2003).

Wallace Friesen²³ alla ricerca dei principali marcatori delle differenti emozioni umane e della loro indipendenza dalla cultura di appartenenza. Il secondo, invece, è quello che considera lo stato affettivo come appartenente a uno spettro continuo, genericamente definito come lo spazio bidimensionale valenza-eccitazione (valence-arousal), basato sul modello di Russell e Pratt²⁴. Il modello a emozioni discrete, popolarizzato dalla serie televisiva *Lie to Me* (il cui protagonista è stato scritto ispirandosi proprio allo stesso Paul Ekman), si è qualificato come il più diffuso a causa di un fondamentale vantaggio: l'esistenza delle reti neurali classificatrici. Con simili interfacce è possibile immaginare, ad esempio, sistemi in grado di riconoscere lo stato affettivo dell'abitante di una casa e, tramite un sistema domotico basato sull'Internet of Things, alterare l'ambiente domestico in modo dinamico. Un'analisi di quest'uso, in teoria fortemente "naturale", della cosiddetta *Affective Internet of Things* è stato operato nel 2020 da Kaklauskas e colleghi²⁵. Si è spinti dunque a chiedersi che cosa significhi, a livello progettuale e tecnico, creare dispositivi concepiti e usati per non essere dispositivi e, specialmente nell'ultimo caso in particolare, non essere nemmeno intenzionalmente usati. Coerentemente alle definizioni da noi esposte all'inizio del presente testo, un'interfaccia è tale in virtù della sua emersione: la sua opera di traduzione non è silenziosa e, in accordo con Galloway, essa può dirsi tale solo fintanto che rimane almeno parzialmente inoperabile. La totale invisibilizzazione del medium in quanto tale non coincide però con la sua sparizione: un dispositivo che agisce da sé, senza bisogno né possibilità di essere interpellato dall'utente, non rappresenta il tipo di interfaccia di cui abbiamo discusso finora. Il problema è però ben più profondo e complesso di una semplice sparizione dei dispositivi tecnologici o delle loro necessità di segnali performativi e specifici per essere operati: la computazione

²³ P. Ekman, W. V. Friesen, *Constants Across Cultures in the Face and Emotion*, in *Journal of Personality and Social Psychology*, Vol. 17(1971), pp. 124–129.

²⁴ J. A. Russell, G. Pratt, *A description of the Affective Quality Attributed to Environments*, in *Journal of Personality and Social Psychology*, 38 (1980) pp. 311–322.

²⁵ A. Kaklauskas et al., *Affective Internet of Things*, in F. Pacheco-Torgal et al. (eds.), *Woodhead Publishing Series in Civil and Structural Engineering*, Woodhead Publishing, online (2020), pp. 203-233.

ubiqua impone all'utente uno stato di costante osservazione che offre il fianco al Tocco di Mida e alla trasformazione performativa di tutti i gesti operati durante tale osservazione. Non crediamo, d'altra parte, che possa darsi un dispositivo tecnologico genuinamente neutrale e "naturale" nella sua interazione, fintanto che la sua destinazione è quella di essere interfaccia. Se è evidentemente possibile registrare il comportamento di un essere umano tramite una telecamera nascosta e ottenerne dati, inferirne informazioni e agire con la conoscenza acquisita, un effetto di interfaccia – o, se vogliamo, un'interazione – tra l'utente e il dispositivo impone che determinati gesti, linguaggi e comportamenti vengano trasformati nella forma, nella frequenza e nella collocazione spaziale dall'emergere di un dispositivo che, inevitabilmente, verrà attivato secondo il desiderio dell'individuo. È difficile argomentare che i telefoni a schermo tattile non abbiano alterato il nostro rapporto con il tocco, ed è altrettanto difficile credere che, in una *smart home* in grado di cambiare la luminosità o la musica in base ai nostri gesti spontanei, il nostro comportamento non verrebbe alterato per evitare gli errori dovuti all'applicazione di un approccio statistico, inevitabili anche nella migliore delle realizzazioni. Di come l'approccio statistico sia alla base della configurazione di ogni normalizzazione e degli effetti di simili processi abbiamo già discusso in altre sezioni del presente capitolo. Nel contesto delle interfacce di interazione naturale, è doveroso aggiungere che il settore progettuale non è certo ignaro del fenomeno né privo di letteratura orientata a indirizzare il problema della normalizzazione e sfidare il concetto di "naturale"²⁶.

²⁶ R.-D. Vatavu, *From Natural to Non-Natural Interaction: Embracing Interaction Design Beyond the Accepted Convention of Natural*, in *ICMI '23 25th International Conference on Multimodal Interaction Proceedings* (2023), pp. 684-688.

Interfacce simulate

In ultimo, un'ampissima categoria teleologica è quella che include tutte le interfacce narrative, immersive e, in generale, finalizzate a simulare un contesto reale o fittizio a scopo di intrattenimento, addestramento o di generare telepresenza. Appartengono a questa categoria le interfacce videoludiche e di Realtà Virtuale, molte interfacce di Realtà Aumentata, i simulatori di veicoli (come i simulatori volo per l'addestramento dei piloti) e tutti gli ambienti di simulazione operativa con cui fruire formazione professionale. Si tratta dunque di tutte quelle interfacce che non si limitano a simulare alcuni comportamenti di dati elementi grafici, acustici o tattili secondo metafore intuitive, ma a simulare intere situazioni circoscritte in spazi virtuali (anche ricorsivamente, come simulando altri computer o interfacce dentro interfacce), leggi della fisica fittizie e relazioni causali di validità circoscritta alla simulazione. Non di rado, al di sopra della situazione simulata viene imposta un'ulteriore interfaccia grafica, testuale o acustica, come liste di operazioni da svolgere, segnalatori audio o differenti tracce musicali a indicare il verificarsi di determinati eventi e barre, contatori e indicatori a fungere da promemoria dello stato della simulazione o dell'utente all'interno della stessa. Simili interfacce possono abbracciare l'uso di diverse componenti WIMP o esserne completamente privi, simulare spazi finiti o generarli proceduralmente²⁷, proporsi su schermo o adottare tecnologie indossabili per coinvolgere maggiormente l'utente nella simulazione. La fondamentale caratteristica che accomuna questo tipo di interfacce è la simulazione di un punto di vista circoscritto. Nei casi più comuni, esso è generato tramite una telecamera virtuale che la macchina utilizza, insieme a leggi fisiche e ottiche più o meno complesse, per situare l'utente all'interno spazio virtuale. Le interfacce pilotate da questa logica sono profondamente cinematografiche, ed ereditano dal cinema numerose strategie comunicative e linguaggi. L'utente può ricevere in

²⁷ L'ampio spettro di algoritmi di generazione procedurale di contenuto (PCG) precedenti l'avvento dell'AI è stato analizzato in M. Hendrikx et al., *Procedural Content Generation for Games: a Survey*, in *ACM Transactions on Multimedia Computing, Communications, and Applications* (TOMM), Vol. 9 (2013), online.

dotazione un avatar, possibilmente modificabile secondo²⁸ le proprie preferenze, e venir coinvolto nella simulazione in prima persona (in soggettiva), in terza persona (la telecamera segue l'avatar) o, più frequentemente nei casi in cui l'utente non abbia alcun singolo avatar, tramite diverse forme della visuale non umana come la telecamera dall'alto o la visuale isometrica. Nei videogiochi come nelle simulazioni, si distinguono innumerevoli generi dotati dei loro linguaggi, delle loro spazialità simulate e delle loro convenzioni semiotiche, sia nelle componenti audiovisive che nell'uso della telecamera, finalizzati a generare un certo senso di immersione.

Il tema dell'immersione è centrale alle interfacce simulative, e in particolare quelle che adottano la prima e la terza persona. Mel Slater e Sylvia Wilbur nel 1997²⁹ hanno formulato una definizione di immersione basata su parametri osservabili per poter misurare la portata del fenomeno in funzione delle tecnologie utilizzate. La protagonista del loro studio per capacità di generare immersione è la Realtà Aumentata, la cui componente cinetica genera una forte illusione di presenza. Altri elementi concorrono a generare quella che lo stesso Slater ha definito un'illusione percettuale ma non cognitiva³⁰: la stereofonia dell'audio, sensazioni aptiche, un preciso tracciamento del corpo dell'utente e un ampio angolo visuale. Lo studio del fenomeno della presenza dal punto di vista psicologico è, come evidenziato da Slater, tornato di grande popolarità nel 2012³¹ grazie al rilascio di Oculus Rift, e si è configurato nel campo fortemente interdisciplinare dei *presence studies*. I temi della presenza dell'immagine e dell'immersione nella stessa, tuttavia, trascendono la popolarità dei caschi di Realtà Virtuale. L'ingresso nell'immagine, nello specchio e nel mondo virtuale si può rintracciare, secondo Pinotti, fino al mito di

²⁸ L. Wiesing, *Pause of Participation. On the Function of Artificial Presence*, in *Research in Phenomenology*, online (2011).

²⁹ M. Slater, S. Wilbur, *A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments*, in *Presence: Teleoperators and Virtual Environments*, Vol. 6 (1997), pp. 603–616.

³⁰ M. Slater, *Immersion and the Illusion of Presence in Virtual Reality*, in *British Journal of Psychology*, Vol. 109 (2018), pp. 431–433.

³¹ Ibid.

Narciso³², passando per *Attraverso lo specchio e quel che Alice vi trovò* di Lewis Carroll. Le interfacce simulative più immersive costituiscono, per Pinotti, un tipo di immagine completamente nuovo: le an-icone, immagini che rifiutano il loro statuto iconico di sezioni di realtà separate dal mondo circostante tramite una cornice (sia essa nella forma di una cornice quadrata, di un piedistallo, di uno schermo) e proponendosi come presenziali, immediate e scorniciate. Le interfacce di Realtà Virtuale, con la loro pretesa di circondare interamente l'utente tramite l'uso di caschi e di simulazioni tridimensionali, coprendo l'intero campo visivo con l'inquadratura e offrendosi come realtà collocate di fronte all'utente senza alcuna mediazione, lo accolgono come lo specchio che una giovane Alice Liddell avrebbe sognato di attraversare.

Nel discutere le interfacce simulative e la loro finalità, l'esempio di Alice è quantomai calzante: l'immersione, lo sprofondare nello specchio e l'immergersi sono metafore molto adeguate a descrivere una relazione fortemente spaziale, che presenta e rappresenta spazi e agentività autonomi a cui accedere attraverso un'interfaccia che ritorna a una viscosità che richiama la sua definizione chimica originale. Faticosamente attraversata quella soglia, le precedenti regole sono sospese e nuove leggi e nuove agentività vengono imposte e devono essere apprese dall'utente. In accordo con queste regole, la razza come differenziatore di capacità diviene un fatto immutabile³³, la salute diviene una risorsa quantificabile e dal valore leggibile, l'apprendimento di nuove abilità e tecniche applicabili al mondo reale diviene non supervisionato e operato tramite una lunga reiterazione di prove ed errori, la gravità un fenomeno individuale da cui svincolarsi a piacere, il suono un fenomeno puramente locale e istantaneo.

Ancora una volta, l'interfaccia può però rimanere tale solo fintanto che si mantiene parzialmente inoperabile, e l'immersione completa coinciderebbe con la sparizione dell'interfaccia in quanto tale. Come Narciso che, nella versione Ovidiana del mito, giunge a riconoscere la sua immagine come tale grazie alle lacrime che turbano

³² A. Pinotti, *Alla soglia dell'immagine. Da Narciso alla realtà virtuale*, Einaudi, Torino (2021).

³³ Ci riferiamo qui alle diverse capacità di razze fantastiche presenti in molti giochi di ruolo. Il tema è stato analizzato anche da Galloway nel già citato *The Interface Effect*.

l'acqua³⁴, inevitabilmente le interfacce simulate devono presentarsi tramite effetti emergenti sul confine tra ciò che dividono. Nelle interfacce simulate, tale rottura dell'illusione si situa nella presentazione gerarchica di informazione saliente atta a direzionare l'agire dell'utente. Fin dalla prima versione del gioco *Doom* nel 1993, il primo caso di gioco sviluppato tramite un software apposito (un *game engine*), la presenza di elementi mimetici non diegetici³⁵ si è mantenuta un'eredità costante delle interfacce simulate. Si tratta di un particolare sottolivello delle interfacce simulate costituito da componenti o individui esteticamente (e a volte diegeticamente) coerenti con il mondo simulato, ma che si rivolgono in maniera più o meno diretta all'utente. Il concetto, formulato da Barbara per classificare alcuni tipi di narrazione vocale all'interno del medium videoludico (come, ad esempio, l'assistente robotica del centro di comando Terran nei giochi della serie *StarCraft*) si può estendere a tutte le scelte progettuali orientate a mantenere il protagonismo dell'informazione saliente pur ricercando un'interfaccia di facile attraversabilità. La soluzione più comune alla soddisfazione delle due necessità (l'immersione e la comunicazione) è quella adottata da Fitts per i caccia dell'aviazione americana: l'HUD. La proiezione, o sovrapposizione, di semplici elementi informativi sullo schermo attraverso cui osservare l'ambiente è difficilmente equivocabile: la presenza di radar, barre e indicatori negli angoli dello schermo di un computer è fattore sufficiente a ipotizzare di essere di fronte a un'interfaccia simulativa. Proprio per questo, tuttavia, l'utente diviene rapidamente pilota, e una simulazione può perdere la sua capacità di immergere l'utente in ciò che simula e divenire solo un luogo d'azione direzionata da elementi grafici. Un eccellente esempio di recupero di questo peculiare equilibrio si può osservare nell'interfaccia dei giochi della serie *Dead Space*³⁶, la cui appartenenza al genere horror impone la necessità di un forte senso di coinvolgimento.

³⁴ A. Pinotti, *ibid.*

³⁵ J. Barbara, M. Haahr, *The Role of Voice in Virtual Reality Interactive Narratives*, in *The Journal of Interactive Narrative*, Vol. 1 (2024), online.

³⁶ *Dead Space*, Visceral Games (2008), xBox 360



Fig. 20: L'interfaccia HUD di StarCraft 2, Blizzard Entertainment. In basso a destra, nell'HUD, è presente il volto animato dell'assistente vocale robotica, che si rivolge direttamente all'utente.

Tra le molte soluzioni adottate dai progettisti della serie, spicca la sottointerfaccia preposta ad informare il giocatore dello stato del protagonista, e in particolare della sua salute (quante volte può essere colpito prima di un *game over*) e delle risorse a sua disposizione (le munizioni e l'energia rimasta alla sua tuta spaziale). Ritenendo che un canonico sistema di barre e indicatori agli angoli dello schermo spostasse l'equilibrio dell'interfaccia verso un fronte più funzionale, gli sviluppatori hanno immerso anche quegli elementi oltre lo strato superficiale dell'interfaccia, rendendo tali indicatori parti dell'equipaggiamento del protagonista. Ponendo le barre di salute ed energia sulla schiena del personaggio, l'utente può sempre osservarli dalla telecamera in terza persona. Per osservare la mappa o il proprio numero di munizioni, invece, dovrà ordinare al personaggio di controllarli da sé, e osservare quei valori quando il protagonista li fa apparire olograficamente.



Fig. 21: L'interfaccia di Dead Space. Il tubo segmentato sulla colonna vertebrale indica lo stato di salute del protagonista, la mezzaluna (qui vuota) sulla sua schiena indica il livello di energia della tuta, e la proiezione olografica di mirino e contatore di munizioni, pienamente diegetica, è osservabile dall'utente e dal suo avatar simultaneamente.

Quando invece la simulazione vuole rafforzare al massimo la propria illusione, tecniche più prettamente cinematografiche possono prendere il posto delle componenti grafiche dell'interfaccia, ma allontanandosi dallo statuto di interfacce digitali e riducendo, a volte fino ad annullare, l'agentività dell'utente, che diviene solo l'operatore di una camera virtuale in grado di muoversi in un ambiente simulato. Ciò però difficilmente può avvenire su ogni livello, e incontra particolare resistenza sul piano materiale: il peso di un dispositivo indossabile che grava sul capo, la forma di un CAVE, i colori e i comportamenti della luce artificiale proiettata sugli schermi, le inevitabili imprecisioni del calcolo elettronico non smettono di costituire categorici indizi della cornice spazio-temporale ed agentiva all'interno del quale le immagini an-icone devono confinarsi per tentare di occultare il loro statuto di simulazione.

Le interfacce simulative sono, come è stato accennato, dotate di linguaggi fortemente standardizzati e loro propri, seppur non privi di eccezioni. L'utilizzo di comandi continui per l'esplorazione, l'utilizzo di luce e codici cromatici per direzionare lo sguardo dell'utente, l'anticipazione acustica di eventi visuali (in una forma che coincide, di fatto, con quella della notifica) e la presenza di una

telecamera virtuale fissa o mobile impongono le priorità sensorie più normalizzate su un pubblico che raramente include abilità fortemente differenti da quelle di un corpo normato, almeno sul piano visuo-tattile. L'interfaccia simulativa è una di quelle composte dal maggior numero di sottointerfacce sovrapposte: un preciso insieme di dispositivi, un preciso sistema di comandi e di requisiti di coordinazione oculo-manuale, un rodato e complesso linguaggio acustico, visivo e tattile, una precisa semiotica degli eventi programmati, complessi sistemi ludici o simulativi di regole e obiettivi. Si tratta nuovamente di interfacce di carattere ergodico, seppur in misura minore rispetto alle interfacce a linea di comando, nella maggior parte delle loro implementazioni. Tuttavia, proprio l'efficacia dei loro linguaggi nell'ingaggiare e spingere all'azione l'utente è causa del loro trasversale successo: le cosiddette pratiche di *gamification* tramite software altro non sono che applicazioni della semiotica delle interfacce simulative per immergere gli individui in una data attività e spingerli verso dati obiettivi, più o meno autodeterminati, secondo la logica agonistica del superamento efficiente dell'ostacolo.

La diffusività della Realtà Aumentata

Prima di definire concluso il lavoro di classificazione, è necessario affrontare una questione presente a più riprese nelle sezioni precedenti: la presenza diffusa della Realtà Aumentata attraverso diversi livelli e categorie di interfacce. Intendiamo qui il termine “diffusa” in entrambe le sue principali accezioni: si tratta di una tecnologia diffusa nel significato di “molto presente nello spazio e nel tempo”, ma anche nel senso di “diluita uniformemente nel suo ambiente”.

Il termine, ad oggi ormai noto e di uso comune, è utilizzato per riferirsi a dispositivi e tecniche digitali molto differenti. Ciò che appare osservando le diverse applicazioni dell’espressione “Realtà Aumentata” è un’eccezionale capacità di emersione entro e attraverso differenti categorie di interfaccia ben definite: in questo stesso testo non si è esitato a definire tecnologie di Realtà Aumentata numerosi casi studio nella categoria delle interfacce geospaziali, delle interfacce tangibili, delle interfacce naturali e delle interfacce simulative. A questo variegato campione è doveroso aggiungere le tecnologie di Realtà Aumentata indossabili e alcune tecnologie domestiche. Le prime, spesso note come “smart glasses” (o, più volgarmente, “occhiali AR”), consistono in occhiali dotati di una piccola telecamera per la cattura di gesti e di immagini e un sistema di proiezione in grado di visualizzare dati ed elementi di interfaccia, un microfono e un dispositivo *output* sonoro. La proiezione avviene in genere sulle lenti degli occhiali, di fronte agli occhi dell’utente, come nel caso dei dispositivi HoloLens e HoloLens 2 di Microsoft, la cui superficie prende il nome di “lente olografica trasparente”³⁷. Simili dispositivi sono molto vicini a un immaginario futuristico, proposto nei decenni al grande pubblico dalla fantascienza in molti media differenti³⁸, e propongono un tipo di immagine e di uso della tecnologia lontano dalle interfacce classiche. Si tratta di

³⁷ Microsoft HoloLens 2 tech sheet, https://news.microsoft.com/wp-content/uploads/prod/2019/02/Fact-Sheet_HoloLens2.pdf, ultima visita 13/03/2025.

³⁸ Oltre ai già citati *Organo Genocida* di Satoshi Itoh e *Dead Space* di Visceral Games, si pensi ad un esempio molto più popolare come i fumetti di Batman di D.C. Comics, in cui l’uomo pipistrello fa frequentemente uso di lenti di Realtà Aumentata per analizzare la scena di un crimine, o al potente visore di Geordie La Forge (LeVar Burton) nella serie televisiva *Star Trek Next Generation* (1987).

immagini ambientali, an-icone vere e proprie, che scelgono la strategia dell'emersione per rendersi presenti all'osservatore³⁹, proponendosi cioè come immagini emerge dalla loro cornice e penetrate nello stesso ambiente di chi osserva, spazialmente e agentivamente. Negli smart glasses olografici, il sistema si appoggia ai dati acquisiti tramite telecamera per generare dinamicamente immagini digitali la cui posizione e il cui comportamento dipendono dallo spazio fisico. Si tratta in tutto e per tutto un processo di tecno-analisi, in cui il dispositivo si propone come soggetto a sé che si relaziona all'utente nel suo stesso spazio e attraverso il suo contenuto, proiettando l'immagine e la sua agentività al di fuori del reame del virtuale e insieme conservando la propria virtualità. Evidentemente, una simile tecnologia appartiene al reame delle interfacce teleologiche, poiché non è possibile situarla sul piano hardware o della pura forma ma solo identificarla secondo gli effetti che essa è progettata per generare – quelli, appunto, di un'interfaccia in grado di agire su e con l'ambiente fisico che ospita utente e dispositivo. Alcune di queste tecnologie (come gli smart glasses olografici appena descritti) si appoggiano inoltre a interfacce multimodali, la cui progettazione è finalizzata a fornire forme differenti e complementari di comunicazione tra utente e dispositivo, comandabili tramite gesto manuale, comandi vocali, rotazione del capo e dello sguardo e pulsanti posizionati sulla superficie degli occhiali. Si tratta, in particolare, di sistemi in grado di trasformare, tramite potenziamento o tramite filtraggio, l'attività cognitiva dell'utente: questa prestazione delle tecnologie di Realtà Aumentata è stata identificata come l'aumento dell'immaginazione da Dalmasso e Pirandello⁴⁰. Le due autrici identificano la Realtà Aumentata come eccellente caso studio delle teorie di Ihde e Malafouris:

In order for AR tools to work when we put into focus a section of space with our mobile phone or smart glasses, and to respond exactly as we want, the software we employ

³⁹ E. Modena et al., 360°. *L'immagine ambientale nelle arti visive tra realtà virtuale e aumentata*, in *PIANO B. Ari e Culture Visive*, Vol. 6, 1 (2021), online.

⁴⁰ A.C. Dalmasso, S. Pirandello, *Augmented Imagination. Thinking Technology Beyond Extension*, in *I Castelli di Yale Online. Annali di Filosofia*, Vol. 12, n.2 (2024), pp.29-42.

requires an extenuating training process. This training is not only for the machine that learns from our movements, but also for the consumers who have to learn how to use it. There is a reciprocal power play between the two poles, in which it is by no means obvious that the user is in control: as Don Ihde and Lambros Malafouris would say, «we make things which in turn make us». On the one hand, AR technologies bring about the fusion of the individual with the other in a broad sense, namely environments, objects, other beings, implementing the layers of reality, in space and time, on the basis of which the mind is constituted and nourished, and thanks to which it transforms itself; on the other hand, they read reality through algorithmic systems, thus conveying a fragmented and regimented understanding of it in a limited set of instructions.⁴¹

Se le tecnologie AR fungono quindi da elemento unificante tra l'individuo e l'altro in senso ampio, stratificando il rapporto con la realtà in senso tanto semio-pratico quanto visuale (interponendosi fisicamente tra l'osservatore e la realtà osservata), il loro effetto è quello di fungere da interfaccia nel senso più ampio possibile: la loro intenzione è di codificare la realtà secondo gli stessi meccanismi con cui, secondo Manovich, l'interfaccia digitale codifica la cultura.

L'effetto è però in questo caso più debole, a causa della resistenza della materia e della natura delle tecnologie digitali e di telecomunicazione: un paio di occhiali AR non può interfacciare un osservatore a ogni cosa. L'esistenza di protocolli, la necessità di progettazione aprioristica delle operazioni della macchina (vincolata dalla sua effettiva costruzione e programmazione) e il limitato campo d'azione a essa imposto dal limite dei suoi linguaggi e codici impone al dispositivo forti limiti settoriali. Così, l'occhiale e il telefono potranno visualizzare ciò che già conoscono, che hanno conosciuto tramite un altro dispositivo in grado di comunicare con loro o tramite l'acquisizione pilotata di dati dall'esterno, come ad esempio visualizzando un codice QR. In alternativa, potranno interfacciare l'utente con un modello o una frazione di realtà predeterminata, come nel caso degli applicativi di navigazione pratici o ludici.

⁴¹ Ibid., p.37.

La teleologia perseguita delle interfacce sfruttate da queste forme di Realtà Aumentata è ampia quanto ben delineata: una forma molto potente di quell'aumento dell'umano come sistema agente nel mondo, quello stesso aumento che Doug Engelbart auspicava di ottenere immaginando per l'aeronautica il personal computer nel 1962. Così definite, appaiono forse come una sorta di fatalista punto di arrivo della tecnica, una presenza permanente in grado di riconfigurare il mondo secondo logiche nuove. Questo non è l'argomento di Dalmasso e Pirandello, ma è certo come le grandi compagnie produttrici di smart glasses cercano di presentare se stesse: fin dai primi trailer di Google Glass, pubblicati nel 2012⁴², gli smart glasses per il grande pubblico sono stati presentati come dispositivi di uso pervasivo in grado di incanalare svariate tecniche del quotidiano come messaggistica, logistica, fotografia, acquisizione di dati e notizie, e insieme la generazione di interazioni e momenti eccezionali (si pensi, ad esempio, alla serenata al tramonto del trailer di Google Glass).

Ben altro tipo di interfaccia teleologica è quello che emerge non appena si volessero indossare dispositivi di altro genere, come nel caso degli occhiali industriali HMT-1 di RealWear⁴³. Questi non presentano lenti ma, seppur dotati di telecamera e microfono come HoloLens, presentano le loro immagini tramite piccolo monitor disposto al di sotto del livello degli occhi. A differenza di HoloLens, gli occhiali HMT-1 sono prodotti e pubblicizzati come strumenti professionali. La loro specialità non è la generazione di immagini fortemente connesse allo spazio abitato dall'utente, ma la comunicazione visuale all'interno tra utenti in forma gerarchizzata. Si tratta di dispositivi progettati per una fondamentale modalità d'uso: permettere a chi indossa gli occhiali di contattare qualcun altro trasmettendo una condivisione parziale del proprio campo visivo, secondo la logica di una "una

⁴² Diversi tutorial didattici sono disponibili sul canale Youtube Google Glass, presso <https://www.youtube.com/@GoogleGlass>, ultima visita 13/03/2025.

⁴³ RealWear HMT-1 tech sheet, <https://support.realwear.com/knowledge/hmt-1-specifications-for-model-t1100g>, ultima visita 13/03/2025.

condivisione dello schermo” sul campo⁴⁴. Tramite questi strumenti, un operatore remoto può fornire istruzioni vocali e visuali tramite manipolazione dell’immagine all’operatore sul campo, ad esempio operando zoom e disegnando linee, frecce e indicatori sul suo monitor indossabile, indirizzandone l’operato. Non si tratta dunque più di un ingaggio materiale con lo strumento che viene integrato nella cognitivtà dell’indossatore degli occhiali, ma di tecnologie in grado di trasferire una delega corporea a un altro soggetto situato altrove, la sostituzione della propria presenza con la telepresenza di qualcun altro.

Tramite la trasmissione di dati che permettano di ricostruzione con buona approssimazione il sensorio dell’agente presente in loco, l’agente remoto può progettare la sua azione secondo le proprie competenze e direttive. L’AR non funge più da agente trasformativo dello spazio locale, ma da trasmittente di soggetto attivo in un dato spazio. Non solo una simile Realtà Aumentata non è più veicolo delle logiche trasformative dell’agentività di chi indossa il dispositivo, ma serve in forma più diretta l’utente che non sta indossando l’occhiale, fornendogli la possibilità di imporre manualmente al lavoratore sul campo effetti simili a quelli del puntatore di Clavis Shepherd ne *L’Organo Genocida* discussi nel capitolo II⁴⁵: aiutante e direttore, giudice e giuria delle sentenze di cui il corpo indossante il dispositivo si fa esecutore. A ciò si aggiunge la funzione di strumento di telepresenza, incorporata però in un altro soggetto.

In ultimo, ad arricchire il ventaglio di tecnologie esemplari della versatilità del modello della Realtà Aumentata, vi sono tecnologie AR domestiche. In una pubblicazione del 2020, Jain e colleghi presentano un sistema protesico per individui privi o deboli di udito⁴⁶, che portiamo qui a titolo esemplificativo. Il sistema, che prende il nome di HomeSound, offre all’utente un’interfaccia vibrotattile su un orologio da polso, connessa a un sistema di microfoni distribuiti in

⁴⁴ Durante il percorso di dottorato l’autore ha avuto occasione di collaborare con un’azienda italiana che si occupa dello sviluppo di software per occhiali industriali, e di esplorarne le differenti funzionalità.

⁴⁵ Cap II, fig. 4.

⁴⁶ D. Jain et al., *HomeSound: An Iterative Field Deployment of an In-Home Sounds Awareness System for Deaf or Hard of Hearing Users*, in *CHI 2020*, Honolulu (Aprile 2020), online.

un'abitazione. Quando un microfono capta un rumore riconoscibile – come il suono di un lavandino rimasto incautamente aperto – una vibrazione segnala all'utente di controllare l'orologio che riporta la posizione e la fonte del suono.



Fig. 22: L'orologio del sistema HomeSound. Immagine tratta dalla pubblicazione.

Nel suo porsi a interfaccia tra l'utente e la realtà fisica per supportare delle decisioni informate nello spazio materiale, il sistema si colloca facilmente sotto l'ampio ombrello della Realtà Aumentata, ma è strutturalmente e teleologicamente differente dalle categorie precedentemente citate: una Realtà Aumentata che non proietta nello spazio materiale le logiche “nuove” che il digitale può veicolare, ma che funge da protesi inversa applicata all'ambiente, permettendo alla casa stessa di trasdurre informazione esistente in una forma più accessibile ai suoi abitanti.

Questa forte varietà di finalità, insieme all'ampiezza dell'insieme di tecnologie che compongono la grande e dinamica famiglia della Realtà Aumentata, permette di osservare come essa si costituisca anacliticamente anche nei confronti delle interfacce stesse. A differenza di una tecnologia che le viene spesso accostata, la

Realtà Virtuale, la cui ampia gamma d'usi si appoggia sulle molte prestazioni e possibilità della singola categoria della simulazione, la Realtà Aumentata appare in grado di realizzarsi al di sopra di un singolo meccanismo, a volte tramite indossabilità, a volte tramite una parziale vena simulativa, a volte solamente tramite uno specifico equilibrio nella coesistenza di elementi fisici e digitali nella logica dei suoi dispositivi – in generale, in una compresenza di digitale e materiale. Quest'idea che non le è certo esclusiva, e la sua forza sta nella sua capacità di differenziarsi dalle altre tecnologie digitali sul livello quantitativo (la misura di questa compresenza) più che su quello qualitativo: potenziare le relazioni che si generano da quella coesistenza, aumentandone l'intensità e moltiplicandone le direzioni. Se davvero il suo scopo è l'aumento delle prestazioni umane nel risolvere problemi e prendere decisioni (affermazione, questa, estremamente semplicistica e possibile solo al di fuori delle logiche politiche, commerciali e sociali che inevitabilmente si impongono sull'uso di una tecnologia pervasiva), allora la Realtà Aumentata altro non è che un modo di intendere ogni interfaccia digitale situata al confine con la realtà fisica, e non sottostà ad alcuna logica o norma corporea, tecnica o culturale, poiché può attingere indiscriminatamente a tecnologie drasticamente differenti tra loro su ognuno di questi piani.

Lo studio delle sue prestazioni è lungo, notevolmente complesso e ancora agli albori, com'è naturale per una tecnologia che, più che un'interfaccia come l'abbiamo sempre intesa, si configura come un nuovo modo di intendere l'interfaccia in senso ampio, e la cui genealogia e archeologia sembrano percorrere lo stesso sentiero dell'interfaccia classica – dalla finestra di Alberti in poi – per giungere a una destinazione differente: quella di una trasversalità totale, una capacità di penetrazione tecnica capace di esaltare la grande versatilità applicativa che è alla base del suo successo e della sua diffusione.

Conclusione

Il presente lavoro è il risultato di molte scommesse. Scommesse umane, scommesse accademiche, scommesse concettuali.

In questo frangente, il termine “risultato” assume un doppio valore. Questo lavoro è di certo ciò che è risultato dalla commistione di tali scommesse, ma è anche la tangibile prova di come tali scommesse siano terminate.

La prima, in senso cronologico, è di certo stata il brusco trasferimento disciplinare dell'autore da scienze molto dure e pratiche a scienze molto umanistiche e teoriche. Martin Heidegger era dell'idea che tra i due settori del sapere, che lui chiamava rispettivamente “scienza” e “pensiero”, vi fosse un abisso, e che l'unico passaggio possibile dalla prima al secondo fosse il salto¹. Sosteneva inoltre che il salto non conducesse solo dall'altra parte, ma in un mondo completamente diverso e riconfigurato. L'esperienza del dottorato che ha condotto a questa tesi ha di certo avuto il carattere del salto sull'abisso, anche se non ci è possibile essere certi che si tratti dello stesso salto o dello stesso abisso a cui si riferiva Heidegger. Il passaggio dall'ICT alla filosofia e all'archeologia dei media foucaultiana significa riconfrontarsi con il concetto di sapere, con nuove logiche epistemiche, e con la scoperta di espressioni come “logiche epistemiche”.

Significa però anche conoscere nuovi modi di concepire ed esplorare le discipline dure, poiché è impossibile costruire qualcosa senza un terreno d'appoggio. Di certo, questa è stata una delle più importanti scommesse alla base di questo dottorato: se l'intersezione di campi e discipline così lontane tra loro le avrebbe condotte allo scontro e a una disordinata commistione, o a una qualche seppur modesta forma di arricchimento di fascino Hegeliano. Ancora una volta, non sappiamo dire con certezza se il metodo ingegneristico sia stato sublato dalle scienze umane – una speranza forse presuntuosa – ma non di meno è l'applicazione di metodologie miste che ci ha condotto alla conclusione di questo lavoro.

¹ M. Heidegger, *Che cosa significa pensare?*, G. Vattimo (a cura di), Sugarco, Carnago (1971).

Il lavoro ha, nel primo capitolo “I. Definire l’interfaccia”, preso in analisi differenti definizioni dello strumento dell’interfaccia come superficie materiale, come fenomeno di compenetrazione trasformativa, come dispositivo facente funzione di codice semiotico e come realtà attraversabile. Attraverso diversi contributi alla costruzione del concetto, è stato possibile delineare quali aspetti abbracciare e quali lasciare a dibattiti e analisi differenti da quelli qui proposti. Abbiamo inquadrato l’interfaccia in un regime tecnico di superfici stratificate dal forte potere di trasduzione del senso, e abbiamo confrontato questo quadro teorico con quello di un regime tecnico in cui la superficie e l’interfaccia non sono sufficienti a descrivere la galassia di simili fenomeni di trasduzione.

Attraverso il secondo capitolo è stato possibile condurre un percorso genealogico dell’interfaccia grafica contemporanea, un percorso che ha attraversato una genesi storica ricca di individui di grande intuito e capacità di immaginare tecnologie a decenni dalla loro nascita, una genesi non lineare ma più spesso composta di salti e momenti perfettamente localizzati nello spazio e nel tempo, che hanno avviato un processo in grado di sconvolgere il mondo. Abbiamo anche operato un’archeologia mediale di più ampio respiro, osservando le logiche dominanti dell’interfaccia visuale come strumenti di misura nella pittura e nell’astronomia, fondamentali contributi accademici, militari e arcaici nella sua costituzione nei secoli, i suoi effetti sullo sguardo esaminati secondo la teoria e le pratiche del cinema, i dispositivi che l’hanno ispirata attraverso la storia e il rapporto tra soggetto e informazione che simili dispositivi hanno delineato e ricostruito nelle interfacce contemporanee. Sulla scia della mediarcheologia abbiamo anche esposto le potenziali agentività che un’interfaccia produce come finestra incorniciata e come vetrina, appoggiandoci a pensatori e pensatrici come Simmel e Friedberg. Abbiamo percorso i passi teorici e tecnici che hanno condotto dai primi schermi radar ai nostri computer e telefoni cellulari, passando per le menti visionarie che hanno saputo immaginare e poi realizzare le tappe più importanti della strada che ha condotto dall’inizio del secondo dopoguerra ad oggi.

Una volta individuate le caratteristiche e le logiche fondamentali dell'interfaccia, è stato possibile procedere con il nucleo del presente lavoro, presentato nel terzo capitolo: la proposta di una tassonomia criticamente informata, avviata con una distinzione sul piano ontologico tra tre famiglie di interfacce definite secondo le loro caratteristiche hardware, le forme e le logiche del loro funzionamento o la loro teleologia.

Abbiamo identificato, per ognuna di queste classi ontologiche, una norma dominante e i suoi caratteri, i meccanismi e i dispositivi normalizzanti che impone sugli individui e quelli che impone alle e attraverso le tecnologie che le pertengono e, come avviene nella maggior parte dei processi normalizzanti, le dinamiche tramite cui questa norma si definisce, andando a definire delle non-norme che riflessivamente delineano la norma iniziale. In questo percorso è stato possibile incontrare molteplici forme di esclusione e di norme: norme corporee, norme culturali, norme tecniche, norme sociali, norme apertamente imposte all'interfaccia per garantirne specifici funzionamenti e norme emergenti dalla tecnica o ereditate da tecniche precedenti. È stato anche possibile considerare diversi casi studio che risultano dalla combinazione modulare di diverse forme dell'interfaccia, così come numerosi esempi di progettazione di interfacce e dispositivi che rivelano uno sforzo attivo nello scardinare le norme imposte e cercano di produrre nuove agentività e nuovi rapporti tra utenti e strumenti.

In ultimo, è stata esposta la natura eccezionale della Realtà Aumentata come forma singolare tecnologia anaclitica, in grado di appoggiare efficacemente i propri processi a tecnologie e interfacce molto diverse tra loro per produrre effetti e dispositivi estremamente differenti. La duttilità della sua definizione, la giovinezza dello strumento e della ricerca su di esso e la sua capacità di diffondersi fluidamente attraverso altre forme tecnologiche rendono la Realtà Aumentata un dispositivo in grado di parassitare e insieme potenziare le più diverse interfacce.

In ultimo, riteniamo che la proposta tassonomica non sia, per sua definizione, un lavoro che possa mai dirsi definitivamente ed eternamente concluso: una classificazione è sempre un lavoro aperto all'emergere di nuove istanze e categorie,

alla scoperta di nuovi esemplari, allo sviluppo di nuovi strumenti critici, all'insorgere di nuove domande di ricerca. D'altra parte, la categorizzazione qui proposta abbraccia modularmente e dinamicamente le tecnologie digitali in maniera aperta e trasversale, permettendo di inquadrare strumenti complessi come la Realtà Virtuale e l'informatica tascabile nella stessa struttura in maniera chiara e stratificata, e insieme di identificare con precisione e mettere in relazione tecnologie che rappresentano ancora fenomeni di complessa e necessaria esplorazione, occasioni per la nascita di nuovi campi di studio e potenziali domande di ricerca future.

Bibliografia

- Aarseth, E. J., *Cybertext: Perspectives on Ergodic Literature*, The John Hopkins University Press, Baltimora (1997)
- Adami, E., *Multimodality*, in O. Garcìa et al. (eds.), *The Oxford Handbook of Language and Society*, Oxford University Press, eBook (2016), pp. 451-472
- Ambrose, R., et al., *Mobile manipulation using NASA's Robonauts*, in *IEEE International Conference on Robotics and Automation Proceedings* (2004), pp. 2104-2109
- Bach B., et al., *The Hologram in My Hand: How Effective is Interactive Exploration of 3D Visualizations in Immersive Tangible Augmented Reality?*, in *IEEE Transactions on Visualization and Computer Graphics* (2018), pp. 457-467
- Barbara, J., M. Haahr, *The Role of Voice in Virtual Reality Interactive Narratives*, in *The Journal of Interactive Narrative*, Vol. 1 (2024), online
- Bardini, T., *Bootstrapping: Douglas Engelbart, Coevolution, and the Origins of Personal Computing*, Stanford University Press, Stanford, CA (2000)
- Bärmann L., et al., *Incremental Learning of Humanoid Robot Behavior from Natural Interaction and Large Language Models*, in *Frontiers in Robotics and AI, Robot Learning and Evolution*, Vol. 11 (2024), online
- Basdogan C., et al., *A Review of Surface Haptics: Enabling Tactile Effects on Touch Surfaces*, in *IEEE Transactions on Haptics*, Vol. 13, No. 3 (2020), pp. 450-470
- Beeson, M. J., *The Mechanization of Mathematics*, in C. Teucher (ed.), *Alan Turing: Life and Legacy of a Great Thinker*, Springer (2004), pp. 77-134
- Benfenati A., et al., *Using Transformers Perception by Exploring Input Manifolds*, in processo di revisione
- Benjamin, W., *L'opera d'arte nell'epoca della sua riproducibilità tecnica*, Feltrinelli, Milano (2021)
- Benvenuto, N., Zorzi, M., *Principles of Communications Networks and Systems*, John Wiley & Sons Ltd., Chichester, UK (2011)
- Bolanowski, J. S., et al., *Four Channels Mediate the Mechanical Aspects of Touch*, in *The Journal of the Acoustical Society of America*, Vol. 84 (1988), pp. 1680-1694
- Bolt, R. A., *"Put-That-There": Voice and Gesture at the Graphics Interface*, report to the Cybernetics Technology Division of the Defense Advanced Research Projects Agency, MIT (1980)
- Boslaugh, D. L., *When computers went to sea. The digitalization of the United States Navy*, IEEE Computer Society, Los Alamos (1999)
- Bousseta, R., et al., *EEG Based Brain Computer Interface for Controlling a Robot Arm Movement Through Thought*, in *IRBM*, Vol. 39, No. 2 (2018), pp. 129-135
- Bruno, G., *Atlante delle emozioni. In viaggio tra arte, architettura e cinema*, Mondadori, Milano (2006)

- Bruno, G., *Surface: Matters of Aesthetics, Materiality and Media*, University of Chicago Press, Chicago (2014)
- Bryan, M., et al., *An Adaptive Brain-Computer Interface for Humanoid Robot Control*, in *IEEE-RAS International Conference on Humanoid Robots Proceedings* (2011), pp. 199-204
- Bruegger, P., Hirsbrunner, B., *Kinetic User Interface: Interaction through Motion for Pervasive Computing Systems*, in Stephanidis, C. (ed.), *Universal Access in Human-Computer Interaction, HCI International Conference Proceedings*, San Diego (2009), pp. 297-306
- Buxton, B., *Sketching User Experiences: Getting the Design Right and the Right Design*, Morgan Kaufmann Pub, Burlington (2011)
- Bush, V., *As We May Think*, in *The Atlantic Monthly* (1945)
- Bush, V., *Oscillating-current Circuits: An Extension of the Theory of Generalized Angular Velocities, with Applications to the Coupled Circuit and the Artificial Transmission Line*, <https://genealogy.math.ndsu.nodak.edu/id.php?id=84507>
- Cai, L., Chen, H., *TouchLogger: Inferring Keystrokes On Touch Screen From Smartphone Motion*, in *6th USENIX Workshop on Hot Topics in Security*, San Francisco (2011), conference paper
- Cantu A., et al., *Does Folding Improve the Usability of Interactive Surfaces in Future Airliner Cockpits? An Evaluation Under Turbulent Conditions and Varying Cognitive Load*, in *IMH '21 Conference Proceedings* (2021), online
- Cardille, J. A., et al., *Cloud-Based Remote Sensing with Google Earth Engine*, Springer (2023)
- Cartwright, W., et al., *Geospatial Information Visualization User Interface Issues*, in *Cartography and Geographic Information System*, Vol. 28 (2001), pp. 45-60
- Ceruzzi, P. E., *A History of Modern Computing*, MIT Press, Cambridge, MA (1998)
- Charman-Anderson, S., *Ada Lovelace: A Simple Solution to a Lengthy Controversy*, in *Patterns*, New York (Oct 2020), online
- Çöltekin, A., et al., *Gaze and Feet as Additional Input Modalities for Interacting with Geospatial Interfaces*, in *ISPRS Annals of the Photogrammetry Remote Sensing and Spatial Information Sciences*, III-2 (2016), pp. 113-120
- Congiu, S., *Architettura degli elaboratori. Organizzazione dell'hardware e programmazione in linguaggio assembly*, Pàtron editore, Granarolo dell'Emilia, BO (2012)
- Dachselt, R., *The Challenge to Build Flexible User Interface Components for non-immersive 3D Environments*, in Bullinger, H.-J., Ziegler, J. (eds.), *HCI (2)* (1999), pp. 1055-1059
- Dalmasso, A.C., Pirandello, S., *Augmented Imagination. Thinking Technology Beyond Extension*, in *I Castelli di Yale Online. Annali di Filosofia*, Vol. 12, n.2 (2024), pp. 29-42
- Descartes, R., *Le passioni dell'anima*, Bompiani, Milano (2003)

- Dreyfus-Graf, J., *Sonograph and Sound Mechanics*, in *The Journal of the Acoustic Society of America*, Vol. 22 (1950), pp. 731-739
- Eichhorn, E., *The Tangible Interaction Framework Cards – Introduction*, <http://www.ehornecker.de/>
- Ekman, P., Friesen, W. V., *Constants across cultures in the face and emotion*, in *Journal of Personality and Social Psychology*, Vol. 17(1971), pp. 124–129
- Engelbart, D., *Augmenting Human Intellect: A Conceptual Framework*, SRI Summary Report AFOSR-3223 (October 1962)
- Eugeni, R., *La condizione postmediale*, La Scuola, Brescia (2015)
- Fitts, P. M., *Psychological Research on Equipment Design - Report no. 19*, Army Air Forces Aviation Psychology Program Research Reports, Washington D.C. (1947)
- Fitts, P. M., *Psychological Research on Equipment Design in the AAF*, in *American Psychologist*, Vol. 2 (1947), pp.93-98
- Fleck, L., *Genesis and Development of a Scientific Fact*, Bradley, F., Trewn, T.J. (trad.), University of Chicago Press, London (1979)
- Friedberg, A., *Window Shopping. Cinema and the Postmodern*, University of California Press, Berkley (1994)
- Galloway, A., *The Interface Effect*, Polity Press, Cambridge (2012)
- Garcia, J., Brock, A. M., *CandyFly: Bringing Fun to Drone Pilots with Disabilities through Adapted and Adaptable Interactions*, in *CHI '22 Conference Proceedings*, New Orleans (2022), online
- Garland Thomson, R., *Extraordinary Bodies: Figuring Physical Disability in American Culture and Literature*, Columbia University Press, New York (1997)
- Garrad, C., *Geoprocessing with Python*, Manning, eBook (2016)
- Gascoigne, W., *Letter to William Oughtred*, c. February 1641, in Sellers, D., *In Search of William Gascoigne*, Astrophysics and Space Science Library, Vol. 390, New York (2012), p. 184
- Gibson, J., *Senses as Perceptual Systems*, George Allen & Unwin, London (1966)
- Grespi, B., Malavasi, L., *Dalla parte delle immagini – Temi di cultura visuale*, McGraw Hill Education Italy, Milano (2022)
- Guidotti, R., et al., *A Survey of Methods for Explaining Black Box Models*, in *ACM Computer Surveys*, Vol. 51 (2018), online
- Hadley, S. W., Lam, K., *Light Field and Crosshair Quality Assurance Test using a Simple Lens System*, in *Medical Physics*, 33 (2006), pp. 930-932
- Hafez, M., Benali Khoudja, M., *Tactile Interfaces: Technologies, Applications and Challenges*, in *REM2005 Conference Proceedings* (2005), pp. 267-272
- Harper, S., *Is There Design-For-All?*, in *Universal Access in the Information Society*, Vol. 6 (2007), pp. 111-113

- Hatscher, B., et al., *GazeTap: Towards Hands-Free Interaction in the Operating Room*, in *ICTMI'17 Conference Proceedings*, Glasgow (2017), pp. 243-251
- Heidegger, M., *Che cosa significa pensare?*, G. Vattimo (a cura di), Sugarco, Carnago (1971)
- Hendrikx, M., et al., *Procedural Content Generation for Games: a Survey*, in *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, Vol. 9 (2013), online
- Hornecker, E., Buur, J., *Getting a Grip on Tangible Interaction: A Framework on Physical Space and Social Interaction*, in *CHI 2006: Designing for Tangible Interaction Proceedings*, Montréal (2006), pp. 437-446
- Hyman, R. A., *Charles Babbage – Pioneer of the Computer*, Princeton University Press, Princeton, NJ (1982)
- Ihde, D., Malafouris, L., *Homo Faber Revisited: Postphenomenology and Material Engagement Theory*, in *Philosophy and Technology*, 32 (2019), pp. 1-20
- Jain, D., et al., *HomeSound: An Iterative Field Deployment of an In-Home Sounds Awareness System for Deaf or Hard of Hearing Users*, in *CHI 2020 Conference Proceedings*, Honolulu (Aprile 2020), online
- Janlert, L. E., and Stolterman, E., *Faceless Interaction — a Conceptual Examination of the Notion of Interface: Past, Present, and Future*, in *Human–Computer Interaction*, 30 (2015)
- Jeon, S., Choi, S., *Haptic Augmented Reality: Taxonomy and an Example of Stiffness Modulation*, in *Presence*, Vol. 18, No. 5 (2009), pp. 387-408
- Johnson, E. A., *Touch Display — a Novel Input/Output Device for Computers*, in *The Institution of Engineering and Technology*, Vol. 1 (1965), pp. 219-220
- Johnson, J., et al., *The Xerox Star: A Retrospective*, in *IEEE standard* (September 1989), pp. 11-26
- Kaklauskas, A., et al., *Affective Internet of Things*, in Pacheco-Torgal, F., et al. (eds.), *Woodhead Publishing Series in Civil and Structural Engineering*, Woodhead Publishing, online (2020), pp. 203-233
- Kapp, E., *Elements of a Philosophy of Technology*, Kirkwood, J. W., Weatherby, L. (eds.), University of Minnesota Press, Minneapolis (2018)
- Kocijan, V., et al., *The Defeat of the Winograd Schema Challenge*, in *Artificial Intelligence*, Vol. 325 (2023), online
- Kyriakou, P., Hermon, S., *Can I touch this? Using Natural Interaction in a Museum Augmented Reality System*, in *Digital Applications in Archaeology and Cultural Heritage*, Vol. 12 (2019), online
- Levy, S., *Insanely Great*, Viking Penguin, New York (1994)
- Lovesque, H. J., et al., *The Winograd Schema Challenge*, in *Thirteen International Conference on Principles of Knowledge Representation and Reasoning proceedings*, AAAI (2012), pp. 552-561

- Lyan, I. P., et al., *On the Issue of Energy Consumption of Vibration Technological Machines*, in *IOP Conference Series: Materials Science and Engineering 747 Proceedings* (2020), online
- Malafouris, L., *The Brain-Artifact Interface (BAI): a Challenge for Archaeology and Cultural Neuroscience*, in *Social Cognitive and Affective Neuroscience*, 2-3 (2010), pp. 264-73
- Manovich, L., *Il linguaggio dei nuovi media*, Merlini, R. (trad.), Edizioni Olivares, Milano (2022), p.64
- McLuhan, M., *Gli strumenti del comunicare*, Capriolo, E. (trad.), Garzanti, Milano (1967)
- McLuhan, M., *La galassia Gutenberg. Nascita dell'uomo tipografico*, Rizzo, S. (trad.), Armando, Roma (2011)
- Meechan, K., et al., *Crosshair, Semi-automated Targeting for Electron Microscopy with a Motorised Ultramicrotome*, in *eLife*, 11 (2022), online
- Minsky, M., *Telepresence*, 1980. Digitalizzato da Philpapers a <https://philpapers.org/rec/MINT>, ultimo accesso 31/11/2024
- Modena E., et al., *360°. L'immagine ambientale nelle arti visive tra realtà virtuale e aumentata*, in *PIANO B. Ari e Culture Visive*, Vol. 6, 1 (2021), online
- Moggridge, B., *Designing Interaction*, MIT Press, Cambridge, MA (2007)
- Monteiro, P., et al., *Hands-free Interaction in Immersive Virtual Reality: A Systematic Review*, in *IEEE Transactions on Visualization and Computer Graphics*, Vol. 27 (2021), pp. 2702-2713
- Mullaney, T. S., *The Chinese Typewriter, a History*, The MIT Press, Cambridge, MA (2018)
- Murad, C., et al., *Designing Voice Interfaces: Back to the (Curriculum) Basics*, in *CHI '20 Conference on Human Factors in Computing Systems Conference Proceedings* (2020), online
- Murray, J. H., *Inventing the Medium. Principles of Interaction Design as Cultural Practice*, MIT Press, Cambridge MA (2011)
- Neumann, M. M., Neumann, D. L., *Touch Screen Tablets and Emergent Literacy*, in *Early Child Education Journal*, 42 (2014), pp. 231-239
- Norman, D., *The Design of Everyday Things*, Basic Books (New York 2013)
- Ortiz-Catalan, M., et al., *A Highly Integrated Bionic Hand with Neural Control and Feedback for Use in Daily Life*, in *Science Robotics*, 8 (2023), online
- Ortiz-Catalan, M., et al., *An Osseointegrated Human-Machine Gateway for Long-Term Sensory Feedback and Motor Control of Artificial Limbs*, in *Science Translational Medicine*, 6 (2014), online
- Pallotta, V., *Unobtrusive Interaction with Mobile and Ubiquitous Computing Systems through Kinetic User Interfaces*, in Cruz-Cunha, M. M., Moreira, F., (eds.), *Handbook of Research on Mobility and Computing: Evolving Technologies and Ubiquitous Impacts*, IGI Global (2011), pp. 689-702

- Picard, R. W., *Affective Computing*, The MIT Press, Cambridge MA (2000)
- Picard, R. W., et al., *Toward Machine Emotional Intelligence: Analysis of Affective Physiological State*, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23 (2001), pp. 1175-1191
- Pieraccini, R., *The Voice in the Machine. Building Computers That Understand Human Speech*, The MIT Press, Cambridge MA (2012)
- Pinotti, A., *Alla soglia dell'immagine. Da Narciso alla realtà virtuale*, Einaudi, Torino (2021)
- Reeves, L. M., et al., *Guidelines for Multimodal User Interface Design*, in *Commun. of the ACM*, Vol. 47 (2004), pp. 57-59
- Rödinger, T., et al., *EarRumble: Discreet Hands- and Eyes-Free Input by Voluntary Tensor Tympani Muscle Contraction*, in *CHI '21 Conference Proceedings*, Yokohama (2021), online
- Rose, D. C., et al., *Technical Note: The Use of Geographical Information Systems Software for the Spatial Analysis of Bone Microstructure*, in *American Journal of Physical Anthropology*, Vol. 148 (2012), pp. 648-654
- Russell, J. A., Pratt, G., *A Description of the Affective Quality Attributed to Environments*, in *Journal of Personality and Social Psychology*, 38 (1980), pp. 311-322
- Schutte, P. C., Trujillo, A. C., *Flight Crew Task Management in Non-normal Situations*, in *The Human Factors and Ergonomics Society 1996 Annual Meeting Conference Proceedings*, Vol. 40 (1996), pp. 244-248
- Seo, H., Lee, C. S., *Emotion Matters: What Happens Between Young Children and Parents in a Touch Screen World*, in *International Journal of Communication*, 11 (2017), pp. 561-580
- Shannon, C.E., Weaver, W., *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana (1964)
- Simmel, G., *L'ansa del vaso. Saggio di estetica*, in G. Simmel, *Stile moderno*, Carnevali, B., Pinotti, A. (a cura di), Einaudi, Torino (2020), pp. 306-313
- Sidenmark, L., et al., *Radi-Eye - Hands-Free Radial Interfaces for 3D Interaction using Gaze-Activated Head-Crossing*, in *CHI '21 Conference Proceedings*, Yokohama (2021), online
- Slater, M., Wilbur, S., *A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments*, in *Presence: Teleoperators and Virtual Environments*, Vol. 6 (1997), pp. 603-616
- Slater, M., *Immersion and the Illusion of Presence in Virtual Reality*, in *British Journal of Psychology*, Vol. 109 (2018), pp. 431-433
- Smith, D. K., Alexander, R. C., *Fumbling the Future. How Xerox Invented, then Ignored, the First Personal Computer*, William Morrow and Company Inc., New York (1988)

- Sol Roo, J., et al., *Inner Garden: Connecting Inner States to a Mixed Reality Sandbox for Mindfulness*, in *CHI '17 International Conference of Human-Computer Interaction Proceedings* (2017), pp. 1459-1470
- Sumak, B., et al., *An Empirical Evaluation of a Hands-free Computer Interaction for Users with Motor Disabilities*, in *Journal of Biomedical Informatics*, 96 (2019), online
- Sutherland, I., *The TX-2 Computer and Sketchpad*, in *Lincoln Laboratory Journal*, Vol. 19, 1 (2012), online
- Turk, M., *Multimodal Interaction: A Review*, in *Pattern Recognition Letters*, Vol. 36 (2014), pp. 189-195
- Turner, F., *From Counterculture to Cyberculture*, The University of Chicago Press, Chicago (2006)
- Ullmer, B., Ishii, H., *Emerging Frameworks for Tangible User Interfaces*, in *IBM Systems Journal* (2000), pp. 915-931
- Valli, A., et al., *Natural Interaction Sensitivetable*, in *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, (2008), pp. 2315-2318
- Valli, A., *The Design of Natural Interaction*, in *Multimedia Tools and Applications*, Vol. 38, 3 (2008), pp. 295-305
- Vaswani, A. et al., *Attention Is All You Need*, in *31st Conference on Neural Information Processing System* (2017), pp. 6000-6010
- Vatavu, R.-D., *From Natural to Non-Natural Interaction: Embracing Interaction Design Beyond the Accepted Convention of Natural*, in *ICMI '23 25th International Conference on Multimodal Interaction proceedings* (2023), pp. 684-688
- Wang, S. et al., *Designing a Generalized 3D Carousel View*, in *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, New York (2005) pp. 2017-2020
- Watanabe, T., S. Fukui, *A Method for Controlling Tactile Sensation of Surface Roughness Using Ultrasonic Vibration*, in *IEEE International Conference on Robotics and Automation Proceedings*, Vol. 1 (1995), pp. 1134-1139
- Wiesing, L., *Pause of Participation. On the Function of Artificial Presence*, in *Research in Phenomenology*, Vol. 41, 2 (2011), pp. 238-252
- Winograd, T., *Procedures as a Representation for Data in a Computer Program for Understanding Natural Language*, The Massachusetts Institute of Technology (1971)
- Wolf, M. J. P., *Before the Crash. Early Video Game History*, Wayne State University Press, Detroit (2012)
- Xiao, B., et al., *Modeling Multimodal Integration Patterns and Performance in Seniors: Toward Adaptive Processing of Individual Differences*, in: *ACM International Conference on Multimodal Interfaces Proceedings*, Vancouver (2003), pp. 265-272
- Yarbus, A. L., *Eye Movement and Vision*, Springer Science+Business Media, New York (1967)
- Devi Mahatmya*, Udai Nath, B. (trad.), online, <https://www.famigliafideus.com/wp-content/uploads/2023/03/DEVI-MAHATMYA-Beatrice-Udai-Nath.pdf>

Micrometer, in *Encyclopaedia Britannica – A Dictionary of Arts, Sciences and General Literature*, Ninth Edition, Vol. 16, Charles Scribner's Sons, New York (1883)

Reference Manual, IBM 7090 Data Processing System, International Business Machines Corporation (Minor Revisions ed. 1962)

XEROX DOCUMENT SYSTEM, Reference Manual, versione 5.1, Xerox office product division (1980)

Sitografia

Letondal, C. et al., *Transface*, <https://lii.enac.fr/projects/transfaces/>, ultima consultazione 13/12/24

Markoff, J., *The real history of WYSIWYG*, <https://archive.nytimes.com/bits.blogs.nytimes.com/2007/10/18/the-real-history-of-wysiwyg/>

Sherif, A., *Nintendo DS lifetime unit sales worldwide as of 2024, by region*, <https://www.statista.com/statistics/1101884/unit-sales-nintendo-ds-region/>

Douglas Engelbart, in *Encyclopedia Britannica*, <https://www.britannica.com/biography/Douglas-Engelbart>

Parc, in *Encyclopedia Britannica*, <https://www.britannica.com/topic/PARC-company>

Vannevar Bush, in *Encyclopedia Britannica*, <https://www.britannica.com/biography/Vannevar-Bush>

Put That There (Original), YouTube, <https://www.youtube.com/watch?v=RyBEUyEtxQo>, ultima visita 31/10/2024

Italian grandmother learns to use Google Home, YouTube, <https://www.youtube.com/watch?v=e2RONSKtVA0> (2018)

How clams help keep Polish water clean, *The Economist*, <https://www.economist.com/europe/2021/01/21/how-clams-help-keep-polish-water-clean>

ACAT, Intel, <https://github.com/intel/acat>

Hololens documentation, Microsoft, <https://www.microsoft.com/en-us/hololens/hardware#document-experiences>

Smart DX Evo, BTS Bioengineering, <https://www.btsbioengineering.com/it/products/smart-dx-evo/>

How are prosthetic arms controlled, *Unlimited Tomorrow*, <https://www.unlimitedtomorrow.com/how-are-prosthetic-arms-controlled/>

Microsoft documentation, HoloLens 2, https://news.microsoft.com/wp-content/uploads/prod/2019/02/Fact-Sheet_HoloLens2

Mindwave documentation, Neurosky, <https://store.neurosky.com/pages/mindwave>

User manual, Psyonic, <https://www.psyonic.io/users>

Epoc X, Emotiv, <https://www.emotiv.com/products/epoc-x>

DS4Windows, <https://ds4-windows.com/>

Material, Google, <https://m3.material.io/>

Remarkable, <https://remarkable.com/>

Natural language, in *The Oxford Dictionary*,
<https://www.oxfordlearnersdictionaries.com/definition/english/natural-language>

Interface, in *The Merriam-Webster Dictionary* <https://www.merriam-webster.com/dictionary/interface>

Bergen, P., blocks world, <https://patrickvanbergen.com/blocks-world/>

iOS documentation, Apple,
<https://web.archive.org/web/20120910202035/http://www.apple.com/ios/>

Mediografia

Itō, S., *L'Organo Genocida* (虐殺器官, Gyakusatsu kikan), regia di S. Murase, Geno Studio (Suginami, Tokyo 2017)

Kojima, H., *Death Stranding*, Play Station 4, Kojima Productions (2019)

Visual Concepts, *NBA 2K15*, 2K Sports (2014)