

Interazione uomo-macchina 2022

Introduzione al corso

Professori

- Federico Cabitza (ingegnere interazionista)
- Pietro Guardini

Per scrivere privatamente usare sempre l'email: federico.cabizza@campus.unimib.it

Mettere sempre per quale corso (IUM) si sta scrivendo, nome, cognome e matricola.

Solo privatamente!!

Usare bene Avvisi, Forum e Spazi Domande per condividere a tutti le domande che possono essere utili a tutti.

Orari

Lunedì 15:30-17:45 U2-07

Martedì 14:00-16:30 U3-09

Venerdì 9:30-11:30 U24-C2

Non c'è distinzione tra le lezioni frontali, esercitazioni o laboratorio

Modalità d'esame

Scritto (12 pnt)

Scritto con domande a risposta chiusa e aperta

Il tipo d'esame dipenderà dalla presenza fisica o meno

- è sempre online
- ci sono soprattutto domande chiuse (3 opzione e ci sono penalizzazioni su solo una delle tua risposte sbagliate)
- Le domande aperte sono su concetti fondamentali e definizione
- i risultati in giornata

Iscriversi cmq all'appello di Giugno anche se volete fare Luglio!!

Documentazione del Progetto (18 pnt)

Componenti: min 2 - max 4 - perf 3

Creare gruppo mandare una mail con soggetto “IUMGRUPPO” e nomi, cognomi e matricole delle componenti.

Documento di progetto va consegnato almeno 5 gg feriali prima della data dell'orale

Punti extra

1. LEGGERE un libro e fare 3-4 cit per capitolo prima del primo scritto.
Leggerlo subito!! (1 pnt se letto prima del primo scritto)

Elenco dei libri:

- La gabbia di vetro - Nicholas Carr
- Internet ci rende Stupidi? - Nicholas Carr
- Il lato oscuro della rete - Nicholas Carr
- Intelligenza artificiale - Floridi e Cabitza
- Il filtro - Eli Parisier
- Internet non salverà il mondo - Evgeny Morozov
- Silicon valley. il signore del silicio - Evgeny Morozov
- Demenza digitale - Manfred Spitzer
- Solitudine digitale - Manfred Spitzer
- Armi di distruzione matematica - Cathy O'Neil

2. Partecipa ad esercitazione “sul campo” (1 pnt)

- Analisi di usabilità trasversale a livello consulenziale di due o più sistemi interattivi commerciali esistenti → ci danno un sistema reale
- Analisi di usabilità longitudinale di un prototipo

Consegno slide

Orale

Temi del corso

Utente: Cosa vuole e se è soddisfatto di quello che ottiene.

Sono più utenti quindi cambia tutto ⇒ statistica campionaria

Valutazione di usabilità di livello professionale (valore di mercato in crescita)

Test con una decina di persone è 20 mila sterline (con test)

Interfaccia utente - User-center Design - Ingegneria dell'usabilità - Ingegneria semiotica -
 Comunicazione - Interaction Design - Persuading Computer - Usabilità - Interazione
 Uomo-macchina - Utente - Etica della progettazione (INFORMATICA)

Libri

Libro: Intelligenza artificiale

Donald A. Norma

L'amore viene quando lo si merita, quando le particolari caratterizzazioni di un oggetto le rendono parte della vera quotidianità, quando ne approfondiscono la soddisfazione, attraverso la bellezza, il comportamento e la componente riflessiva.

Tenere in casa cose utili e cose Belle!

Non progettate nulla che non vorreste usare voi stessi, o meglio, che un vostro parente non trovi utile e facile da usare.

L'interfaccia efficiente non è solo per migliorare la frustrazione dell'utente ma anche la giusta usabilità del macchinario e dei suoi effetti. (ex: Therac-25)

Incidenti nucleari per colpa di un interfaccia.

Obiettivi del corso

Fornire una introduzione (pratica e teorica) alle problematiche del design dell'interfaccia uomo-macchina, per la progettazione di sistemi interattivi usabili e per la loro valutazione.

Terminologia

- Design: progettazione
- Progettazione: “ideare qualcosa e studiare il modo di attuarla, così che porti valore dove collocata\usata per qualcuno”
- Sistema interattivo: ...
- Valutazione: ...
- Interazione: processo in cui due o più oggetti (agenti e sistemi) agiscono in modo da influenzarsi reciprocamente.

INTERAZIONE UOMO MACCHINA

08/03/2022

HUMAN COMPUTER INTERACTION (HCI)

E' una disciplina che si occupa della progettazione, realizzazione e valutazione di sistemi interattivi con capacità computazionali destinati all'uso umano; e dello studio dei principali fenomeni che li circondano.

La valutazione prende i requisiti e le aspettative e valuta il sistema.

Viene fatta a fianco del cliente e inserita in un contesto reale.

Guarda le caratteristiche non funzionali.

La verifica invece è formale ed è basata sulle funzionalità del sistema



Si parla di IUUMM: Interazione Uomo-Uomo Mediata dalla Macchina, poiché l'uomo è sociale.

UNA BREVE STORIA

LHCI nasce negli anni '80.

Ha le sue origini nell'intersezione di due aree disciplinari molto diverse: ergonomia e informatica. Le conoscenze e le competenze di chi lavora in questi ambiti si integrano.

Progettare per l'interazione, quindi È progettare per l'utente (che fa qualcosa).

Ergonomia: studio delle attività.

Ergonomia cognitiva: studio dell'interazione tra l'uomo e gli strumenti per l'elaborazione di informazione studiando i processi cognitivi coinvolti (percezione, attenzione, memoria, pensiero, linguaggio, emozioni), e suggerendo delle soluzioni per migliorare tali strumenti.

E' bene creare il sistema migliore per l'utente medio, NON per se stessi (introspezione).

Perchè vanno fatti molti test per rendere l'utente soddisfatto e il sistema di facile usabilità.

Un'applicazione ha tre aspetti da considerare:

- Varietà dei sistemi interattivi: sensori, schermi, tastiere, cabine di pilotaggio, elettrodomestici, dispositivi per medici...
- Varietà degli utenti: età diverse, culture diverse, conoscenze diverse
- Varietà degli scopi e degli usi: obiettivi scopi e vincoli temporali, curiosità, molte situazioni differenti

TEMI DELL'HCI

- Criteri, metodi e strumenti per la progettazione dell'interazione fra esseri umani e sistemi interattivi
- Criteri, metodi e strumenti per la valutazione dell'usabilità dei sistemi interattivi

Definizione dell'usabilità da sapere a memoria

- Progettazione di nuove tecniche di interazione
- Valutazione dell'impatto dell'automazione nei contesti umani

Impatto diretto nel breve, sui singoli e nel qui-e-ora (**usabilità**) e impatto indiretto nel medio-lungo sulla collettività (conseguenze inattese).

⚠ L'usabilità è un tipo di effetto/impatto sui singoli utenti, NON una caratteristica intrinseca di un sistema. L'usabilità quindi non si può valutare senza il coinvolgimento degli utenti e sarebbe meglio valutarla in un contesto il più simile possibile alla realtà. ⚡

Non esistono metodologie per insegnarci e progettare una “buona” interazione.

Esiste un approccio, che per quanto semplice è comunque difficile da applicare: usare la creatività, usare le tecniche insegnate e verificare che le idee avute siano accettate e utili dall'utente. Qualsiasi valutazione è meglio di nessuna applicazione.

SISTEMA SOCIO-TECNICO (SOCIO-TECHNICAL SYSTEM – STS)



Strumento usato per forza a due, i minatori conoscono bene come usarlo, tubo che porta a qualcosa: sistema socio tecnico quando vedo un'azione umana con dei ruoli e degli oggetti tecnici.

Insieme di elementi interrelati ed eventualmente mutualmente dipendenti che, agli occhi di un osservatore esterno, appaiono come una entità unitaria ma collettiva, con caratteristiche e comportamento proprio, solitamente autonomo ed intenzionale (cioè volto ad un obiettivo).

Importante la differenza tra osservatore e osservato.

E' l'osservatore che identifica il perimetro del sistema.

Sistema in cui la componente umana (sociale) e quella tecnica (tecnologica) sono inestricabilmente legate tra loro e la loro interazione porta a fenomeni emergenti impredicibili.

L'interazione tra loro porta a dei fenomeni emergenti imprescindibili

E' un concetto di invarianza si scala dal piccolo team di lavoro alla società umana

Le proprietà/comportamento del sistema non sono desumibili da quelle sue parti, prese isolatamente, ma bensì emergono proprio in virtù delle interazioni tra quelle (analisi solo post-hoc) in un determinato ambiente (analisi locale).



Gödel, Escher, Bach : Hofstadter

Proprietà emergenti:

- Funzionali, che riguardano il funzionamento dell'intero sistema una volta che tutte le sue parti, assemblate come devono, funzionano bene
- Non funzionali, che riguardano quanto bene opera il sistema in un determinato ambiente/contesto (es: reliability, security, performance, safety, usability, comfort). Ad esempio:
 - Le informazioni sono giuste ma non aggiornate
 - Il sistema richiede la password per il log in ma questa è stampata su un post-it accanto al monitor
 - Il sistema prevede la verifica con barcode, ma il filo dello strumento è troppo corto
 - Font troppo piccolo

Prendersi del tempo per osservare da chi come e quando il sistema che creo sia usato.

STS THINKING

Un approccio che è consapevole che le due componenti si integrano bene (fit) e si “ottimizzano” solo congiuntamente, in configurazioni sub ottime (joint optimization).

→ Attenzione rinnovata a job satisfaction, workers' needs, and skill enhancement.

L'introduzione di una tecnologia in un contesto (sociale) è parte di un **processo di cambiamento** operante su più piani.

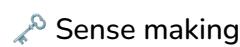
Bauman filosofo della liquidità: tempi liquidi (tempo a cui è difficile dare forma)

L'elemento umano è fondamentale per poter completare un sistema velocemente e produttivamente

METAFORA DELLO YIN YANG

Il sociale e la tecnica non sono del tutto divisi:

- Persone: non sono ergonomia (uomo-macchina) ma anche la dimensione sociale, la conoscenza tacita, le convenzioni, le “unwritten rules”
- Tecniche/Tecnologie: tra cui procedure, protocolli, divisione del lavoro, written policies and rules



Un sistema socio-tecnico è un sistema complesso, in cui è impossibile prevedere ogni possibile evoluzione.. il concetto di **conseguenze inattese** delle tecnologie.

Impatto nel medio e lungo periodo non solo nella dimensione individuale, ma anche sociale

Esempi:

- Covid-19: 18-65 anni le linee sono peggiorate; in Italia i non vaccinati non aumentano mentre i malati aumentano..com'è possibile? È una conseguenza inattesa? **Effetto Peltzman**: le persone stanno abbandonando le misure di sicurezza in alcune situazioni, per cui si ammalano di più anche se sono vaccinati.

 Risk compensation

- La app *Immuni* ha un nome sbagliato perché porta troppo a pensare che l'applicazione possa salvare dai contagi, quando li traccia e basta.
- Ciclisti: aggiungendo l'obbligo del casco gli incidenti gravi sono aumentati, perché le persone si sentono più sicure e fanno azioni più rischiose; succede anche con le ciclabili, in cui aumentano gli incidenti con le macchine agli incroci, quindi per migliorare la cosa, hanno messo la ciclabile allo stesso piano della strada, così stavano tutti e due più attenti.

Per effetto *Peltzman* si intende il fenomeno per cui, là dove si introduce una protezione obbligatoria, la percezione del rischio si abbassa, incentivando comportamenti che vanno nella direzione opposta agli effetti auspicati dalla norma stessa.

Questo effetto nell'ambito delle tecnologie si definisce come Overreliance.

OVERRELIANCE

- **Overdependence:**
 - Mancanza di autonomia
 - Abuso (uso al di là dei bisogni effettivi)
 - Mancanza, dimenticanza o ignoranza di un piano B
- **Overconfidence:**
 - Pensare che non andrà mai giù
 - Pensare che non ne deriverà mai un danno
 - Pensare che non può sbagliare mai

Esempio:

- Troppa fiducia nel navigatore e due macchine sono finite in acqua per un errore tecnologico.
- Goal segnato perchè l'orologio dell'arbitro ha avvisato che era entrata la palla, ma in verità quando si è alzato il portiere ha passato la palla dentro la rete. La linea immaginaria della tecnologia usata è approssimata, ci sono dei millimetri di differenza che possono ingannare gli arbitri nei casi estremi.

Un'idea per poter fare a meno Overconfidence o Overdependence è dichiarare che una tale tecnologia ha una percentuale di errore. Molti macchinari medici hanno infatti queste percentuali, che sono anche molto alte, usare tali macchinari equivale quasi al lancio di una moneta.

Automation-related complacency (overconfidence) e automation bias (overdependence) di Parasuraman:

- La **complacency** riguarda la fiducia che il sistema computazionale funzionerà sempre per come è stato progettato e quindi fa abbassare l'attenzione quando lo si usa
- L'**automation bias** invece è l'eccessiva fiducia nella risposta del supporto alle decisioni e quindi causa errori di omissione o di azione quando i sistemi sono imperfetti

 Startle effect

Riepilogo

Progettare sistemi usabili È progettare per l'uso, quindi per qualcosa che il progettista non controlla e che dipende dall'utente e da quello che fa.

Per progettare sistemi usabili non esistono metodologie, si deve piuttosto:

1. Imparare per (non) imitazione e dall'esperienza
2. Essere creativi (ma non troppo)

3. Valutare il proprio sistema coinvolgendo utenti veri

4. Essere progettisti responsabili

 "I'm only an engineering" → la banalità del male

 Dark patterns

Essere progettisti responsabili significa sapere che il proprio sistema sarà il componente di un sistema socio-tecnico, che può stravolgere o comunque modificare (ad esempio la pratica di persone che lavorano).

Aspettatevi che il vostro lavoro possa avere conseguenze inattese, e lavorate per minimizzarne l'impatto.

BASSA USABILITÀ = DANNI E PROBLEMI

- Gli utenti non capiscono come svolgere i propri compiti con il sistema
- Il sistema presenta un'eccessiva quantità di funzionalità e opzioni (**low use**)
- Gli utenti non capiscono cosa il sistema sta facendo (poca trasparenza)
- Interazione con le tecnologie spesso fonte di emozioni negative (frustrazione)
- Frustrazione, ansia, disorientamento, ira, portano l'utente a compiere errori (più o meno gravi)

Gli oggetti ben progettati sono facili da interpretare e comprendere: contengono indizi visibili (**affordances**) del loro funzionamento.

Gli oggetti progettati male possono essere difficili e frustranti da usare: non offrono indizi o ne danno di sbagliati, oppure sono stati progettati curando l'estetica più che la funzionalità.

Esempi:

- La maggior parte degli errori successi nelle centrali nucleari, sono errori umani per queste cause.
- E' stato verificato che i giudici sono molto più severi prima di pranzo, mentre sono molto più lascivi dopo pranzo.

Per questo vengono create tecnologie per mitigare il rischio di soluzionismo, non eliminarlo.

METAFORA DEDALO E ICARO

Dedalo, i progettisti che danno i mezzi (le ali), avvisano dei vari pericoli (il sole).

Icaro, gli utenti che usano il sistema (le ali), ma non fanno attenzione ai i pericoli (il sole).

Di chi è la colpa???

Dibattito sempre in corso.



Gli oggetti ben progettati sono facili da interpretare e comprendere: contengono indizi visibili (**affordances**) del loro funzionamento.

Anche dare delle indicazioni mano mano per avvisare di più l'utente, può aiutare.

Il concetto di affordance è molto importante nella IUM (e anche nella IUUMM)

Gli oggetti progettati male possono essere difficili e frustranti da usare non offrono indizi o ne danno di sbagliati, oppure sono stati progettati curando l'estetica più che la funzionalità



Un esempio

Per acquistare il biglietto è necessario:

1. Selezionare la tariffa (ridotta o intera) premendo il primo pulsante blu
2. Selezionare la zona di viaggio premendo il secondo pulsante blu
3. Inserire le monete (o la carta di credito – vedi passo 3)
4. Confermare premendo il pulsante verde
5. Attendere circa 20 secondi (!!)
6. Ritirare il biglietto dalla fessura a sinistra.

rpolillo@unimib.it

Se in precedenza hai già scelto una password, inseriscila qui sotto.

Altrimenti procedi cliccando su CONFIRMA

Potrai scegliere la tua Password nella pagina successiva, dopo aver inserito i tuoi dati.

Password

CONTINUA

Non ti ricordi la tua password? [Clicca qui](#)

Ecco la pagina di un sito di commercio elettronico che compare quando si tenta di ordinare un prodotto...

Spot the horror!

14/03/2022

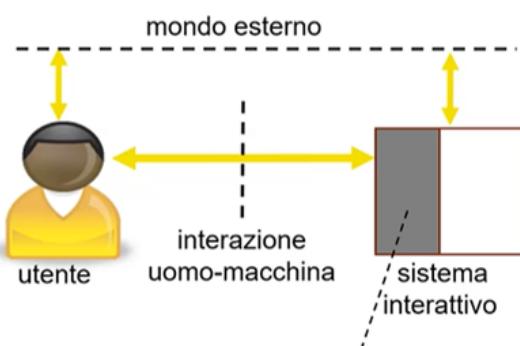
DUE CONCETTI “ESTREMI”

Norman door: tendenza del progettista a concepire oggetti o sistemi che non invitano all’uso (e all’uso corretto) sulla base di elementi visuali o indicazioni chiare (affordances), affidandosi quindi solo alla memoria, all’esperienza o all’intuizione e incontrando spesso il disorientamento o il fraintendimento dell’utente

Sono oggetti di design, fatti male perché non viene capito il loro uso pratico. Questo è ciò che dobbiamo evitare.

Floyd toilet: tendenza del progettista a complicare affari semplici e a pretendere che basti scrivere una procedura per renderla ovvia a qualsiasi utente

Se c’è bisogno di far uso del manuale di istruzioni, c’è un problema sul design



ELEMENTI “BASE”

Sistema interattivo: combinazione di componenti hardware e software che ricevono input da un utente umano, e gli forniscono un output allo scopo di supportare l'effettuazione di un compito

Interfaccia: insieme dei componenti di un sistema interattivo (software o hardware) che forniscono all'utente informazioni e comandi per permettergli di effettuare specifici **compiti** (task) attraverso il sistema

💡 **Halo effect (effetto alone)**

Una persona viene condizionata da elementi piccoli ma che saltano all'occhio. Può essere anche un effetto di discriminazione, non solo in situazioni reali con persone reali ma anche con i siti. Viene notato ciò che interessa “vanno dritti al sodo”.

Per l'utente l'interfaccia È il sistema interattivo, infatti non ha conoscenze sul back-end.

⚠️ Non si parla di progettare interfacce “belle”, ma di progettare elementi che permettono e vincolano l'azione dell'utente sul sistema computazionale (AFFORDANCES e CONSTRAINTS).⚠️ [Fitt's law]

TRE CONCETTI CENTRALI

- **Affordance** ("intuitività"): elementi visibili che invitano qualcuno all'uso corretto
- **Mapping** (mappatura)
- **CONSTRAINT (vincolo)**
 - Strutturale (o passivo): ad esempio concepisco un form con 3 campi e non 4 per vincolare l'inserimento di soli tre tipi di dati, non prevedo un campo “note”, non prevedo la possibilità di feedback

- Funzionale (o attivo): ad esempio concepisco dei controlli per cui un utente non può proseguire se non compila tutti i campi, o non li compila “correttamente” (sulla base di regex, come per l’indirizzo email o il codice fiscale)

Sono strumenti per garantire una migliore qualità dei dati e una standardizzazione delle pratiche, ma non bisogna esagerare con i vincoli, altrimenti si verifica il fenomeno dei workaround.

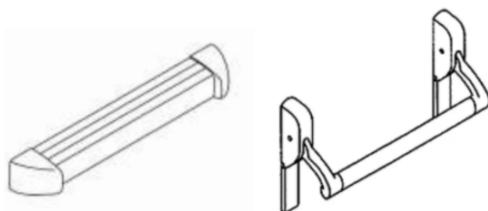
Workaround: azione relativa all'esecuzione di un processo o di un compito (supportati dal sistema informatico) che



non è prevista o descritta nei manuali di uso del sistema informatico e/o nei manuali che descrivono tale processo e procedura e che può bypassare l'uso del sistema o piegarlo ai propri fini.

Invece di Workaround si può anche parlare di **Walkthrough** o “**usì creativi del sistema**” (definizione da usare con l'utente).

AFFORDANCE



Quale delle due invita di più a spingere?

E' meglio quella a sinistra perchè l'unica azione che si può fare è spingere, mentre con quella a destra, la barra è prendibile, e quindi tirabile

Qual'è la traduzione in italiano di Affordance?

Negli anni sono state date diverse definizioni:



Gibson Eleanor: risorsa o supporto che l'ambiente esterno offre all'animale, l'animale deve possedere le capacità per capirlo e usarlo
ex: vedono delle ossa e ne fanno delle armi.

Donald Norman: qualsiasi qualità o proprietà di un oggetto che definisce i suoi possibili usi o rende chiaro come può essere o dovrebbe essere usato

Infine la definizione che si trova tuttora è: *qualsiasi proprietà di un oggetto che invita una persona competente all'azione mediata da tale oggetto*

- Affordance **cognitive**: aiutano gli utenti nelle loro azioni cognitive (pensare, decidere, imparare, ricordare, sapere...)
- Affordance **fisiche**: aiutano gli utenti nella loro azione (cliccare, toccare, muovere, puntare...)***
- Affordance **sensoriali**: aiutano gli utenti nelle loro azioni sensoriali (vedere, sentire...)
- Affordance **funzionali**: aiutano gli utenti ad eseguire dei lavori
- ...
- Affordance **fisica**: ...



*** Calcolatrice: che su pc è analogica (somigliante) uguale a quella fisica, anche per le ombre

Scheumorfismo: deriva dal greco e indica un ornamento fisico o grafico apposto su un oggetto allo scopo di richiamare le caratteristiche estetiche di un altro, come ad esempio una ceramica ornata con dei rivetti per far in modo che ricordi una pentola di metallo.

— [Parte aggiuntiva presa da dispositivi mobili]

Un esempio più tecnologico è un'app per leggere libri su un cellulare, che mostra una libreria con dei libri selezionabili.

Inizialmente è stato usato perché l'utente nuovo era facilitato nel riconoscere delle forme conosciute. Una rappresentazione della realtà in virtuale. (2008)

Non è solo grafica, ma anche il sonoro. Ad esempio l'app per le foto.

Microsoft invece lanciò il metro Design o flat Design con windows 8, perché iniziava ad esser pesante lo scheumorfismo

Elimina texture, sfumature e ombre semi realistiche proprie dello scheumorfismo

Introduce sagome geometriche "flat" (piatte), spazi netti e definiti, colori brillanti, illustrazioni e interfacce bidimensionali e minimaliste. —

- Affordance **create dagli utenti**: ...



- Affordance **emozionali** (o sociali): aiutano gli utenti a esprimere emozioni

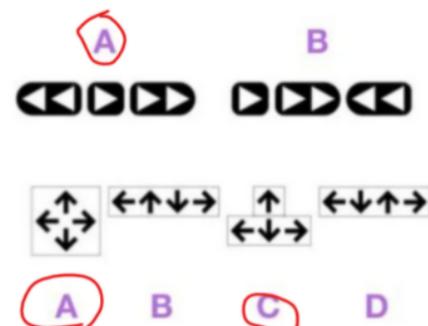


- Affordance sociali: aiutano gli utenti nelle azioni sociali

Riepilogo

Affordance Type	Description	Example
Cognitive affordance	Design feature that helps users in knowing something	A button label that helps users know what will happen if they click on it
Physical affordance	Design feature that helps users in doing a physical action in the interface	A button that is large enough so that users can click on it accurately
Sensory affordance	Design feature that helps users sense something (especially cognitive affordances and physical affordances)	A label font size large enough to be discerned
Functional affordance	Design feature that helps users accomplish work (i.e., usefulness of a system function)	The internal system ability to sort a series of numbers (invoked by users clicking on the Sort button)

18/08/2022



MAPPING

Nel primo esempio è meglio la A? Lo diciamo solo per abitudine? Per altre persone, come gli arabi o gli orientali, che leggono in modo diverso, valo lo stesso?

Nel secondo caso diciamo la C per familiarità, mentre diciamo la A per facilità d'uso, anche se può occupare più spazio.

Quale sarebbe la collocazione più efficiente e intuitiva? -> Dipende dal contesto e dall'uso che ne va fatto.

La convenzione di lettura ci potrebbe dire che il pulsante che va verso il basso sia quello al suo fianco, ma non è così.

Il simbolo della freccia è frantendibile. (Ex: indigeno)

[Non sono sicuro di quale sia il giusto controllo (nell'esempio sotto: il pulsante da premere è quello più vicino alla freccia in verticale o in orizzontale?).]



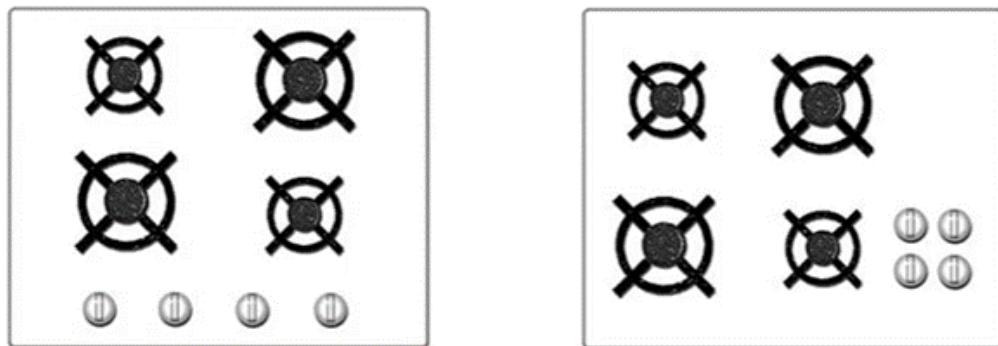
Il mapping è una caratteristica dell'affordance.

Mapping: relazione tra controllo ed effettore.

Nell'esempio sotto:

- Controlli e effettori se messi nella stessa configurazione sono molto più intuitibili, ma occupano più spazio. Nel altro esempio classico magari si può leggere in maniera occidentale
- il mapping **topologico** di sinistra è buono, ma l'affordance può essere migliorata, perché la rotazione in senso orario non ha logicamente nulla a che fare con l'intensità della fiamma.

- La relazione tra, da una parte un controllo e l'azione che esso “afforda” e, dall'altra,

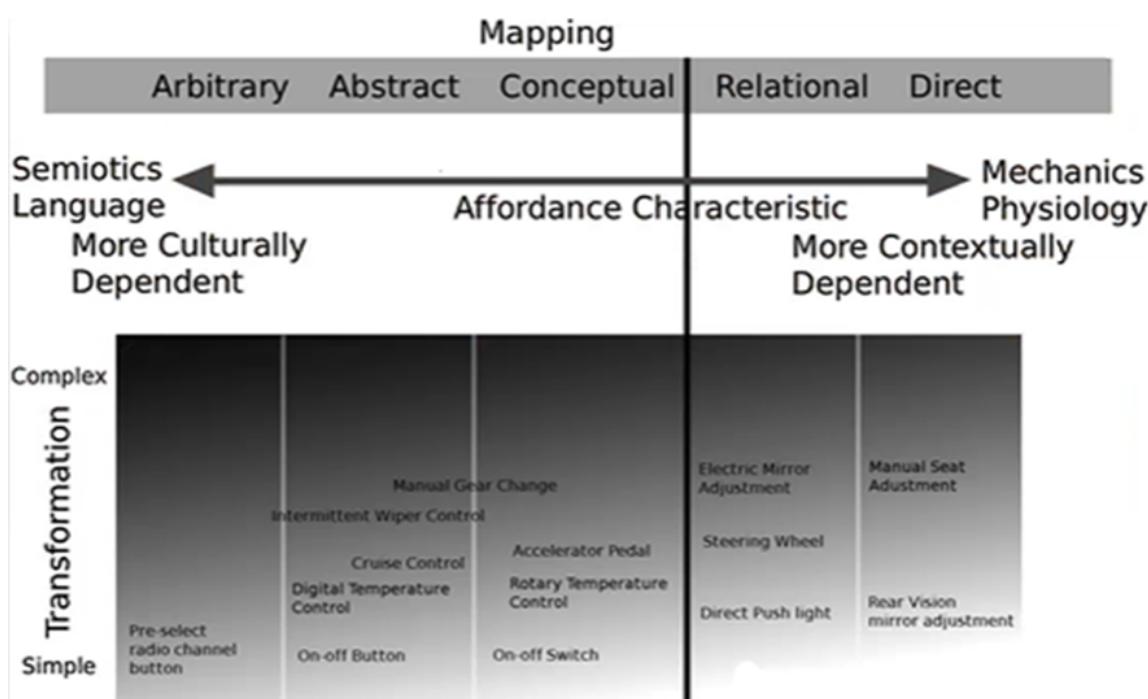


l'effetto che tale azione avrà nel mondo reale/applicativo.

Mapping arbitrario (convenzionale): una relazione del tutto immaginata dal progettista e che l'utente adotta solo per prove ed errori e poi impara a memoria, senza riferimenti culturali o indicazioni fisiche.

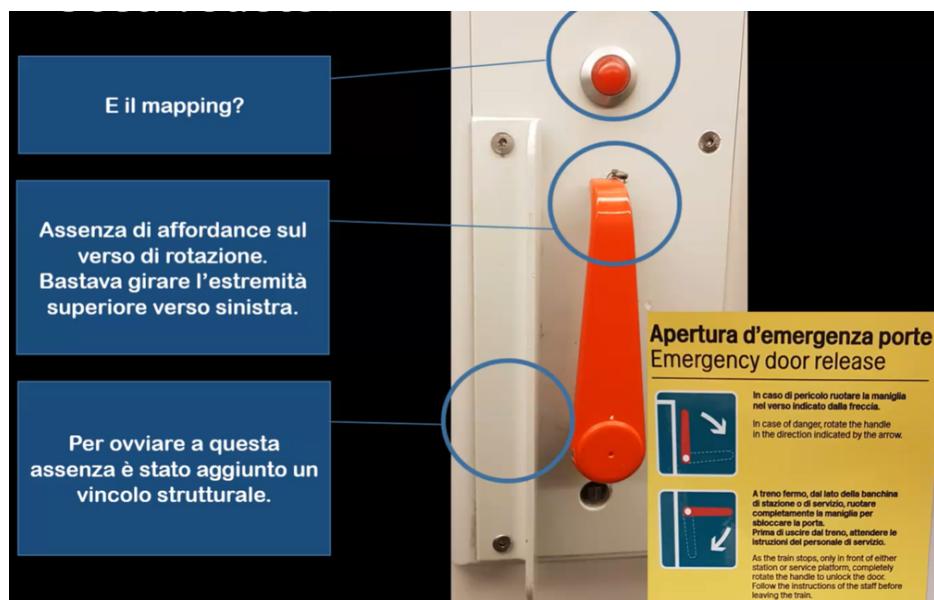
Mapping diretto (naturale): c'è una relazione fisica, visibile, spesso legata alla posizione, tra controllo (che afforda una azione) e l'effettore controllato che produce un effetto nel mondo.

Relazione tra mapping e affordance:



→ Non puoi avere mapping senza affordance, ma puoi avere affordance senza mapping.

Esempio: pedale dell'acceleratore (mapping convenzionale), freno (mapping naturale). In



entrambi i casi affordance di "schiacciare".

Nell'esempio sopra: per mapping si intende che la luce (che si accende se ha funzionato l'azione) è dello stesso colore della maniglia per evidenziare la correlazione.

Se un design dipende da etichette, è probabile che sia fallato. Le etichette sono importanti e spesso necessarie, ma l'utilizzo appropriato del mapping naturale minimizza il numero di etichette o di scritte. Quando le etichette sembrano necessarie, si potrebbe pensare a un'altra soluzione di design.

Ricapitolando: un affordance di un oggetto (tangibile o intangibile, come un'applicazione) è un "invito visuale" (testuale, simbolico o iconico) ad un certo tipo di azione, eseguita la quale l'oggetto che la esibisce abilita l'utente a raggiungere un determinato scopo.

Cosa significa progettare "rispettando" le affordance?

1. Segui le maggiori convenzioni già stabilite per immagini e azioni e adotta mapping naturale ogni volta che è possibile.
2. Se appropriate, usa le parole in aggiunta alle azioni e alla grafica.
3. Utilizza metafore riconoscibili (ad esempio utilizza il cestino per buttare i file).
4. Sii consistente e coerente nell'uso del modello concettuale dietro al design.

MAPPING DEL DISTRIBUTORE

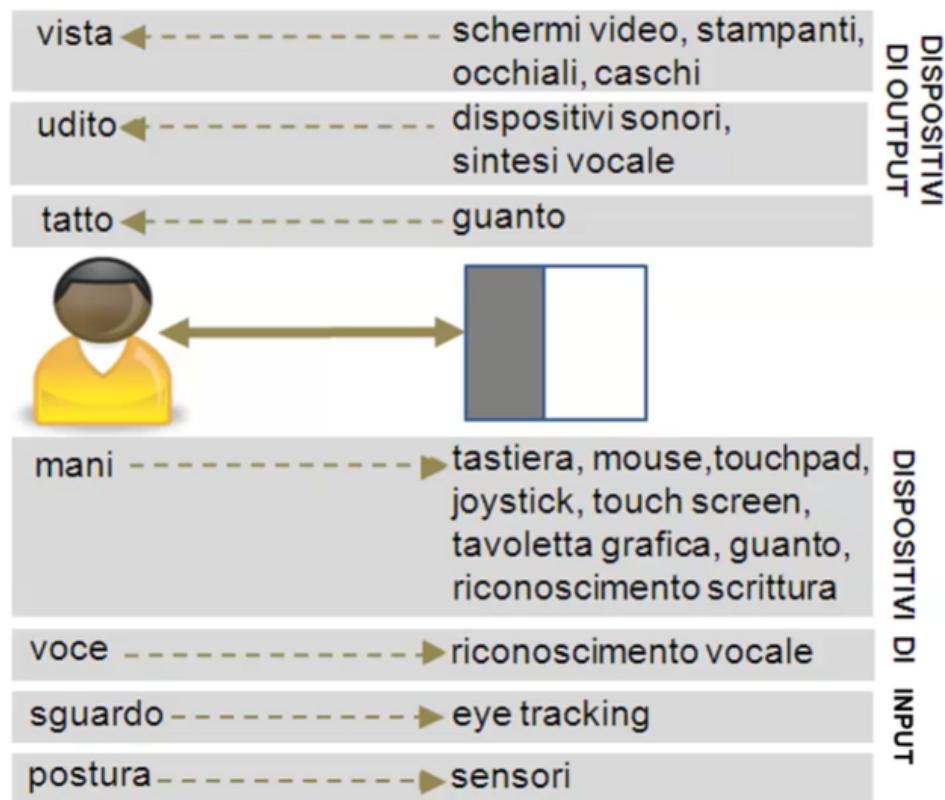
Mapping del distributore:

- Tre effettori per fare uscire l'acqua Sono premibili??
- Consistenza: perché dei dati diversi che fanno la stessa cosa?
- Badge, non è stato pensato a come aiutare a non scordarlo.. lontano dagli occhi
- C'è un manuale sul muro, poco usabile e intuibile



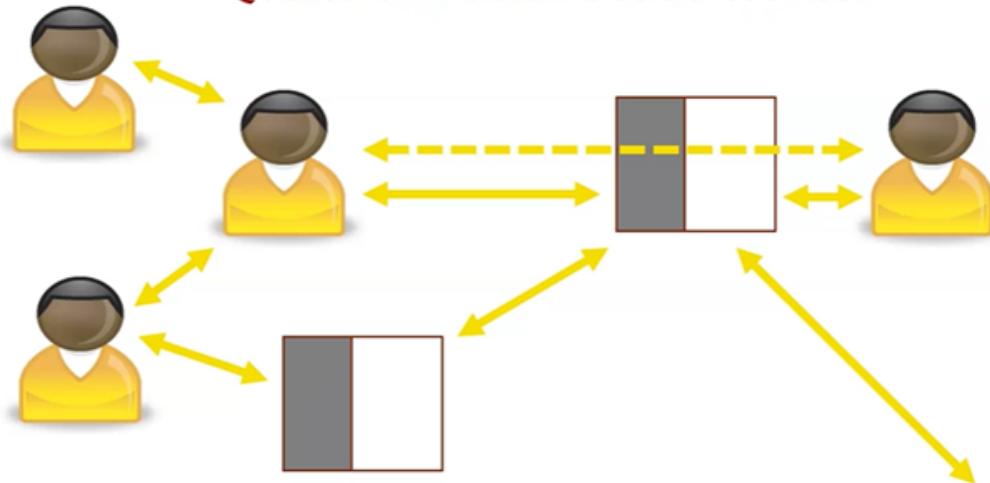
Bisogna esercitarsi, osservare in giro tutto per vedere l'usabilità!!!!

DISPOSITIVI PER L'INTERAZIONE



- [manca una definizione corta che ha dato]
- Sensore: il sistema interattivo coglie degli elementi
- Attuatore o effettore: dispositivi di output e di input

Questo è un sistema socio-tecnico!



Essere più superficiali conviene, infatti:

- Con una “buona” interfaccia l’utente può svolgere il suo compito dimenticandosi che sta usando un computer (teoria del flusso).
- Una “buona” interfaccia dà benefici in termini di più alta produttività, più basso turnover, più alto morale dello staff, minore frustrazione, stress, insoddisfazione, più alta soddisfazione data dal lavoro, che si traducono in più bassi costi operativi.
- E si riducono gli errori introdotti dalla tecnologia.

Un sistema usato in una azienda da **250 utenti**, costo orario **20 Euro**.

Certi moduli elettronici sono usati in media **60 volte** al giorno da ogni utente **230 giorni** all’anno.

Ogni compilazione del modulo potrebbe essere fatta risparmiando **3 sec** per ogni compilazione (ad es. con degli shortcut da tastiera).

$250 \text{ utenti} \times 60 \text{ comp.} \times 230 \text{ giorni} \times 3 \times (20/3600) =$
57500 Euro
(risparmiabili in un anno)

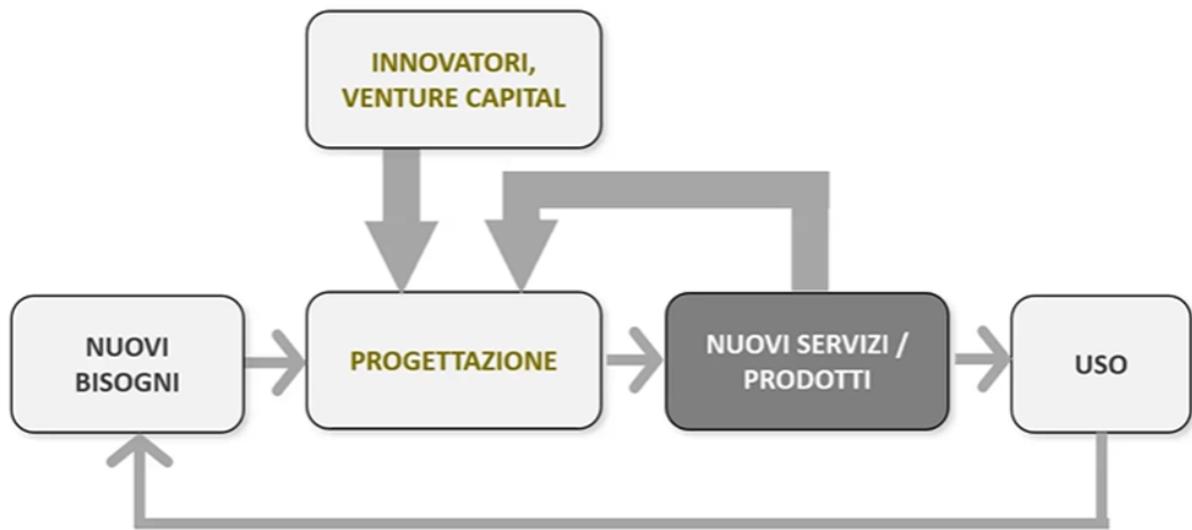
Il tempo è denaro, ottimizzare anche di poco un sistema, in contesti ampli può essere essenziale

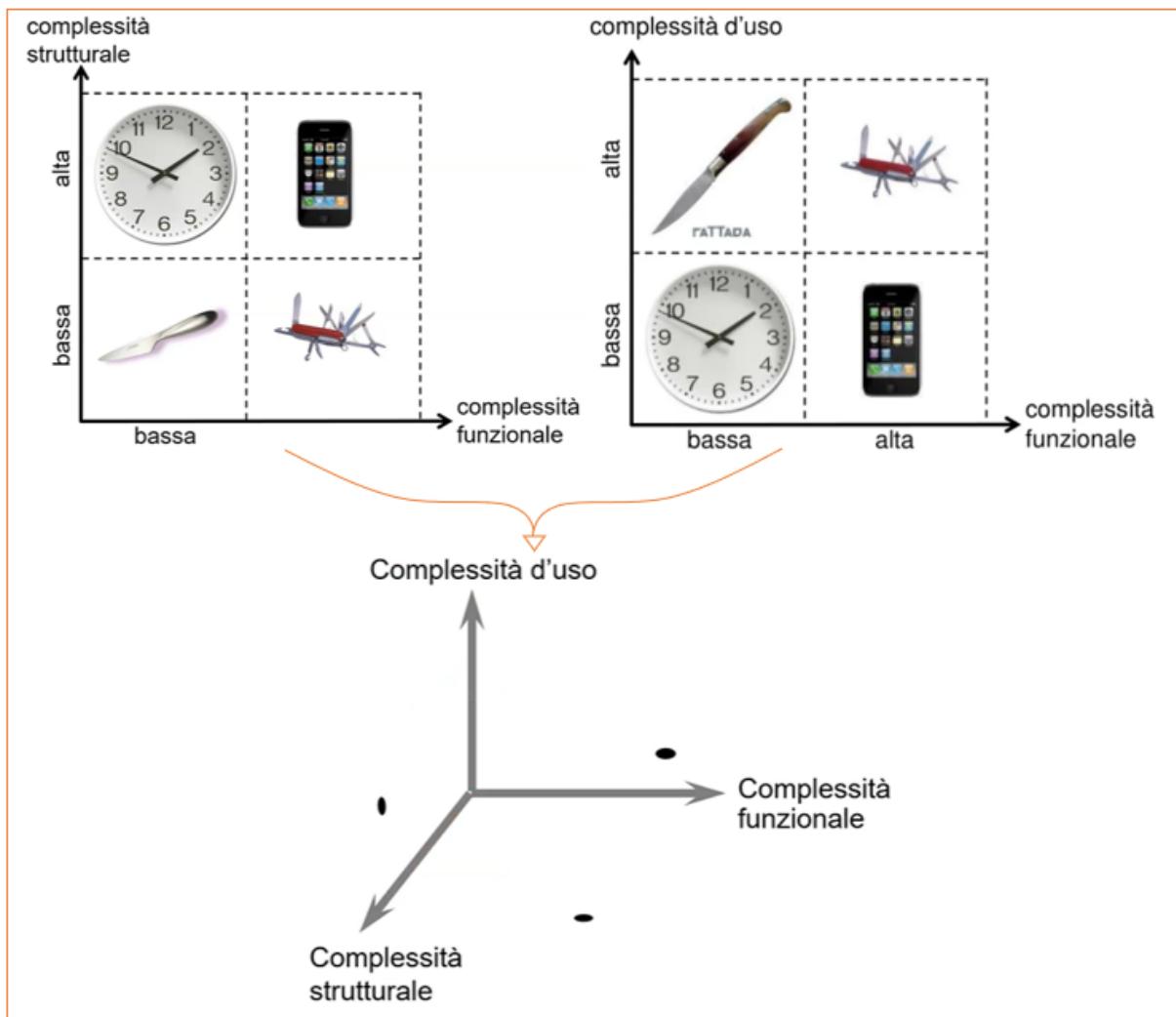
Possono esserci errori di giustapposizione: premo un bottone, ma in verità per poco se ne apre un altro

⚠️ Un buon sistema deve dire in che stato si trova il sistema

🔑 CPOE

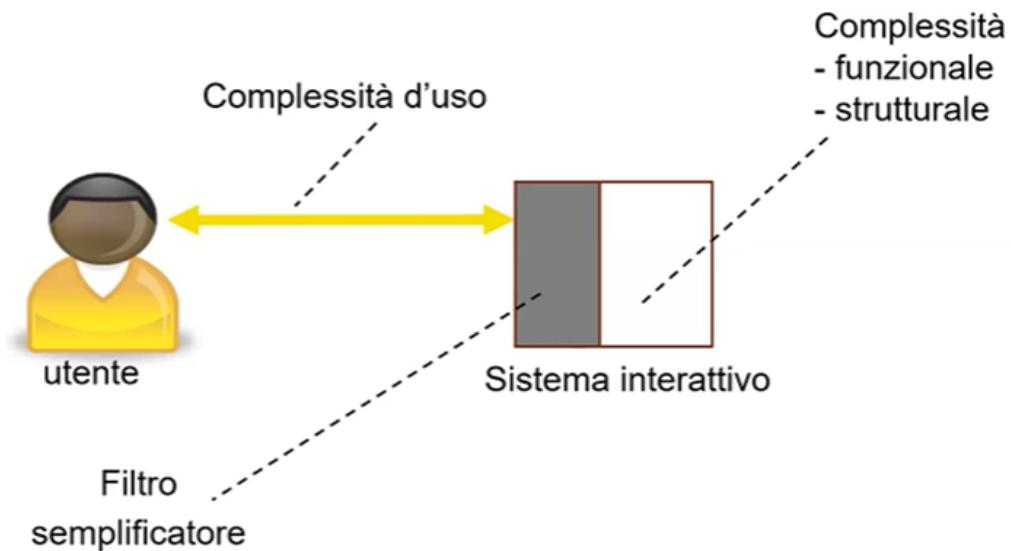
NUOVE TECNOLOGIE





Dimensioni della complessità:

Show off: voler complicare inutilmente il design di un interfaccia. Invece per efficienza, efficacia e soddisfazione per l'utente è meglio costruire un'interfaccia semplice



Riepilogo

Concetti cardine:

- Constraint, mapping, affordance, workaround
- Abbiamo capito che non esistono metodologie per progettare una buona interazione se non questa ricetta: osservare, immaginare, prototipare, valutare con gli utenti e così via
- Evitare errori e orrori attraverso l'osservazione (la toilette di Floyd, le porte di Norman) e tanto sandaflor
- E poi, a volte, basta chiedersi: ma come la vedrà l'utente, gli incasinerò la vita?
Progettazione responsabile!

21/03/2022

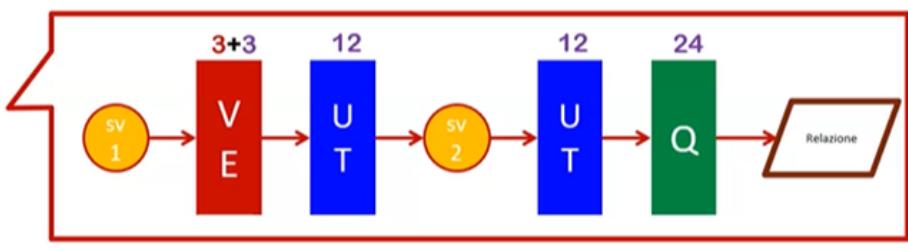
METODOLOGIA PER L'ATTIVITÀ PROGETTUALE

VALUTAZIONE USABILITÀ

CHI?  
voi utenti reali

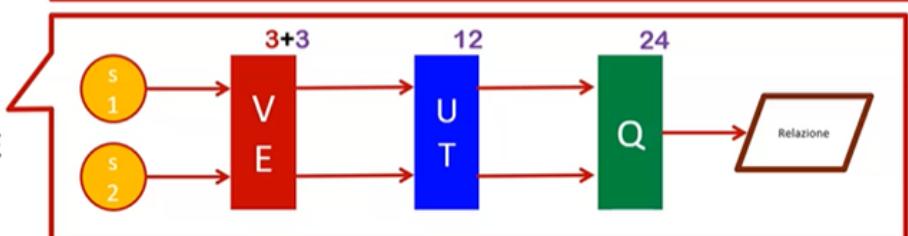
COSA? **COME?**

1)
CONFRONTO
LONGITUDINALE



oppure

2) CONFRONTO
TRASVERSALE



 **SISTEMA ALLA VERSIONE ***

 **SISTEMA X**

 **VALUTAZIONE EURISTICA**

 **USER TEST**

 **QUESTIONARIO**

- Questionario: almeno 24 utenti più o meno rappresentativi della popolazione di riferimento a cui somministrare un questionario di usabilità. Due elementi importanti per avere un buon gruppo di utenti sono randomicità e rappresentatività.
- User test: 12 utenti
- Valutazione euristica: 3 utenti esperti di usabilità (noi) + 3 utenti esperti di dominio

 Se si usano gli stessi utenti per i vari test i campioni devono essere accoppiati e dipendenti

Una metodologia povera ma onesta:

- Affina sensibilità per il brutto e il cattivo
- Progetta in equilibrio tra familiarità e originalità
- Sappi che parli all'utente, anche di te stesso
- Non guardarti l'ombelico ma coinvolgi gli utenti
- Trova i problemi, prioritizzali e risolvili
- Valuta usabilità in termini di efficienza, efficacia e soddisfazione
- Migliora il tuo artefatto
- Torna dagli utenti

→ Ogni volta che ti sembra di avere una buona idea chiediti: cosa ne penseranno gli utenti?

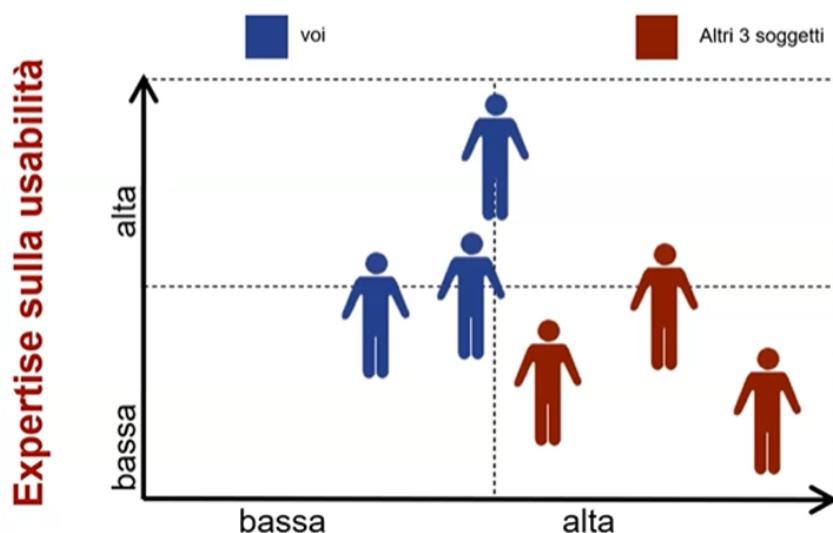
Chiediglielo!

Valutazione qualitativa

Si coinvolgono utenti esperti del task e/o esperti di usabilità per valutare sistemi interattivi, a qualsiasi livello di prototipazione e maturità, alla luce di principi di buona progettazione (si procede secondo un metodo, cioè sistematicamente e riportando i risultati in modo strutturato e completo).

Il fine non è una misura, ma farsi un'idea sufficientemente completa dei possibili problemi di usabilità che un sistema può presentare se usato veramente.

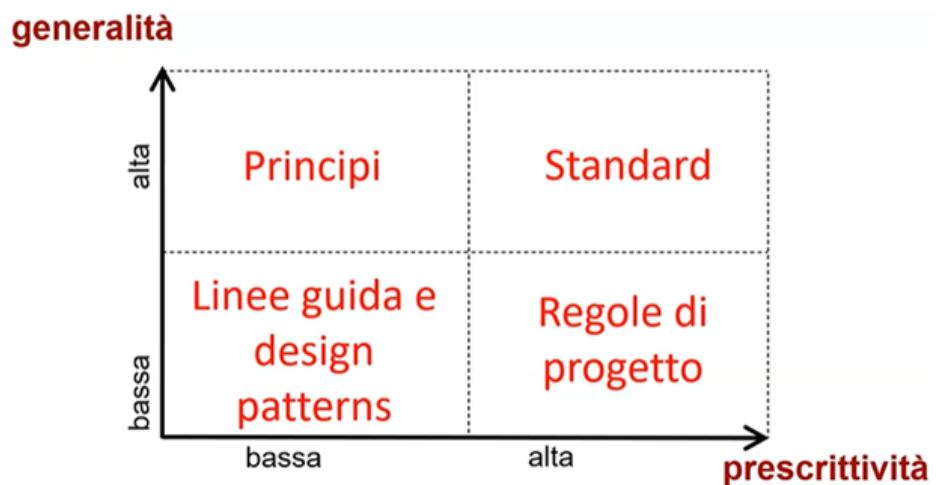
L'obiettivo nella valutazione qualitativa euristica il risultato voluto è un insieme di problemi di usabilità, pesato per gravità e collegato ad un principio di usabilità negato.



⚠ La valutazione è qualitativa se riguarda l'opinione (interpretazione) di un esperto/utente, e questa non viene trattata statisticamente.

- **Principi:** indicazioni generali per la progettazione di interfacce utente usabili, basate su evidenza

scientifica o sul generale consenso.
Derivano dalla conoscenza degli aspetti psicologici, computazionali e sociali e sono indipendenti dalla tecnologia. Sono espressi spesso in forma molto generale



- **Standard:** insieme di regole da applicare nel progetto di una classe di sistemi. Possono essere vincolanti. Sono di norma emesse da un Ente di standardizzazione (es ISO). La conformità allo standard deve essere valutabile in modo preciso.
- **Design pattern:** insieme di raccomandazioni per il progetto dell'interfaccia utente per una particolare classe di sistemi, espresse in modo generale ma meno astratte dei principi, con esempi e motivazioni. Non sono mai vincolanti, sta al progettista decidere sulla opportunità di applicarle caso per caso.
- **Regole di progetto:** insieme di regole o specifiche da applicare nel progetto di un particolare sistema. Sono vincolanti all'interno del progetto, ad esempio perché dettate da precisi requisiti del cliente (committente) o per motivi di budget e vincoli progettuali.

Una **euristica** è un insieme di concetti, riferimenti e, soprattutto, strategie che si sono rivelate particolarmente adatte a risolvere un determinato problema, nel nostro caso la progettazione di sistemi interattivi usabili.

La valutazione euristica è la valutazione (di usabilità) svolta alla luce di un determinato insieme di euristiche (ben definite, possibilmente ben note, prese a riferimento) per identificare soluzioni di design che o si conformano o violano una o più euristiche di tale insieme.

→ Preferiamo le euristiche ai design pattern perché si vuole valutare l'usabilità senza creare vincoli eccessivi contro l'originalità e l'innovatività.

Diversi livelli di autorevolezza (o forza dell'evidenza) delle euristiche:

- A. Completamente supportate da risultati di ricerca e dati empirici
- B. Basate su pratica generalmente accettata (in modo documentato)
- C. Non ben documentate, ma supportate dall'opinione di professionisti esperti
- D. Opinione individuale (del progettista)

Principi di Nielsen: a cavallo tra C e B.

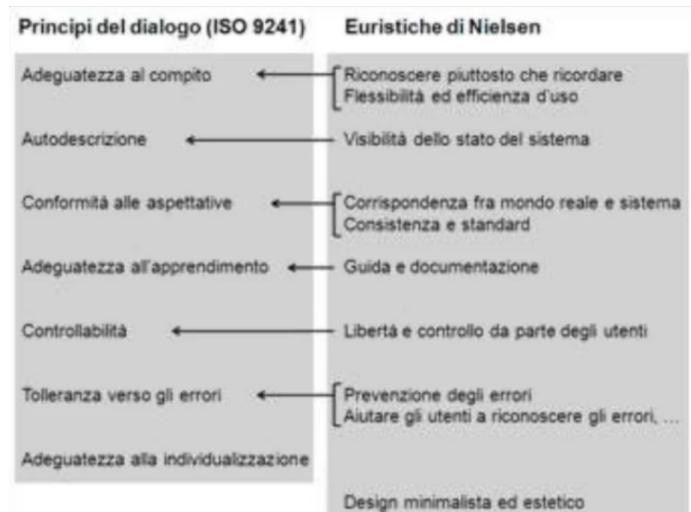
→ Le linee guida e i design pattern che apprezzate e trovate efficaci li applicate. I principi (della buona progettazione di sistemi usabili) li seguite.

EURISTICHE

IXD checklist: <https://idxchecklist.com/>

7 principi del dialogo uomo macchina secondo lo standard ISO (non a memoria):

1. Adequatezza del compito
2. Autodescrizione
3. Conformità alle aspettative dell'utente



4. Adeguatezza all'apprendimento
5. Controllabilità
6. Tolleranza verso gli errori
7. Adeguatezza alla personalizzazione

Cercare diverse fonti, una che consigliata è: i principi euristici di Jacob Nielsen

Basati su una analisi fattoriale di 249 problemi di usabilità, con cui si è derivato l'insieme minimo che massimizza la potenza esplicativa

www.nngroup.com/articles/ten-usability-heuristics/

22/03/2022

PRINCIPI EURISTICI DI JACOB NIELSEN (non a memoria)

Nielsen J. (1994b) Heuristic evaluation. In Nielsen J., and Mack R.L. (Eds.), Usability Inspection Methods, John Wiley & Sons, New York NY.

Euristiche di usabilità per il design di interfacce utente (usability heuristics for user interface design)

[Visibilità dello stato di sistema]

- Allineamento tra sistema e mondo reali
- Adeguato controllo e libertà all'utente
- Garantire consistenza e aderenza agli standard
- Prevenire l'errore
- Fare affidamento sull'intuizione piuttosto che sulla memoria
- Flessibilità ed efficienza
- Design minimalista e estetica

- Aiutare l'utente a riconoscere, fare una diagnosi e ripristinare uno stato coerente in caso di errori
- Curare la documentazione e il supporto

Mezzi per associare il problema all'euristica: a ogni problema è associato uno o più euristiche (non più di 4)??

Principi euristici di Nielsen **rielaborati da Mussio**:

1. Percezione
 - a. Far vedere lo stato del sistema (feedback)
 - b. Adeguare il sistema al mondo reale (parlare il linguaggio dell'utente)
 - c. Controllo dell'utente e libertà (uscite indicate chiaramente)
2. Cognizione
 - a. Assicurare consistenza (nell'applicazione, sistema, ambiente)
 - b. Riconoscimento piuttosto che uso della memoria dell'utente
 - c. Assicurare flessibilità ed efficienza d'uso (acceleratori)
 - d. Visualizzare tutte e sole le informazioni necessarie (estetica e design minimalista)
3. Errori
 - a. Prevenire gli errori
 - b. Permettere all'utente di correggere gli errori e non solo di rilevarli
 - c. Help e documentazione

METODO PER UNA VALUTAZIONE EURISTICA DI USABILITÀ

Tecnica sistematica di ispezione predittiva (per fare prevenzione, cercando di capire cosa potrà capitare una volta rilasciato il sistema).

Un gruppo di valutatori esamina il sistema e identifica eventuali problemi di usabilità rispetto ad un insieme prefissato di principi. Può essere:

- Olistica o task-oriented
 - Se è guidata da un task (o uno scenario) è necessario anche specificare tale task (o scenario). Potete riutilizzare gli scenari\task delle valutazione quantitativa

Obiettivo e compito: un compito (task) porta alla realizzazione di un obiettivo, con una procedura, un algoritmo. Un obiettivo rappresenta solo il risultato del compito senza nessuna informazione su come raggiungerlo.
- Con o senza osservatori

=> Risultato: elenco di problemi di usabilità, associati ad una o più euristiche, pesati per gravità

Olistica: non guidata, basata su una libera navigazione ed esplorazione del sistema, per viste di insieme. Tipico dei sistemi “walk-up-and-use” (sistemi auto-esplicativi).

Task-oriented: se è guidata da un task (o uno scenario) è necessario anche specificare tale task (o scenario). Si possono riutilizzare i task per la valutazione quantitativa.

Osservatori: siamo noi che presidiamo la sessione, ed eventualmente qualcuno che “da fuori” prende appunti.

Compito: passi da eseguire per raggiungere un obiettivo

COME PROCEDERE

1. Scegliere una o due applicazioni
2. Concepite 2 o 3 task (per ognuna)
3. Ciascuno valuta l'applicazione (valutazione esperta olistica) con uno dei form a portata di mano
4. Ciascuno scrive un elenco di problemi di usabilità, riconducendoli alla valutazione di una o più euristiche

5. Presentare l'applicazione a 3 utenti esperti di dominio, dandogli 3/4 minuti per esplorare l'applicazione in maniera olistica. Nel frattempo, registrarlo (bisogna fargli firmare un modulo) e dirgli di pensare ad alta voce mentre usa l'applicazione (think-aloud protocol analysis). Questo può servire ad esempio per valutare quante espressioni positive/negative ha nella prima e nella seconda applicazione. Qualcuno dovrebbe prendere appunti, mentre un altro no (a meno che il soggetto si senta in soggezione). Si potrebbe prima creare una presentazione da mostrare a tutti gli utenti per poter avere una valutazione uniforme.
 - Aggotto sopraccigli → negativoContare le varie espressioni
 - Contare tutte le espressioni che le persone notano in modo differente e fare una media.
 - Non stare a segnare tutto mentre ha l'app in mano, riguardare video per ricontrizzare tutti i dati. Segnare l'ora di inizio, e se c'è qualcosa di particolare segnare l'ora per ritrovare il punto nel video
 - Suggerimento: chiarire bene alla persona, che si potrebbe trovare in soggezione, che chi è sotto esame è il sistema, e non è lei. A questo fine può essere utile fare una piccola presentazione con anche questo aspetto del test.
6. Presentare all'utente un task di quelli definiti al passo 2. Meglio farlo con una presentazione/filmato, come prima, spiegando sia l'obiettivo sia i macro-passi.
 - Segnare quando l'utente ha sbagliato e quando ha eseguito il compito ma è stato aiutato
7. Fare eseguire il task all'utente (non più di 3/4 minuti – valutazione utente task-based) e continuare con la protocol analysis. Se ha bisogno di aiuto bisogna dargli suggerimenti, segnandosi di averli dati.
8. Dalla sessione di uso di ciascun utente estrarre collaborativamente un elenco di problemi di usabilità.
9. Mettere insieme le relazioni del punto 4 e quella prodotta dalla sessione con gli utenti, decidendo cosa togliere e cosa lasciare.

10. Predisporre un questionario con il quale prioritizzare voi e far prioritizzare agli utenti i problemi trovati.

11. Associare ad ogni problema una o più euristiche, indirizzare le prime n.

TASK

Concepire task realistici:

- User goal: esplorare offerte dei prodotti e acquistarne uno
- Task concepito male: acquista un paio di scarpe da corsa della Nike arancioni (l'utente normalmente non ha fattori esterni che gli impongono il colore e o la marca delle scarpe, inoltre è un obiettivo troppo specifico, l'utente non è un automa, gli deve essere lasciata una certa libertà)
- Task migliore: compra un paio di scarpe sotto i 40€

Concepire task eseguibili:

- User goal: cercare un film e i suoi orari
- Task concepito male: vuoi vedere un film di domenica pomeriggio, vai sul sito e dimmi dove cliccheresti come prima cosa
- Task migliore: usa il sito per trovare un film che vorresti vedere domenica pomeriggio

Evitare suggerimenti e descrizioni dei singoli passi:

- User goal: consultare i voti online
- Task concepito male: vuoi vedere i voti del tuo esame, vai sul sito del corso, loggati e dimmi dove cliccheresti per vedere il tuo voto
- Task migliore: vuoi andare a vedere il tuo voto dell'esame

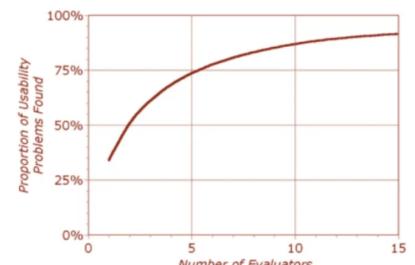
Metodo ibrido per fare entrambe le valutazioni euristiche, consigliato:

Per 5 minuti far esplorare l'applicazione liberamente

Proporre un task tra i più importanti, che possono essere svolti con l'applicazione o sistema

OSSERVATORI

Quanti valutatori coinvolgere (utenti a cui sottoporre l'applicazione)? Ne basterebbero 5 per trovare i $\frac{3}{4}$ di tutti i problemi del sistema. Ogni utente ha la probabilità del 30% di individuare un problema.



Numero totale di problemi (T) $T = (x * y) / z$

Prodotto del numero totale di problemi trovati da due persone (indipendenti l'una dall'altra), diviso il numero di errori trovati da entrambi.

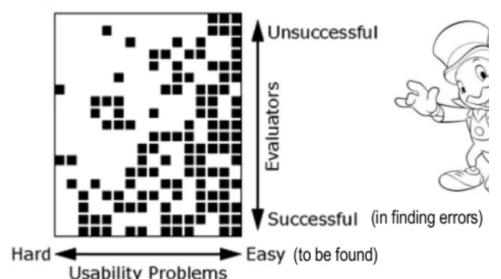
MATRICE PROBLEMI/VALUTATORI (IMPORTANTE DA FARE)

Nera quando il valutatore i -esimo trova il problema

k -esimo. In questo caso la matrice è ordinata sia per valutatore che per problemi.

Valutatori hanno un ID

Cartella condivisa con tutto: video dei valutatori...



Autopromuovere il proprio lavoro senza indugio nel progetto:

- 90% problemi trovati → siamo bravi perché abbiamo trovato tanti problemi
- 50% problemi trovati → siamo bravi perché abbiamo individuato tutti i problemi principali

COME FARE LA VALUTAZIONE EURISTICA?

Come prioritizzarli?

L'elenco così fatto deve essere proposto ai valutatori (almeno 20 in teoria, o anche solo ai 3 più coinvolti) (e possibilmente al campione di utenti) per la valutazione della severità da 0 a 4, cioè per ciascun problema, i partecipanti alla valutazione della severità devono assegnare un grado di severità (Nielsen suggerisce la seguente scala):

0. Non sono d'accordo che questo sia un problema di usabilità
1. È solo un problema “cosmetico” (accessorio): non deve essere risolto, a meno che nel progetto non sia disponibile del tempo extra
2. Problema secondario: alla sua soluzione risoluzione bisognerebbe dare bassa priorità
3. Problema rilevante: è importante risolverlo, bisognerebbe dare alta priorità alla sua risoluzione
4. Catastrofe di usabilità: è imperativo risolverlo prima che il prodotto possa essere rilasciato

Sulla base di questa valutazione collettiva si prioritizza l'elenco, individuando quelli di fascia prioritaria più alta

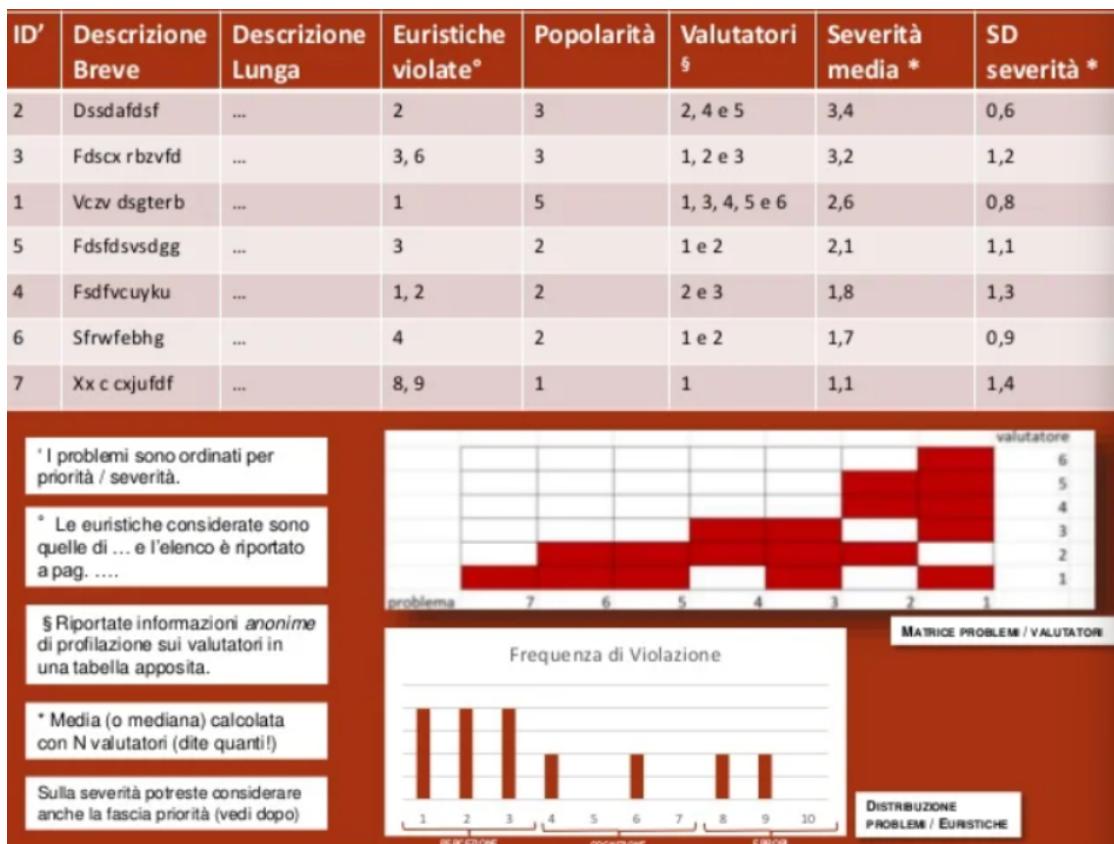
Il rapporto di valutazione deve riportare anche il numero di persone coinvolte in questa valutazione della severità

Come valutatori di severità notate (e fate notare agli altri partecipanti) che il punteggio dovrebbe essere scelto tenendo a mente tre fattori:

1. La frequenza o probabilità di occorrenza: è comune o raro?
2. L'impatto del problema SE esso occorre: è facile da superare? Quali sono le conseguenze all'atto pratico?

3. La persistenza: è un problema in cui lo stesso utente può incappare più volte anche essendone consapevole e a conoscenza, oppure è di fatto one-time?

In contesti professionali la severità è sostituita dal rischio = impatto (severità) x probabilità (frequenza)



Ad ogni problema bisogna assegnare (oltre che fare i 2 grafici):

- ID: utile per comunicare con altri team in ambiente lavorativo, ma anche internamente al team
- Descrizione breve
- Descrizione lunga
- Euristiche violate (riferirsi al numero associato all'euristica violata)
- Popolarità: numero di valutatori che l'hanno trovata

- Valutatori: quali valutatori l'hanno visto (anche a loro assegnare un ID, un'età, un "ruolo", ad esempio esperto di dominio, esperto di valutazione, ...)
- Severità media (o mediana): media dei giudizi di severità dati dagli utenti (in realtà lui stesso dice che sarebbe meglio la mediana dopo avere ordinato le severità in ordine crescente) 
- SD severità (o range interquartile): deviazione standard
- Fascia di priorità: ordinati all'interno delle fasce per severità, all'interno ordinati per popolarità

Ai problemi più gravi mostrare degli screenshots relativi al problema e pensare a una possibile soluzione (scriverla solo se si è sicuri).

 Moda: valore (o insieme dei valori) ripetuto più volte

Media: somma dei valori / numero dei valori

→ Deviazione Standard

Mediana: numero alla posizione pari al numero dei valori / 2 (se non è intera fare la media dei due valori a destra e sinistra)

→ Range Interquartile

(Moda-Media-Mediana da sapere)

QUINDI LA VALUTAZIONE EURISTICA

La valutazione euristica è una tecnica di approccio più generale detto di “discount usability engineering” (veloce, efficiente, efficace). Molto utile, identifica i veri problemi e li prioritizza.

Pro:

- Più veloce e meno costosa
- Risultati già in forma interpretata

Contro:

- Meno accurata: può non individuare dei problemi o produrre “falsi positivi”
- Non tiene conto degli utenti veri e dei loro compiti situati

RISORSE NECESSARIE

- UN sistema funzionante, oppure un prototipo oppure un mock-up
- n valutatori
- un osservatore, quando coinvolge utenti esperti
- Tempo:
 - In ambito professionale una sessione individuale di osservazione dura tra i 30 e i 90 minuti (a seconda dei task). Nel nostro caso 5-15 minuti.
 - Ogni valutazione dovrebbe riesaminare l’interfaccia più volte

Workflow proposto da Nielsen per impostare una valutazione euristica (sulla colonna di destra ci sono le ore uomo):

Having outside evaluation expert learn about the domain and scenario	8
Finding and scheduling evaluators, 1.8 hours + 0.2 hours per evaluator	4
Preparing the briefing	3
Preparing scenario for the evaluators	2
Briefing, 1 system expert, 1 evaluation expert, 11 evaluators @ 1.5 hours	19.5
Preparing the prototype (software and its hardware platform) for the evaluation	5
Actual evaluation, 11 evaluators @ 1 hour	11
Observing the evaluation sessions, 2 observers @ 11 hours	22
Debriefing, 3 evaluators, 3 developers, 1 evaluation expert @ 1 hour	7
Writing list of usability problems based on notes from evaluation sessions	2
Writing problem descriptions for use in severity-rating questionnaire	6
Severity rating, 11 evaluators @ 0.5 hours	5.5
Analyzing severity ratings	2
Total	105

ERRORI COMUNI

Tipicamente errori di distrazione o di organizzazione.

Condurre la valutazione aggregando i problemi e non analizzandoli uno ad uno

I problemi di usabilità vanno descritti uno ad uno

Granularità di un solo problema atomico: un utente ha trovato un grande problema, ma altri utenti trovano dei problemi che sono compresi, stanno nel problemone trovato prima.....

Nella discussione di un problema proporre soluzioni

Questo lavoro va fatto nella fase dopo la valutazione euristica

Considerare un principio alla volta e indicare i problemi che violano tale principio

In realtà ogni problema può violare più principi: si parte dai problemi

Ogni problema può violare più euristiche!

Mischiare problemi e soluzioni

L'output della valutazione è un elenco di problemi prioritizzati

PRIORITIZZAZIONE

Per ogni problema identificato dal gruppo, far valutare indipendentemente la severità, e coinvolgere anche una decina di colleghi dell'azienda e una decina di potenziali utenti (non colleghi dell'azienda)

In poche parole, cercare di ottenere 20/30 (per noi bastano 6) valutazioni ordinali di severità per ciascun problema.

Il fine è identificare i problemi più urgenti da indirizzare con le risorse a disposizione.

1. Estrarre dal sistema di rilevazione (il questionario online) la tabella delle valutazioni (i problemi indicati sulle colonne, le valutazioni di severità nelle righe, una per singolo valutatore)
2. Per ciascuna riga considerare la classifica di severità adottando la strategia di competizione standard (aka 1224), costruendo una nuova tabella delle posizioni (a partire da quella delle valutazioni)

Il (o i) problemi in testa sono quelli che hanno ottenuto il punteggio più alto (indipendente da quale sia)

prob1	prob2	prob3
3	3	1
4	5	3

→

prob1	prob2	prob3
1	1	3
2	1	3

3. Considerate 2 fasce: quella dei problemi a priorità più alta considera le prime n posizioni, con n circa 20% del numero totale dei problemi, t e poi la fascia delle restanti t-n posizioni

→ Solitamente nella prima fascia si considerano le prime 3/5 posizioni, cioè il podio o una specie di podio allargato. È ragionevole pensare che questi problemi siano responsabili di circa l'80% di tutti i guai

prob1	prob2	prob3
3	3	1
4	5	3

→

prob1	prob2	prob3
1	1	3
2	1	3

→

Prob1	Prob2	Prob3
2 v 0	2 v 0	0 v 2

Matrice della valutazioni dei valutatori dei problemi: Matrice Del podio dei problemi
Matrice di ? (num volte in prima fasci, num volte in seconda fascia)

4. Contate, per ciascun problema, il numero di volte che esso è in prima fascia (p) e il numero di volte che è in seconda fascia (s). In questo caso la prima fascia è formata dalle prime due posizioni, quindi visualizziamo per ciascun problema p vs s.

5. Assegnate il problema ad una fascia in base al confronto tra p ed s

Se $p > s$, il problema è in prima fascia, altrimenti è in seconda fascia

6. Eseguite un test binomiale per capire se la differenza tra p ed s è dovuta al caso. In base a questo test stabilite 3 fasce in base al P value del test:
- Fascia alta con priorità significativa (A)
 - Fascia intermedia di non significatività (B): non si riesce a raggiungere una classificazione significativa statisticamente
 - Fascia bassa con priorità significativa (C)

In questo caso A indica la prima fascia con significatività statistica; C la seconda con significatività statistica; e B la fascia intermedia in cui mettiamo tutti i problemi per cui non riusciamo a raggiungere una classificazione significativa statisticamente.

The diagram illustrates a process flow. On the left, there is a table with three columns labeled Prob1, Prob2, and Prob3. The first row contains the values 20 v 6, 23 v 4, and 2 v 25 respectively. An orange arrow points from this table to the right, indicating a transformation. On the right, there is another table with the same three columns. The first row now contains the values A, A, and C respectively, representing the prioritized categories.

Prob1	Prob2	Prob3
20 v 6	23 v 4	2 v 25

→

Prob1	Prob2	Prob3
A	A	C

25/03/2022

- Tutti i problemi nella fascia di alta priorità significativa dovranno essere risolti prioritariamente e scegliete alcuni problemi dalla fascia di non significatività.
- Nella tabella dei problemi aggiungete due colonne: In una indicate la fascia di priorità, nell'altra la mediana/media dell'indice di severità.
 - Per chiarezza ordinate i problemi mettendo prima quelli che sono in prima fascia e risultati anche significativamente severi. Riportate come nota il numero delle persone coinvolte nella prioritizzazione.

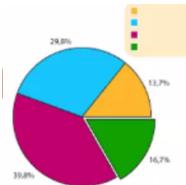
ID'	Descrizione Breve	Descrizione Lunga	Euristiche violate°	Popolarità	Valutatori §	Priorità	Severità media (SD)
2	Dssdafdsf	...	2	3	2, 4 e 5	A	3,4 (0,6)
3	Fdscx rbzvfd	...	3, 6	3	1, 2 e 3	A	3,2 (1,2)
1	Vczv dsgterb	...	1	5	1, 3, 4, 5 e 6	B	2,6 (0,8)
5	Fdsfdsvsdgg	...	3	2	1 e 2	B	2,1 (1,1)
4	Fsdfvcuyku	...	1, 2	2	2 e 3	B	1,8 (1,3)
6	Sfrwfebhg	...	4	2	1 e 2	C	1,7 (0,9)
7	Xx c cxjufdf	...	8, 9	1	1	C	1,4 (1,1)

Ricordare di accompagnare al numero della media o mediana un numero che indica la dispersione.

IL RAPPORTO FINALE DELLA VALUTAZIONE EURISTICA

Il rapporto finale della valutazione euristica contiene:

- Descrizione del design sperimentale: cosa abbiamo valutato e perché, quante persone abbiamo coinvolto, quali task, aerogramma profilazione, diagramma cartesiano delle expertise...;
- L'elenco finale dei problemi prioritizzati: tabella dei problemi, screenshot dei problemi più gravi, matrice problemi/valutatori, distribuzione problemi/euristiche;
- Considerazioni finali legate al numero totale stimato di problemi, alla distribuzione dei problemi, alle citazioni notevoli (sia in positivo che in negativo) dalle sessioni olistiche ed esperte...



NB: importante la prioritizzazione dei problemi in tre fasce (priorità più alta, priorità intermedia, priorità più bassa).

“chi ha poco da dire scrive tanto” cit. Cabitza, “per questo lui preferisce le immagini e i grafici che sono anche più rappresentativi”

Le n schede di analisi da cui è stato redatto l'elenco finale come allegato/appendice.

Consiglia di fare uno slideware, non un documento scritto.

Nella cartella di progetto mettete anche le registrazioni e i moduli consenso.

La parte statistica la approfondiremo in seguito, con degli strumenti online fatti apposta (se uno strumento chiede la probabilità aspettata, noi dobbiamo inserire 0.5 cioè che la probabilità che un problema sia ad alta priorità è la stessa che il problema non lo sia (tanto poi vedremo tutto meglio)).

EURISTICHE

1. Far vedere lo stato del sistema (feedback)

Il sistema deve informare continuamente l'utente di ciò che sta facendo e come sta interpretando gli input dell'utente

L'utente deve poter capire l'effetto dell'azione passata sullo stato presente e individuare le possibili nuove azioni

Fornire indicazioni sui malfunzionamenti

Esprimere le informazioni in maniera concreta e non generale (es. NO copia file, Sì quali file da dove a dove)

Il feedback può avere differenti gradi di persistenza

Il feedbec: mostrare che c'è un'attività in corso

- Se tempo di risposta > 3ack è importante nel caso di operazioni che richiedono una lunga elaborazione...

Esempi:

Mostrare il “tempo”

- Se tempo di risposta < 0.1 sec: non serve feedback, ma mostrare semplicemente l'output
- Se tempo di risposta 1/3 s sec: mostrare la progressione del lavoro

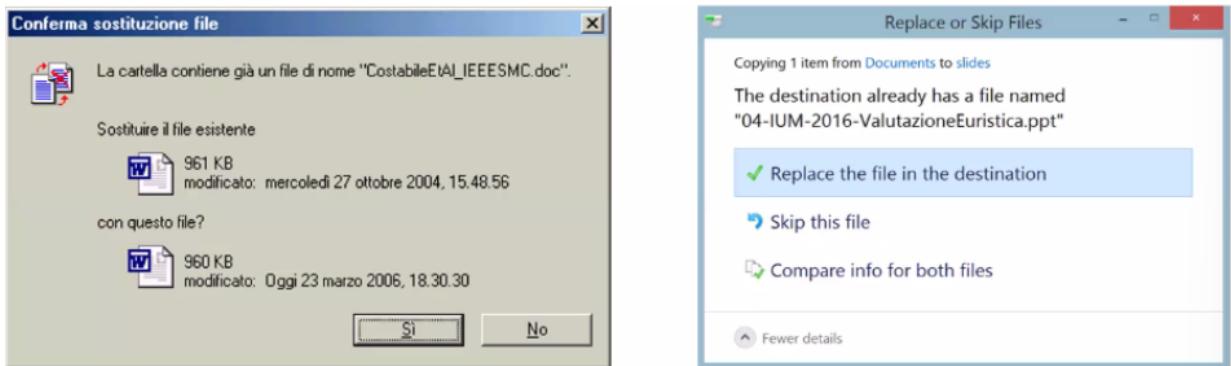


È comune che le barre di questo tipo si fermano alla fine, questo è dato da una predizione della macchina che non è corretta. Un trucco (effettivamente utilizzato) per evitare il blocco al 100% è fermare la barra al 99% o al 95% in modo che l'utente capisce che effettivamente c'è ancora del lavoro in corso.

Mostrare i “cambiamenti”

Copia di file con lo stesso nome su un altro disco:

- Feedback peggiore: dice che si sta sovrascrivendo un file senza dire quale
- Feedback migliore: dire quale file si sta sovrascrivendo indicando altri dati (es. data ultima modifica)



Feedback con diversi gradi di persistenza

- Persistenza breve: un messaggio che la stampante ha finito la carta, un “tooltip” che appare spostandosi su un pulsante
- Persistenza media: sta sullo schermo finché l’utente esplicitamente non lo riconosce (es. un messaggio che dice che una stampa è stata re-diretta su un’altra stampante per indisponibilità di quella scelta)
- Persistenza alta: rimane una parte permanente sullo schermo (es. stato della batteria nei portatili quando non collegati alla corrente elettrica)

2. Adeguare il sistema al mondo reale (parlare il linguaggio dell’utente)

Permettere all’utente di sfruttare la sua esperienza nell’interagire con il sistema

Familiarità: adotta le notazioni di utente, convenzioni sui nomi

Proiezione mondo reale → metafore familiari (richiede di capire l’utente e il contesto)

I dialoghi, se possibile, devono essere nella lingua nativa dell’utente, così come le icone devono rispettare la cultura (affordance percepita)

Esempi:

- **Coerenza linguistica** (tutti gli elementi devono essere scritti nella stessa lingua)
- Le caselle della posta a Milano non sono mai come quelle nella immagine, sopra un paletto, quindi è più raccomandabile la icona della posta a destra.

3. Controllo dell'utente e libertà

Gli utenti non si devono sentire “intrappolati” dal computer

Devono sentire che sono loro che **controllano il dialogo** (es. usare le finestre modali solo quando servono)

Offrire sempre all'utente un modo semplice per uscire dalla situazione corrente → **chiare vie di uscita**

Offrire sempre delle possibilità di “**undo**”

Vie di uscita e uno permettono all'utente di imparare esplorando il sistema

Permettere di **interrompere o cancellare operazioni** che durano troppo (limite dei 10 secondi)

Chiare vie di uscita



Undo (annullamento di ciò che si è fatto)

L'utente deve sempre, per quanto possibile, poter ripristinare lo stato precedente (e così via, all'indietro, "senza limitazioni")

- Non sempre si individuano gli errori immediatamente
- Il numero di funzioni di undo possibili deve essere pari a quello dei comandi presenti

Alcuni comandi sono molto difficili da recuperare

Es. nelle mail dopo la scelta di invio per l'invio della posta elettronica, l'utente non può recuperare lo stato precedente a tale comando (retract)

I comandi tempo-dipendenti, in particolare, hanno side-effects (effetti collaterali, irreversibili): non si può tornare indietro nel tempo

Flip undo

L'ultimo undo annulla l'effetto dell'undo precedentemente effettuato.

- Un flip undo ha senso se tutti i comandi devono essere trattati in maniera uguale: tutti i comandi devono essere recuperati compreso l'undo stesso
- La ragione perché l'undo venga trattato uniformemente è dovuta a scelte fatte da quei progettisti che ritengono sufficiente un undo relativo solamente all'ultimo comando immesso (anche perché è più facile da realizzare)

Multiple-step undo

I meccanismi di undo del tipo multiple-step (al contrario del flip-undo) consentono di eseguire undo un numero infinito di volte, o almeno fino a che il sistema ha informazioni disponibili sull'interazione passata

Per poter correggere eventuali errori nell'utilizzo di undo viene inserito il comando redo che consente di ripristinare l'azione annullata dall'ultimo comando di undo (es. ctrl+z e ctrl+y)

Es:

- 
- Ctri + z. annulla l'ultima modifica
- Ctrl + y. ripristina l'ultima modifica annullata

Su gmail dopo avere inviato una mail hai ancora qualche secondo per annullarla tramite un pop-up, che strani che siamo noi esseri umani che abbiamo bisogno di un sistema del genere.

4. Assicurare consistenza

Uno dei principi base dell'usabilità

Lo stesso comando o la stessa azione hanno sempre lo stesso effetto à l'utente acquista confidenza con il sistema ed è incoraggiato ad esplorare

Somiglianza in situazioni simili o perseguitando scopi simili

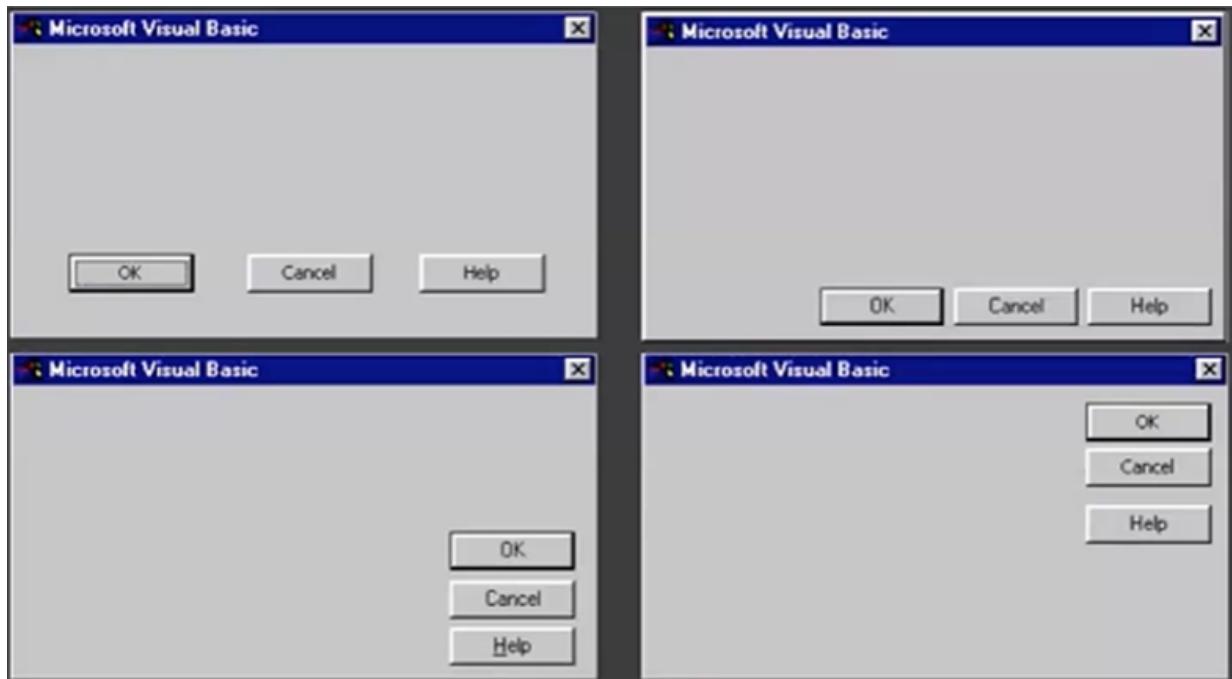
- Nell'ambito di un'applicazione, tra applicazioni diverse
- Nell'ambito di un compito, tra le aspettative dell'utente (modello mentale) e il modello concettuale del sistema

Generalizzabilità: supportare gli utenti ad estendere la loro esperienza da situazioni note a situazioni mai incontrate

Un tipo di consistenza la stessa informazione presenta sempre nella stessa locazione à in generale: mantenere consistenza nel layout dell'interfaccia

Esempi:

- **Abduzione semiotica:** se si sposta la posizione dei pulsanti apposta per ingannare l'utente si parla di dark patterns



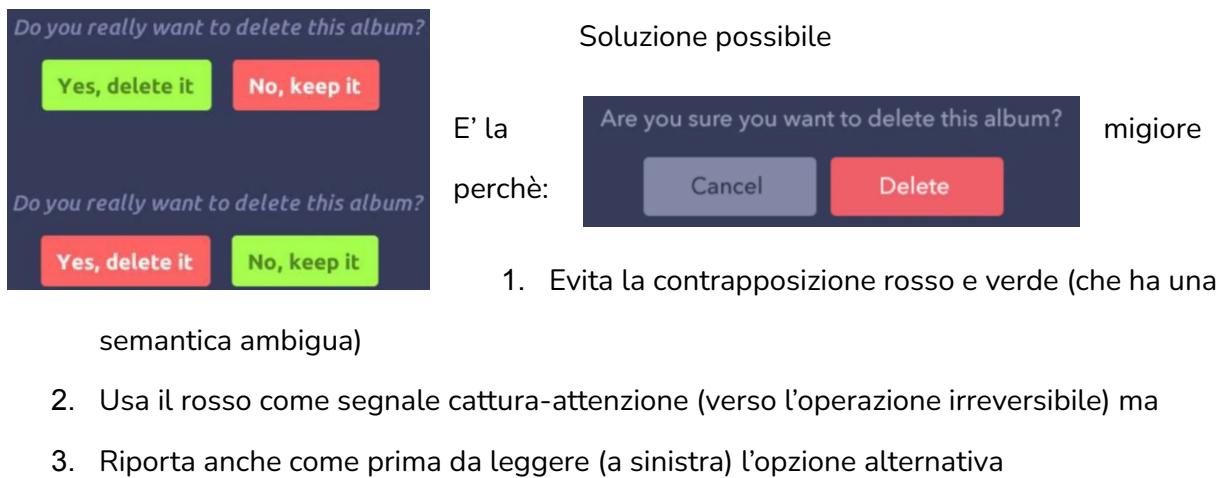
- Le **toolbars** sono i luoghi del controllo da parte dell'utente, dove gran parte dell'interazione "attiva" avviene. Nella valutazione euristica ci si deve prendere tempo per osservare la pagina principale, dove si trovano le aree, che comandi offrono, se ci sono ridondanze o inconsistenze...
- Le "barre", o toolbars, sono i luoghi del "controllo" da parte dell'utente, dove gran parte dell'interazione attiva avviene.

Nella vostra valutazione olistica, prendetevi tempo (un paio di minuti) per osservare la pagina principale, dove si trovano le barre, che comandi offrono, se ci sono ridondanze o inconsistenze.

Non iniziate subito a cliccare, ma guardate l'interfaccia per fare quante più ipotesi (abduttive) sul funzionamento probabile del sistema.

GARANTIRE LA CONSISTENZA

Nell'ambito del gruppo di sviluppo è importante adottare o definire convenzioni interne di rappresentazione e di interazione con rigore e..., farle rispettare.



Attenzione

Il punto è sempre quello. Non c'è una soluzione giusta o una sbagliata. Una bella o brutta. Una originale o tradizionale (trita). Il punto è: non affezionatevi ad una idea solo perché è la vostra. Siate anche originali MA: chiedete il parere dei colleghi, soprattutto se sono fuori dalla cosa. Ancora meglio: coinvolgete degli utenti (6-7, ma vari) e chiedetegli quale soluzione preferiscono (andate a maggioranza). Ancora meglio, coinvolgeteli in un rapido test utente e vedete quale soluzione induce MENO errori. SIATE EMPIRICI!

A/B testing (sometimes called **split testing**) is comparing two versions of a web page to see which one performs better. You compare two web pages by showing the two variants (let's call them A and B) to similar visitors at the same time. The one that gives a better conversion rate, wins!



Esempio:

Questa pagina poteva avere 5 immagini o 3 video diversi al posto del media. 4 Alternative per la scritta sul bottone.

Più di 310.000 visitatori, ogni variazione p stata testata da 13000 persone.



La variante vincente ha avuto

un tasso di iscrizione dell'1,6%.

La pagina originale aveva un tasso di iscrizione dell'8,26%.

Si tratta di un miglioramento del 40,6% nel tasso di iscrizione. In cosa si traduce un miglioramento del 40,6%?

[..]Ciò significa 2.880.000 indirizzi email aggiuntivi tradotti in 288.000 volontari in più.

[...] Gli ulteriori 2.880.000 indirizzi email sulla nostra mailing list si sono tradotti in ulteriori 60 milioni di dollari in donazioni.

CONVENZIONE DEL GORILLA

I risultati dei test devono essere spiegati in maniera da essere comprensibili anche a chi non ha competenze nel campo di interesse

💡: Un test binomiale è stato eseguito per confrontare il tasso di conversione della [soluzione A] (proporzionaA +/- [intervallo di confidenza]) rispetto a quello associato alla [soluzione B] (proporzionaB +/- [intervallo di confidenza]).

In base a tale test, si è trovata una differenza significativa [o non si è trovata differenza significativa] del tasso di conversione 8differenza-proporzione +/- intervallo-di-confidenza) tra la [soluzione A] e la [soluzione B]; Chi-squared = []), p = [il valore di P esatto].

Questo risultato suggerisce che gli utenti [non] preferiscono una soluzione rispetto all'altra quando usano il sistema.

[si può aggiungere un grafico]

5. Riconoscimento piuttosto che uso della memoria

Rendere **visibili** gli elementi del processo

→ Mostrare elementi del dialogo e permettere all'utente di **scegliere** (menu)

Usare **comandi generici** (facilitano trasferimento da un sistema ad un altro – facilitano la consistenza tra applicazioni)

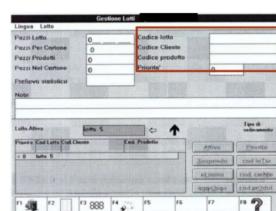
→ Es. comando “incolla” funziona (quasi) sempre allo stesso modo per testi e immagini

Facilitare la creazione di nuovi elementi mediante modifica di elementi esistenti (nomi, definizioni, programmi)

Attenzione all'**information overload**

→ Mostrare troppi oggetti prende tempo e offusca la scelta (vedi principio 7)

Esempi:



- Ricordare dati ricorrenti (es. nome, email...)

Obbliga l'utente a ricordare dati che sono già stati inseriti o che sono calcolabili (i codici sono di solito noti, e potrebbero essere richiamati da un elenco)

- Se si richiede all'utente di fornire un dato, descrivergli il formato, magari anche con un esempio (es. data)

L'utente non deve necessariamente conoscere l'intervallo dei valori legali di ingresso, il sistema deve fornire questa informazione

Collegare affordance simili a comportamenti applicativi simili.

6. Assicurare flessibilità ed efficienza d'uso

L'**utente esperto** deve poter svolgere le operazioni più frequenti

Uso di shortcuts (**acceleratori**): es. cltr+c, cltr+v, cltr+z

Flessibilità nel senso che il sistema deve **permettere di fare le stesse cose in maniere diverse** (es: per un utente novizio e per un utente esperto)

Esempi:

- Segnalibri
- Breadcrumb trail (percorso del file)
- Uso della storia dell'interazione (es. file usati di recente)
- Definizione di templates e macro
- Type-ahead (poter inserire il prossimo input prima che il computer sia pronto ad accettarlo)
- Fornire valori di default

7. Visualizzare tutte e sole le informazioni necessarie

L'interfaccia deve essere semplice (minimalist design, material design by Google) e gradevole (aesthetics matters) e rispecchiare nella maniera più naturale possibile i compiti dell'utente.

Presentare solo l'informazione che l'utente vuole quando l'utente la vuole come l'utente la vuole dove l'utente la vuole

.. quindi non accompagnata da informazioni non rilevanti, o fuorvianti.

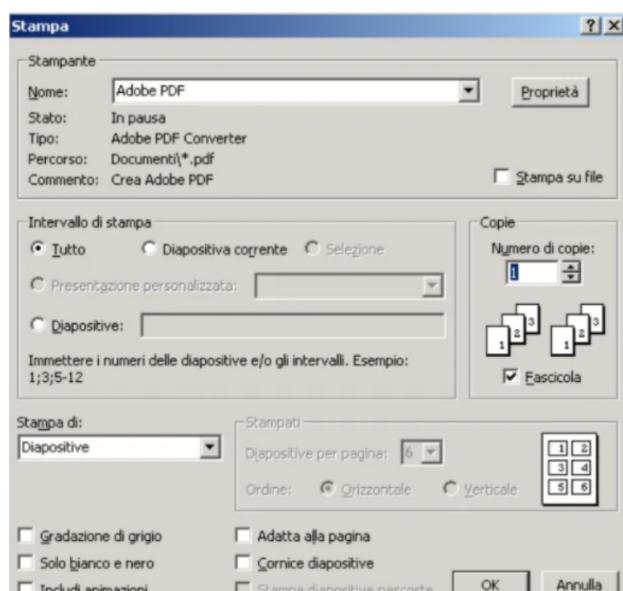
Raggruppare le informazioni che vengono usate assieme.

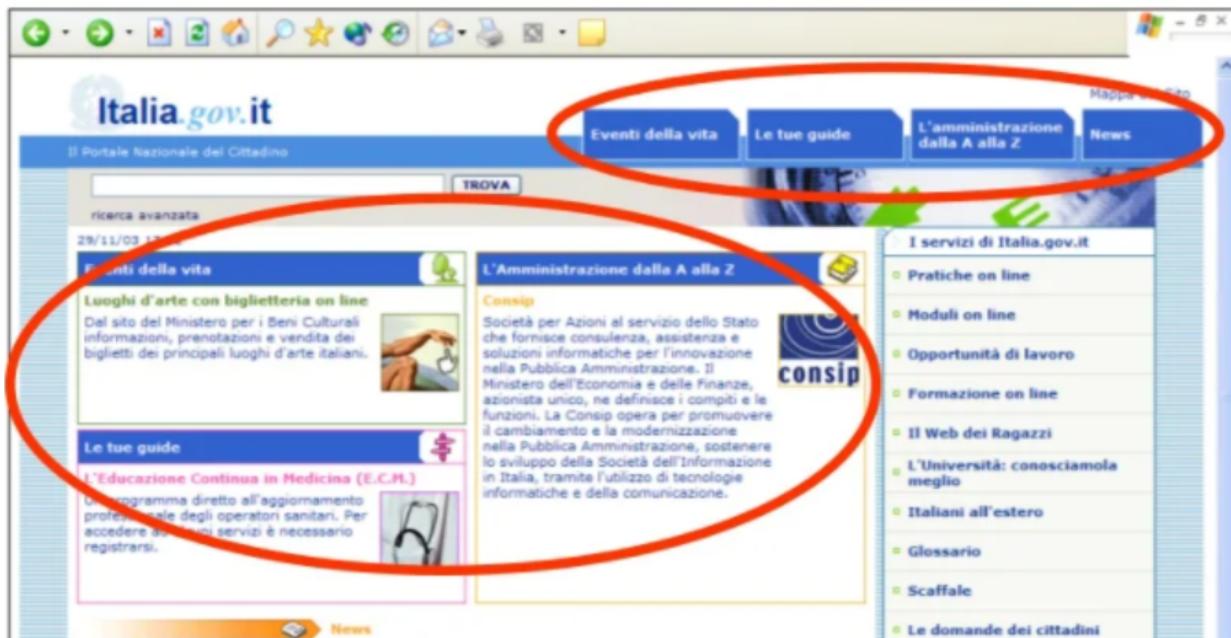
Presentare le informazioni in modo da suggerire la sequenza di azioni (es. i campi di una finestra di dialogo dall'alto in basso)

Raggruppare le informazioni

Una tale disposizione degli elementi è ottima per un utente esperto (se le shortcut come TAB funzionano). Per un utente non esperto chiaramente questa configurazione non è ottimale, la cosa migliore sarebbe una procedura del tipo Wizard.

Una tale disposizione degli elementi è ottima per un utente esperto (se le shortcut come TAB funzionano). Per un utente non esperto chiaramente questa configurazione non è ottimale, la cosa migliore sarebbe una procedura del tipo Wizard.





Cosa c'è che non va?

Sito ufficiale della pubblica amministrazione

Problema di leggibilità del testo, font

“ricerca avanzata” non sembra un bottone ed è più attaccato alla sezione sotto che la ricerca sopra

“Trova” è diverso da “Cerca”: la azione è il cercare, posso anche non trovare niente

Evitare la ridondanza!

Minimalità essenzialità, coerenza...

Ridondanza, information overload

“Più opzioni offro e più modi di fare le cose metto a disposizione, più utenti soddisfo” (noto come il “Bells and whistles bias” (in realtà può riguardare anche aspetti e dettagli estetici in eccesso)).

Ogni volta che si aggiunge un'opzione l'utente ha una cosa in più da imparare (se la trova...) e da usare in maniera scorretta.

Un'alternativa: offrire diversi "modi utente": es. per novizio ed esperto, per operatore di linea e manutentore, etc.

Low Use

Fenomeno del "Low Use" = fenomeno per il quale sono offerte funzioni non utilizzate.

Fenomeno del "Low Use", nella eHealth coinvolge fino al 70% dei medici.

Ma voi quante funzioni di MS Word usate correntemente? quante usate tutte le volte (80%)?

Use Index

Riguarda il rapporto tra funzioni usate spesso o a volte e quelle "funzioni che non sono note, oppure sono note ma usate praticamente mai". Simon et al. (2007, 10.1001/archinte. 167.5.507) considera solo quelle note e calcolano uno "Use Index" così:

$$\text{UI} = ((\text{Numero di funzioni usate spesso}) * 1) + (\text{numero di funzioni usate a volte} * 0.5) + (\text{numero di funzioni mai usate} * 0) / \text{numero di funzioni considerate}.$$

Definire "spesso" e "a volte"

Nota: non conta "quale" funzione specifica stiamo considerando, conta solo il numero totale.

I low users sono quelli con $\text{UI} \leq 0.5$

- Ragionare in termini di low users ci aiuta a capire quanto potenziale della nostra applicazione gli utenti stanno sfruttando (vogliamo che aumenti).
- Identificare le funzionalità a maggiore uso serve per prioritizzare attività di manutenzione correttiva (correzione bug) ed evolutiva (introduzione miglioramenti).
- L'idea è puntare sulle funzionalità più usate e lasciare indietro quelle meno usate...

Importanza della grafica, del layout del colore (estetica)

Mettere "prima" le informazioni che devono essere lette prima.

Oggetti lampeggianti o CARATTERI MAIUSCOLI, solo se necessario, anche perché rallentano del 10% la velocità di lettura (con il font adatto si può scrivere in maiuscolo, sono casi particolari).

In generale si considera un velocità di lettura media di 200-300 pam (parole al minuto, o wpm) per la lettura generica (reading) ma la velocità dipende dal task (memorizzazione, M=100; skimming M=700), dall'illuminazione, dai font usati, e dal livello di scolarità (scuole superiori: M-6 sillabe al secondo, SD=1; di solito si intendono parole di 5 lettere, quindi M-1500 (lettere al minuto).

Per calcolare la persistenza di messaggi di allerta o help considerate 200 pam

Test: a che velocità leggete: <http://www.readingsoft.com/>

Non usare più di 5-7 colori, non usare colori brillanti per lo sfondo

E considerare il daltonismo (riguarda in generale la accessibilità) che è un fenomeno diffuso tra il 2% e l'8% della popolazione (gli uomini circa 8 volte più frequentemente delle donne) e rende difficile distinguere il rosso e il verde dagli altri colori

Assicurarsi che gli elementi dell'interfaccia si distinguano anche senza i colori

Non dedicare troppo spazio a titolo, nome del venditore, numero di versione, e informazioni simili

8. Prevenire gli errori

Individuare le situazioni di possibile errore e cercare di aiutare l'utente

Creare meccanismi di prevenzione: es. selezionare un file da un menu piuttosto che richiedere di digitare il nome permette di prevenire una digitazione errata

Riduzione errori tramite richiesta di conferma di un'azione (non abusare delle richieste di conferma)

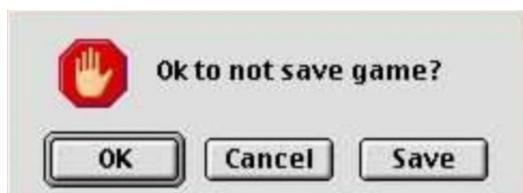
Evitare comandi troppo simili con significati diversi

Creare meccanismi di comunicazione utente - progettista: errori dell'utente rilevati nell'uso portano al ri-progetto

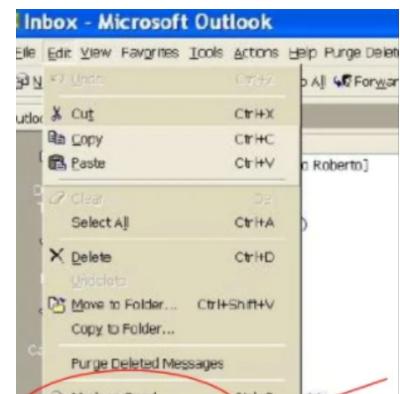
Tecniche di prevenzione

- Diversificare le azioni dell'utente
- Amodilità
- Funzioni vincolate
- Avvertimenti e richieste di conferma
- Default inoffensivi
- Bypass sicuri

Esempio 1



Esempio 2



Diversificare le azioni dell'utente

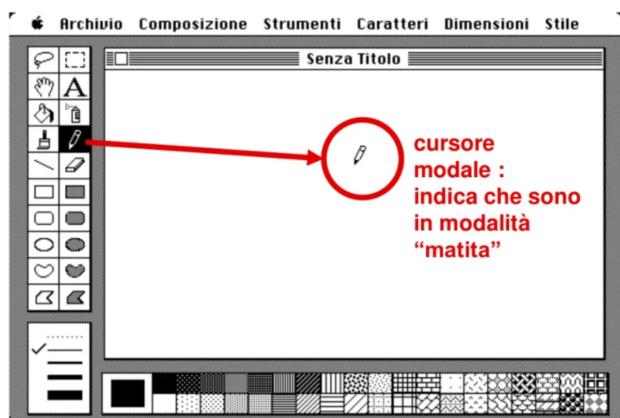
Tecnica per prevenire i lapsus ed evitare che l'utente esegua inavvertitamente un'azione al posto di un'altra

Se invece di selezionare Mark as Unread clicco su Mark all as Read non posso tornare indietro

Amodalità

Comportamento modale: il sistema si comporta diversamente a seconda dello "stato" (o "modalità") corrente

- modalità il trade off con Principio 7 (minimalismo)
- oppure evidenziare chiaramente la modalità corrente



Funzioni vincolate (obbliganti)

Progettare le vari funzioni in modo da rendere impossibili le azioni non lecite nel contesto corrente



Avvertimenti e richieste di conferma

Livelli di allerta:

1. "Tieni presente che ..." → Note
2. "E' proprio questo che vuoi?!" → Caution
3. "Fermo!" → Stop



Richieste di conferma

Chiedere sempre conferma prima di effettuare azioni irreversibili o comunque pericolose, ma senza esagerare!

Poco usabile!!

Ma serve per proteggere l'utente da se stesso

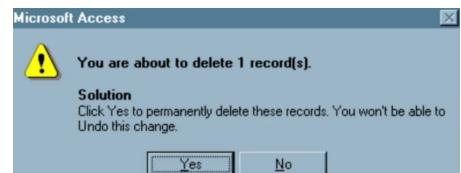
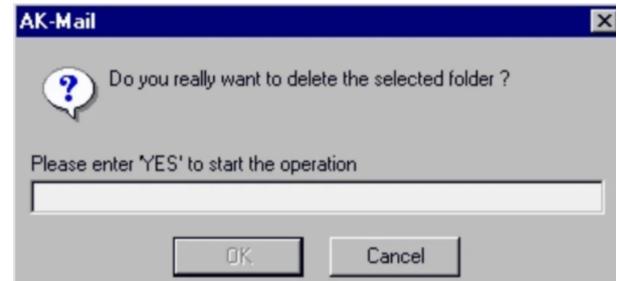
“Solution” implicita che c’è un problema: quale?

Default inoffensivi

Un valore, uno stato o un’azione che un dispositivo o un programma seleziona automaticamente a meno che l’utente non specifichi un valore sostitutivo

Esempi:

Word processor stabilisce i margini e il formato della pagina Alcuni programmi creano copia di scorta (back up)



BIGLIETTI	ABBONAMENTI
<input type="checkbox"/> Da dove vuoi partire? <input type="text"/>	
<input type="checkbox"/> Dove vuoi arrivare? <input type="text"/>	
<input type="checkbox"/> Quando vuoi partire? giorno <input type="text" value="24"/> mese <input type="text" value="01"/> anno <input type="text" value="2008"/>	
<input type="checkbox"/> A che ora vuoi partire? ora <input type="text" value="12"/> min <input type="text" value="37"/> <input type="button" value="Invia"/>	
Per acquistare un pass InterRail	

Usare per quanto è possibile default inoffensivi

Ex: Data e ora corrente

Bypass sicuri

Suggerire un’azione “sicura” in alternativa ad azioni che possono condurre ad errori



9. Permettere all'utente di correggere gli errori e non solo di rilevarli

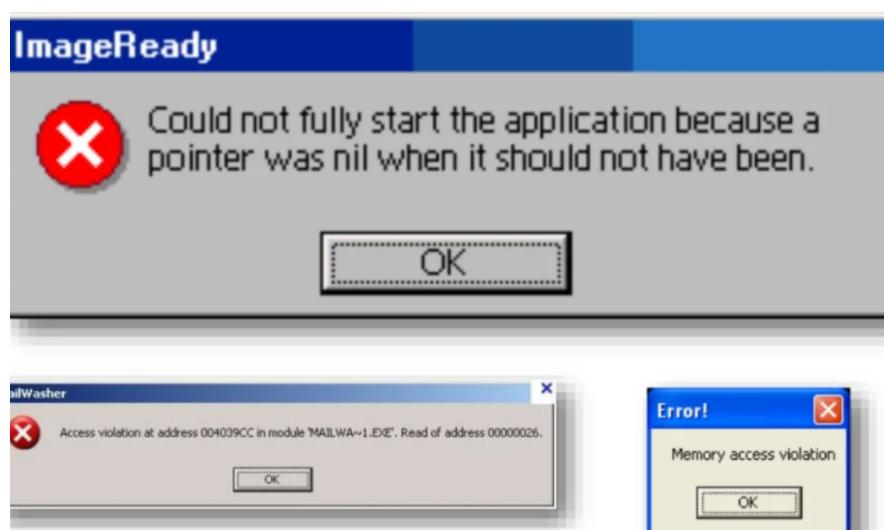
Buoni messaggi di errore

- in linguaggio chiaro e senza codici
- indicazioni precise
- indicare come risolvere il problema
- essere gentili

Recoverability: capacità di raggiungere il risultato voluto, dopo aver commesso l'errore

Azioni difficili da correggere dovrebbero essere difficili da fare

Messaggi poco chiari:



Messaggi precisi e costruttivi

Invece di dire “non posso aprire questo documento”

Dire “non posso aprire ‘Capitolo 5’ perché non è sul disco”

Google

Ricerca

Web

Immagini

Maps

Video

Notizie

hum-a-computer interaction

Cerca 32.900.000 risultati (0,44 secondi)

Risultati relativi a [human-computer interaction](#)
Cerca invece hum-a-computer interaction

[Articoli accademici per human-computer interaction](#)
[Human-computer interaction](#) - Dix - Citato da 3742
[Human-Computer Interaction](#) - Carroll - Citato da 297
[Human-computer Interaction](#) - Alan - Citato da 35

Essere costruttivi anche tramite correttori ortografici

Messaggi gentili e non intimidatori

ILLEGAL USER ACTION, JOB ABORTED!

Evitare l'utilizzo di termini come fatale, illegale....

In caso di errore

1. Alert

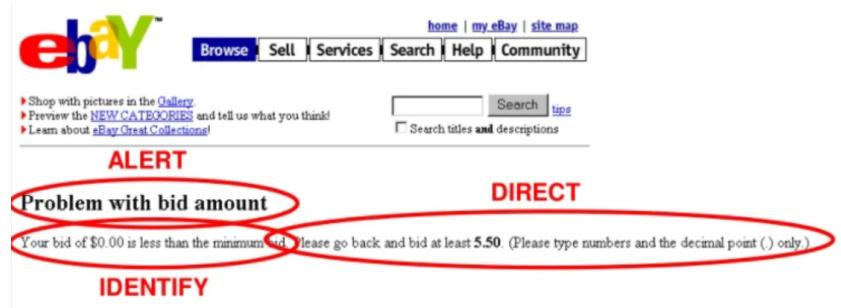
“Qualcosa non va”

2. Identify

“è questo che non va”

3. Direct

“ora devi fare questo”



Granularità dei messaggi di errore

Un messaggio di errore alla volta, appena ho commesso l'errore oppure tutti i messaggi di errore insieme, alla fine?

Vuoi saperne di più ?
Click per maggiori informazioni

Movie Music Adult Home Theatre

ONLINE MENU

- HOME PAGE
- REGISTRAZIONE
- CERCA/TROVA
- E SAMINA CARRELLO
- SALVATAGGI
- COMUNICAZIONI
- INFORMAZIONI E
- CONDIZIONI DI VENDITA
- NEWS --

ERRORE
Il campo User-ID è vuoto.

Tornare indietro e specificare un User-ID.

Ricorda lo User-ID può contenere solo caratteri alfabetici maiuscoli o minuscoli (0-9 / a-Z / A-Z) e i seguenti simboli:

1. @
2. -
3. _
4. —

1. chioceola
2. punto
3. segno meno
4. sottolineatura

Info line 059.4908027

Se hai bisogno di aiuto clicca . I campi contrassegnati da sono obbligatori.

Controllare i campi sottostanti

Codice Fiscale non corretto

Data di rilascio del documento: Si è inserito una data non valida

E' necessario inserire un indirizzo di E-Mail corrente su cui essere contattati

Il Cap e la provincia di residenza non sono congruenti

Inserire il numero di carta di credito

: La carta è scaduta

Torna indietro

Nome: ROBERTO
Cognome: POLILLO
Indirizzo: Manca l' indirizzo
Codice Postale: Manca il codice postale
Città: Manca la città
Provincia: Manca la provincia
Nazione: Italia
E-mail: rpolillo@unimib.it
Telefono:
Altro telefono:
Data di nascita: Manca il recapito telefonico o e' invalido

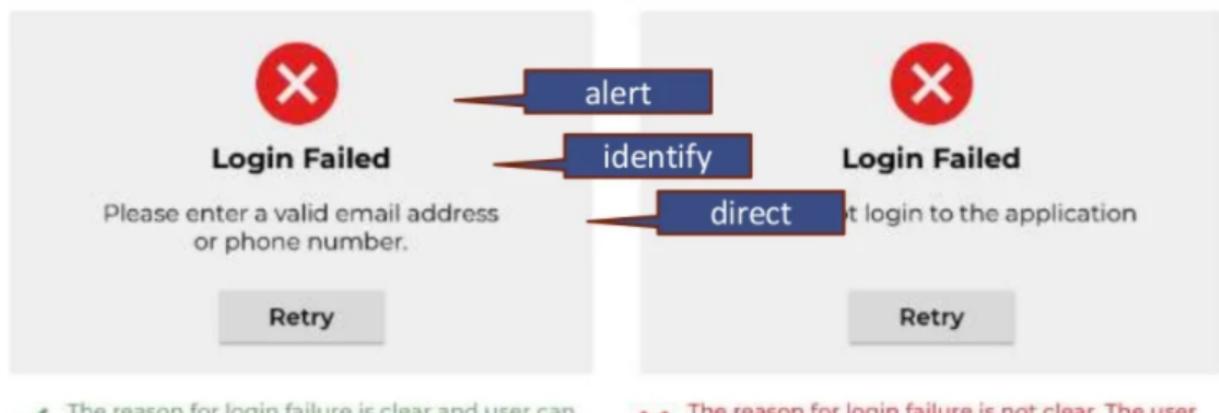
Indietro Avanti

Fine

Non creare trappole

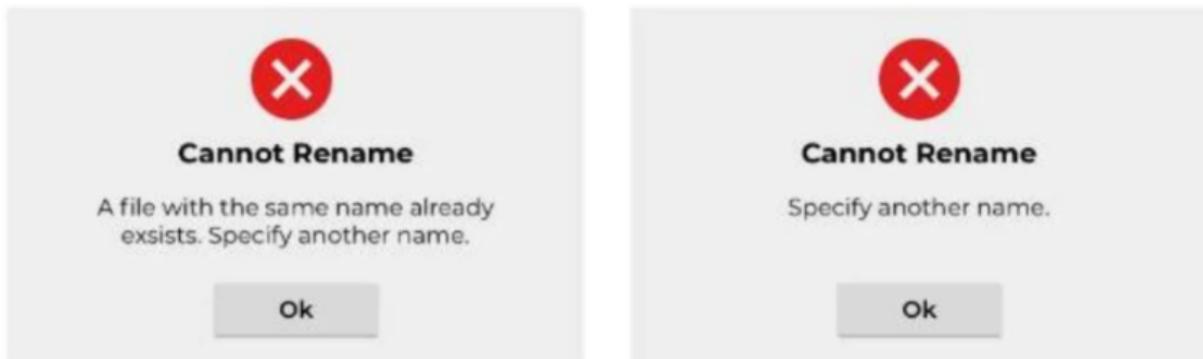
1-Dire cosa succede e perché

L'utente dovrebbe essere in grado di comprendere il problema durante la lettura di un messaggio di errore. Se il messaggio di errore non è chiaro e l'utente non è in grado di trovare il motivo del messaggio, non è di alcuna utilità. L'utente non può fare nulla per risolvere il problema e influisce negativamente sull'esperienza del prodotto.



2-Suggerisci il passaggio successivo

Ogni volta che si verifica un errore, l'utente desidera risolverlo il prima possibile. Il messaggio di errore dovrebbe avere informazioni sufficienti per l'utente che lo guida su come uscire dalla situazione errata.



✓ The message contains the problem, cause of error and solution.

✗ The cause of the error is missing in this error message.

3-Trova il tono giusto

Quando scrivi un messaggio di errore, scegli il tono che meglio si adatta al pubblico e al contesto.

- Se si tratta di un problema stressante o serio, un tono sciocco sarebbe inappropriato.
- Leggere il messaggio ad alta voce può aiutarti a individuare le parole o le frasi che devi rivedere.

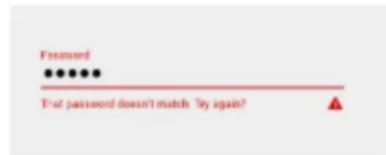
Troppo robotico

Decente

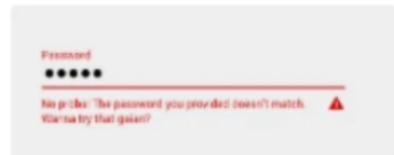
Troppo sciocco



Words like "bad request" and "supplied" make it sound robotic.



This one is pretty clear and approachable. Nice.



Would you actually say this? It's a bit too silly.

4-Non incolpare l'utente, sii umile

Trasmetti i problemi con garbo all'utente senza incolpare lui per le sue azioni. L'utente può eseguire un'azione errata ancora e ancora. Ma la responsabilità del design è di informarlo bene dei suoi errori.

"Un buon modo per incorporare un tono più umano nei tuoi messaggi di errore è pensare di spiegarlo ad alta voce a qualcuno. Come suona quando lo parli durante una conversazione."
 -Sonia Gregory



5-Leggibile dall'uomo

Utilizzare un linguaggio semplice e chiaro senza fare riferimento ai dettagli di implementazione. Se un messaggio contiene termini tecnici o linguaggio tecnico, l'utente si confonde. La maggior parte degli utenti non è interessata ai dettagli tecnici del problema verificatosi.



6-Evitare il testo in maiuscolo

Un messaggio di errore è un luogo in cui l'utente viene informato su alcuni scenari critici, quindi l'utilizzo del testo in maiuscolo può dargli un senso di scoraggiamento. Inoltre, il testo maiuscolo è difficile da leggere e dà un impatto di urla sull'utente.



"Il miglior messaggio di errore è quello che non compare mai" - Thomas Fuchs

Sintesi: progettare per l'errore

1. Comprendere le cause di errore, e minimizzare (prevenzione)
2. Facilitare la scoperta degli errori che comunque accadono, e facilitare la correzione
3. Rendere ogni azione reversibile, o rendere difficili le azioni irreversibili (undo)
4. Cambiare atteggiamento verso l'errore: non giusto/sbagliato, ma approssimazioni verso l'obiettivo

10. Fornire aiuto e documentazione

Il sistema perfetto è così facile da usare che non ha bisogno di help e di manuali

Siccome il sistema perfetto ancora non esiste è sempre bene prevedere help e manuali (...e per i sistemi walk-up-and-use?)

Ma... gli utenti non leggono i manuali, preferiscono provare a usare il sistema senza aver letto alcuna spiegazione

Quindi

- Mai dire: "è tutto scritto nel manuale d'uso"
- Fornire help in linea con documentazione raggiungibile in modo preciso

Manuali necessari nel momento in cui l'utente vuole diventare più esperto

L'help deve essere minimalista e contestuale (es. tooltip)

Anche l'help, visto come sistema a sé stante, può presentare problemi di usabilità

Esercitazione: Introduzione - 29/03/22

Professore

dott. Pietro Guardini

pietro.guardini@unimib.it

Etica... (l'ho perso)

Di chi è colpa? E' l'utente che non capisce quello che deve fare su un programma, o è il designer che deve costruire interfacce più intuibili?

L'utente spesso si da la colpa, ma è davvero colpa sua?

Il ciclo del design:



Con i quattro momenti che succedono quando costruisci un design

- Finding the need: trovo un problema
- Brainstorm alternatives: invento delle alternative di design
- Prototyping: costruisco con degli strumenti
- Evaluation:
 - valutazione predittiva: senza utenti
 - valutazione empirica: con utenti

Esempio prototipico, su cosa piace all'utente



Chi non ha la sedia a rotelle o altri impedimenti, per fare più in fretta, in questo caso usa un workaround che esce dagli schemi: quel pezzo di terra era pensato per fare da prato, non da marciapiede.

La strada creata che non è pavimentata in realtà è un'altra scelta di Design, che però non è stata scelta.

Vincoli, affordance, mappe, principi di design..cosa si vede?

Perché si preferisce tagliare?

Per far prima...ma per arrivare dove?

Piace stare vicino agli alberi? è più piacevole

Sono così create due strade, due modi diversi.



Il designer pensa agli aspetti riguardo la cognizione, l'emozione e il comportamento. Le cose studiate dal designer sono quelle che passano in modo diretto agli utenti. Gli utenti nel caso migliore usano lo stesso "modello mentale" del designer rispettando e seguendo tutto ciò che il designer ha pensato. Ma a volte succede che gli utenti non soddisfino le aspettative.

Il designer pensa a tutta la struttura

- Cognizione
- Emozione
- Comportamento

L'utente quando ha in mano il prodotto, lo può usare correttamente come aveva pensato il design, ma può essere usato anche in diverso modo, o non saperlo usare, non capire alcune funzionalità



Console: Come ci si avvicina ad un oggetto sconosciuto?

Si interagisce cercando di costruirsi un modello mentale interagendo anche “fisicamente”. Il modello può non avere niente a che fare con ciò che fa effettivamente l’oggetto.

Attenzione! il modello mentale non ha niente a che fare con la funzione per cui l’oggetto è stato veramente pensato. Esempio: usare un bicchiere come strumento musicale (senza capire che è fatto per contenere liquidi).

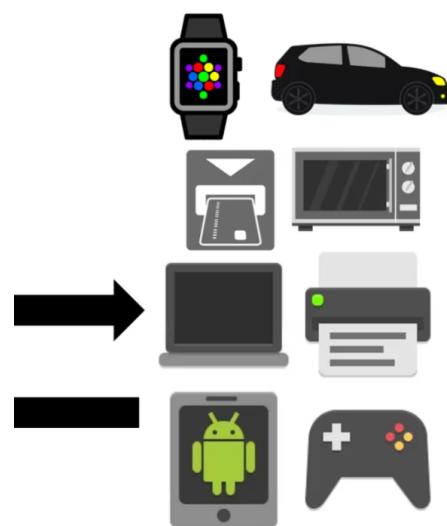
Canale YouTube “come i bambini interagiscono con...” e “come i teenager si approcciano a...” (teens react) o qualcosa di simile. PS: il nome in realtà è in inglese ma lui lo ha tradotto.... Questi video chiedono a chi è filmato di pensare tutto ad alta voce, questo permette di avere una finestra importante su quello che pensano le persone.

I modelli mentali si basano spesso, del tutto o in parte, su modelli mentali precedenti che sono nella nostra memoria. Quando abbiamo a che fare con qualcosa che conosciamo già usiamo un modello mentale nella mia memoria. Quando abbiamo a che fare con qualcosa di nuovo usiamo un modello mentale più simile possibile, e sperimentiamo per confermare il modello mentale (in tutto o in parte) e per aggiustare quel modello mentale in modo che si adatti all’oggetto nuovo, creando un modello mentale nuovo.

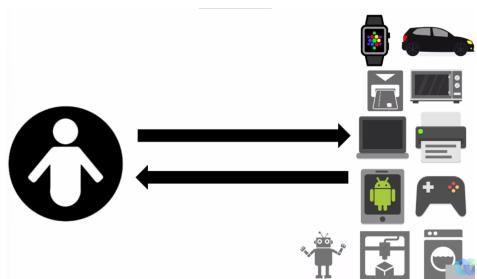
La scaramanzia nasce per comportamenti che, casualmente, vengono associati a esiti positivi di qualche altra azione. O meglio, ci si trova in una situazione in cui è impossibile negare con sicurezza che la scaramanzia è stata inutile.

Interazione Uomo Macchina:

- PC
- Stampante
- Tablet
- Console
- Smartphone
- Smartwatch



- Automobili
- Bancomat e altre stazioncine
- Elettrodomestici
- Stampanti 3D
- Robot
- ... altro



Utente → Task

Icona gender neutral (bambino?)

IUM non è da studiare come l'interazione tra l'uomo e la macchina, ma come all'utilizzo della macchina per raggiungere un fine generico, che (in un modo o nell'altro) si trova quasi sempre fuori dalla realtà della macchina.

Figma!!

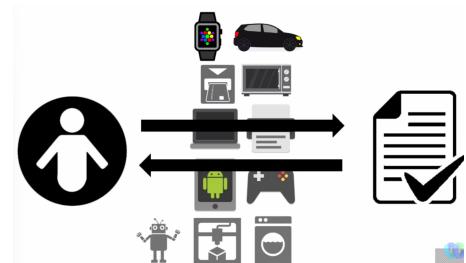
Come mi approccio a una richiesta di fare un'applicazione:

- Uso di figma per iniziare a creare l'interfaccia di una app
- ...

Esempi



Forno



Utente → Interfaccia → Task

Interfaccia: con due manopole...

Temperatura a cui cucinare? nel libro..informazione arriva da un'altra parte

Forno molto semplice con scanner che legge codice a barre. Prodotto che risolve molti bisogni, ma non è comunque conosciuto.

Abbonamento invece che comprare sempre...



Perchè non ha funzionato?

- Manca la fiducia
- Poco marketing
- Troppa pigrizia
- Troppo costoso, e posso usare solo pochi prodotti

Bisogna tenere presente: motivazioni, emozioni..

<https://www.youtube.com/watch?v=vJG698U2Mvo>

Non si vede tutto ciò che ci capita davanti l'occhio. Se un task ci chiede di tenere molta concentrazione, non si fa attenzione a cosa succede attorno: la nostra capacità di attenzione è limitata.

In nostro sistema cognitivo applica delle scorciatoie per elaborare meno dati visivi

Complemantamento modale: quando non c'è lo stimolo visivo, il cervello completa...

Un po' di illusioni ottiche: quello che vediamo e che il nostro cervello percepisce a volte inganna.

Design e magia sono collegati, perché il design porta l'attenzione nel posto giusto, e il mago fa lo stesso mentre svolge il trucco.

Informazioni utili per fare ricerche online

Siccome noi paghiamo la retta della bicocca abbiamo accesso ai sistemi di ricerca bibliografica con molti articoli VERIFICATI che trattano svariate ricerche (non ci sono tutte le cavolate che normalmente si trovano online). Cercare su Google "Bicocca accesso alle risorse da remoto" dovremmo accedere al sistema Prometeo.

Lui ha fatto una ricerca demo "videogame heuristics".

La prossima lezione continueremo con i sistemi cognitivi.

Euristiche

<https://www.nngroup.com/articles/ten-usability-heuristics/>

Conoscere un po' di psicologia e caratteristiche dell'utente

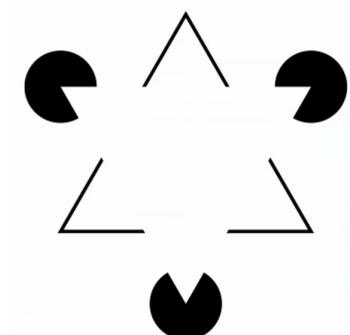
1. Visibilità del sistema
2. Corrispondenza tra il sistema e il mondo reale
3. Controllo e libertà dell'utente
4. Coerenza e standard

Tesla: macchina parte e sente che sei presente nella macchina Engy start: a dx vicino al piantone dello sterzo, perchè la chiave si infilava in quella direzione

5. Prevenzione degli errori
6. Riconoscimento piuttosto che ricordo
Icone della UI
7. Flessibilità ed efficienza di utilizzo
8. Design estetico e minimalista
9. Aiuta gli utenti a riconoscere, diagnosticare e recuperare dagli errori
10. Aiuto e documentazione

Ci dobbiamo chiedere, una ad una, se l'app che stiamo guardando le verifica o no.

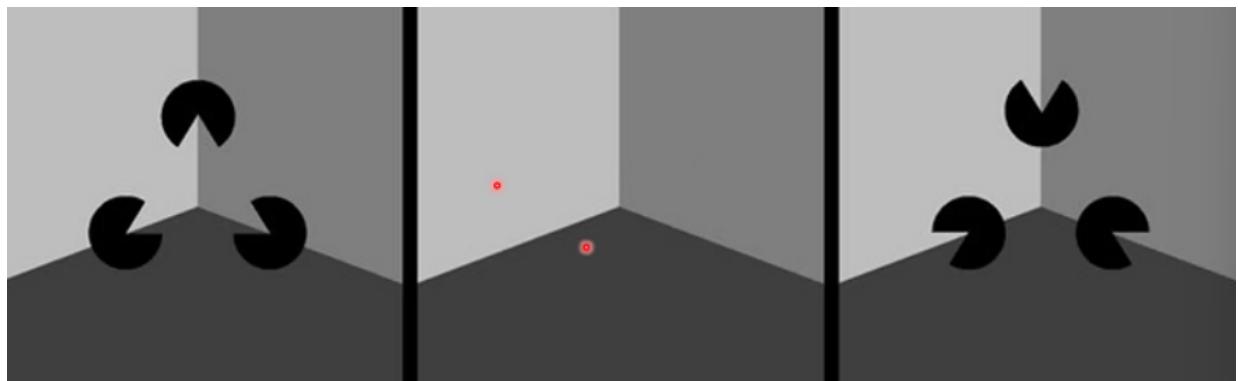
Illusioni



Tre pac-man e tre angoli fanno sì che vediamo due triangoli, uno bianco e uno contornato dal nero.

Quello bianco ci sembra pure più bianco dello sfondo

Il contorno non esiste: per il cervello è più facile vedere questi elementi uniti, invece che sei elementi diversi tra loro.



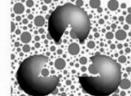
Piramide in fuori, con ombre Angolo di una stanza in dentro o in fuori. Angolo di una stanza in dentro



C'è un margine che non esiste nel secondo piramide



Cambia la luminosità: le punte smussate o appuntite, la profondità cambia



Rendendo più solidi i pac-man diventa tutto più stratificato.

Esperimento

Osservare bene per poco un'immagine e contare:

Quante icone hanno lo sfondo verde: 3 - 5

Quante icone hanno degli animali: 1 - 2



C'era tra le altre questa app? No



Sì Cambia lo sfondo o animale?



Icona con entrambe le quest



Boh x6 anche da casa



Si Boh - da casa



Boh



3 sì



2 boh

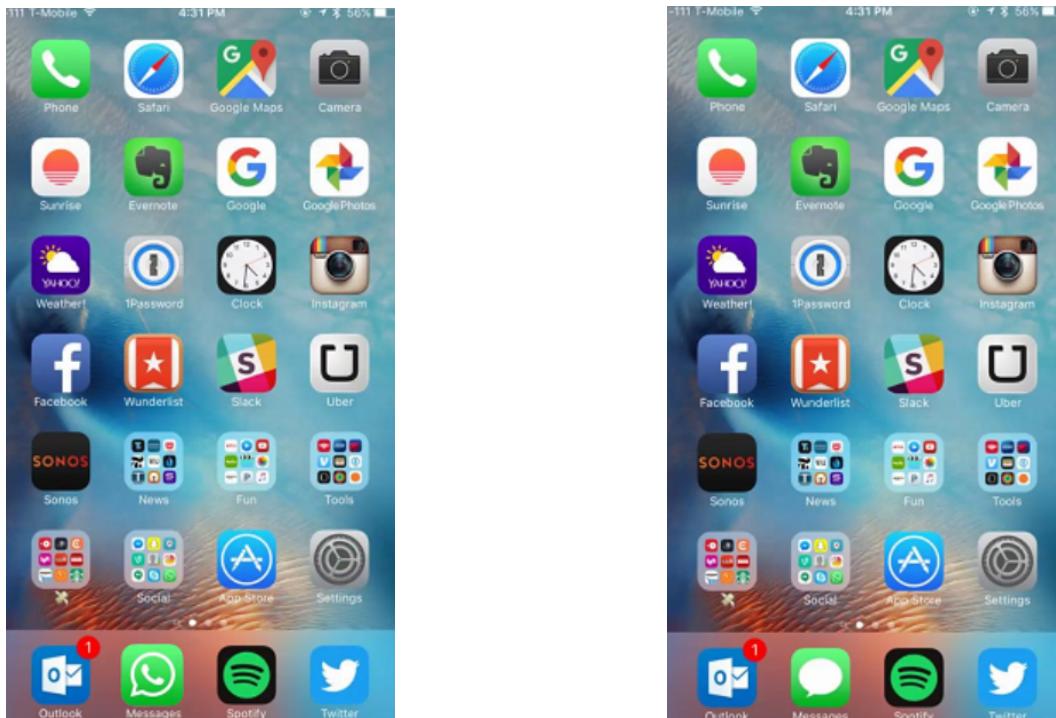
Cambiano tutte le prospettive e i risultati.

Se ad esempio si guarda una app velocemente, non si notano alcune cose perchè mi aspetto qualcos'altro, sto cercando altro

Anche le istruzioni possono modificare con l'utente interagisce

Se le cose chieste cambiano di poco, non viene notato

Quindi attenzione al disegno sperimentale



2 Animali e 3 a sfondo verde

Insta non c'è, è quella vecchia. L'animale c'è....

Questa è l'immagine che ha fatto vedere prima? (SX)

Ordine delle app diverso?

Il compito che quello che la persona sta cercando fare ci mette su diversi piani

Esempio menù barilla

Ristorante barilla:

con menù super tecnologico



Se piove non è
comodo, disturba
sensore, è



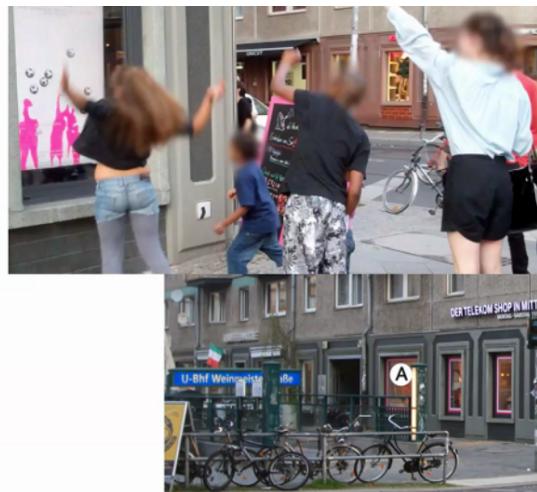
il
scomodo con qualcosa in mano.

C'è molta folla e può essere fastidioso

C'è un buco nella vetrina perché serviva per lo schermo.

Non è stato testato...

Esempio App per strada



Studio molto approfondito per creare questo gioco.

Le persone interagiscono con i palloni

Pericoloso perché si gesticola molto

Appeal limitato: ci si gioca per poco

Non ti porta ad entrare

Una delle sfide principali durante la creazione di display pubblici accattivanti è che le persone che passano di solito non sono a conoscenza di alcuna capacità interattiva.

Abbiamo progettato un'installazione interattiva che utilizza il feedback visivo dei movimenti accidentali dei passanti per comunicare la sua interattività.

Come attirare l'attenzione?



3 versioni:



Scrive cosa devo fare

Si muove quando mi muovo, subito

Si vede la persona stessa

Sono tutte delle ipotesi che vanno sperimentate per sapere il vero risultato

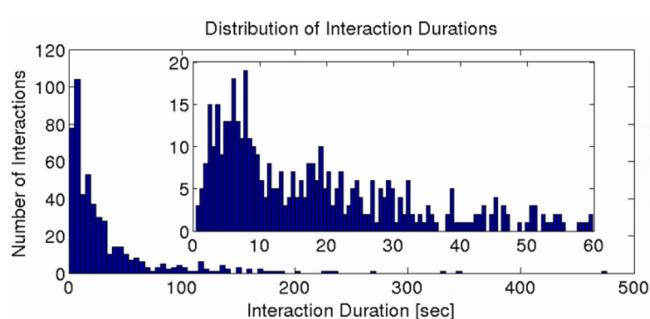
Rappresentazione dell'utente: 3 condizioni

- Niente
- Silhouette
- Image

Tipo di interazione: 2 condizioni

- Call to action
- Interazione involontaria

	No-representation	Silhouette	Image
Call-to-action	67	59	79
Inadvertent interaction	60	87	150



Per quanto sono durate le interazioni e quanti hanno interagito

Osservazioni qualitative



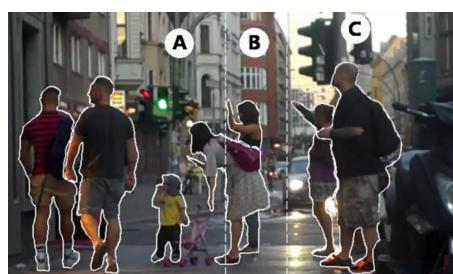
Dato qualitativo di cosa fanno gli utenti

E' più facile che si fermino gli utenti se sono almeno due persone
⇒ multiplayer è più efficace



Vedere una persona che interagisce, da più visibilità al progetto

Honeypot



Differenti piani di interazione. Per quanto il marciapiede è grande, le persone passano davanti e fanno perdere l'immagine all'app

Risultati:

- l'invito all'azione non è il massimo.
- Usare un'immagine dei passanti è la cosa migliore.

Più risultati:

- non hai bisogno di contenuti per più di 30".
- Effetto atterraggio.
- Effetto Honeypot.
- Robusto alle interruzioni.

I nostri risultati possono essere utilizzati per progettare applicazioni di visualizzazione pubblica e vetrine che comunicano in modo più efficace l'interattività ai passanti.

(Riferimento Müller J, Walter R, Bailly G, Nischt M, Alt F (2012) Looking glass: uno studio sul campo sulla notazione dell'interattività di una vetrina. CHI '12 Atti della Conferenza SIGCHI sui fattori umani nei sistemi informatici, pp 297-306.)

Interfaccia nella fantascienza

Come funzionano esattamente?

Il moro design può essere migliorato?

Alcune sono finte (tecnica del "Mago di Oz" o "Wizard of Oz"), mentre altre sono realmente esistenti

—



<https://www.youtube.com/watch?v=Pjgbivkm0Ms>



Interfaccia dei gesti di Minority Report

- Super popolare, quasi stereotipata.
- Progettato da John Underkoffler,
 - MIT Media Lab
 - Oblong <https://vimeo.com/214088355>

Lezione 1. Omissione ingannevole. Poche persone possono tenere le mani al di sopra del livello del loro cuore e spostarle per un periodo prolungato.

Lezione 2. Un'interfaccia gestuale dovrebbe comprendere l'intento. Il corpo dell'utente è il meccanismo di controllo, ma l'utente intende controllare l'interfaccia solo per una parte del tempo. Il sistema deve ospitare diverse modalità.

Errore di progettazione dal film: <https://youtu.be/PlqbivkmOMs?t=83>

Analizzatore di Immagini - Blade Runner

<https://youtu.be/tFrnY8fLrNE>

Interfaccia che si controlla con comandi vocali: parla e il pc fa quello che gli chiede, dando dei comandi specifici e misura specifiche

Altri Esempi:

Tipo: Alexa, Playstation, Google

Capire il contesto è il problema maggiore per le interfacce vocali. (Quando attivarsi)

Magari un controllo può essere utile per dire quando attivarsi o meno

Come so che comandi utilizzare?

Viene scritto a schermo “Prova a dire....”, Perchè farlo se mi basta cliccare (dipende da cosa chiedere)

Deve essere su internet per funzionare

Privacy: dico ad alta voce una password per ricordarmela e scriverla, mi ascoltano.

Come si svolge l'interazione

- L'utente inserisce la fotografia e accende la macchina



- Linee giallo/arancioni ad alto contrasto che scorrono: indicatore che si è accesa Tesla: non sanno come spegnere la macchina...Problema di usabilità (?) → se consuma batteria mentre l'utente è via, è possibile che si rimanga a piedi
- Schermo si accende, immagine compare con griglia
- L'utente dice "Esalta 22 41 76"
- Compaiano 3 elementi visivi "ZM 0000", "NS 0000", "EW 0000"
- Compare un cursore che al centro dello schermo
- La posizione del cursore ed i 3 elementi si aggiornano in risposta al comando vocale (i tre numeri menzionati non corrispondono perché i doppiatori italiani hanno cambiato i numeri pronunciati per questioni di labiale)
- Il rettangolo selezionato sulla griglia viene ingrandito
- Il tempo speso tra l'emissione del comando e il suo completamento è di 11 secondi



Ex: Fare il the è un compito semplice, ma descrivendolo è molto più complesso

⇒ Questa è task analysis, l'analisi di tutti i passi per arrivare con successo ad un certo obiettivo.

Primo confronto: battere a macchina o scrivere a computer

Ad ogni possa viene collegato il tempo che ci si impiega per compierlo, ottenendo un tempo tot. Infine vengono confrontati i tempi, il tempo più breve mi dice qual'è il migliore.

- L'utente dice "Esalta.....Stop"
 - Vengono utilizzati altri comandi vocali. Alcuni sono più efficaci di altri
 - "Indietro a destra"
 - Centralo e muoviti indietro; Un momento
 - Ascissa, 45 a dx; Esalta da 34 a 46
- La macchina riconosce comandi con e senza parametri. Quando il comando viene emesso senza parametri, la macchina continua a fare l'azione finché non viene fermata

Come si potrebbe migliorare?

- Elencare l'elenco di comandi sullo schermo, ma testo si sovrappone all'immagine, quindi è complicato
- Premere un tasto per fare una azione
- ⇒ Dare un numero o lettera alle righe e le colonne

101	102	103	104	105	106	107	108	109	110
201	202	203	204	205	206	207	208	209	210
301	302	303	304	305	306	307	308	309	310
401	402	403	404	405	406	407	408	409	410

Ex Amazon: sono nella pagina con il prodotto che voglio acquistare

1. Se sono già loggato e se il sistema ha salvato i dati della spedizione e del pagamento
Basta cliccare “compra ora”
 - Revisione in cui controllo dove spedirlo e con che carta pagare (step opzionale che può essere saltato)
 - Conferma (one click, 1 step)
2. Se non sono loggato, devo loggarmi, quindi inserire nome e password
 - E se non ho i dati salvati: sono molti da inserire, molto più complesso (5 step)

Nel primo caso, essendo molto più veloce il poter comprare, permette ad amazon di guadagnare di più

Nel secondo caso è più lento ma magari non si spendono soldi nell'acquisto compulsivo.

[Ex Tiger]: anche lui usa il tempo minimo necessario per comprare una cosa come punto di forza, ma soprattutto la poca “frizione” perché tutto costa poco ed è carino.

È etica questa cosa?

Sfruttare la psicologia delle persone

- Alcuni vogliono controllo assoluto, quindi più opzioni per verificare che sia tutto corretto.
 - Amazon ti permette di non attivare il “compro a un click” quindi c’è la revisione di tutto

Quindi il sistema permette stili differenti di usabilità(?)

Prossimi argomenti (?)

- Eye tracking
- Cognitive Walkthrough
- Journey mapping
- Strumenti per la prototipazione
- Dark Patterns: pattern per un beneficio dell’azienda più che per l’utente

Ex Tesla



Maniglia, finché è chiusa la macchina è interna, mi avvicino con la chiave in tasca e la maniglia salta fuori

È un affordance? NO, non è intuibile

No perché essendo liscio non invita a fare nessuna azione (al massimo premere), ma resta riconoscibile per il fatto che si trova dove si trova la maniglia in tutte le macchine normali.



È un signifier? SI quando è fuori

Luce che permette di vedere la maniglia anche al buio

Eye tracking

Come funziona il tracker: ci dice dove stiamo guardando



<https://www.realeye.io/> → “Launch demo”

<https://gazerecorder.com/> → GameReecorder “Online Live Demo”

Si può montare sugli occhiali e si può vedere dove ho messo gli occhi. Ma ci vuole anche l'attenzione: torcia elettrica

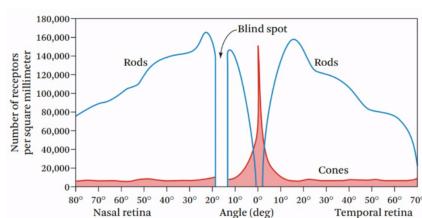
Com'è fatto l'occhio

“L'occhio è come una fotocamera”

Sotto sopra

Al centro si vede il colore e più a fuoco.

Macchia ceca: non ci sono recettori e non si vede nulla (occhio sx)



Numero di recettori per millimetro quadrato

Cones: coni - vedono a colori

Rods: bastoncelli - bianco e nero ma più sensibili

Salsicciotto volante:



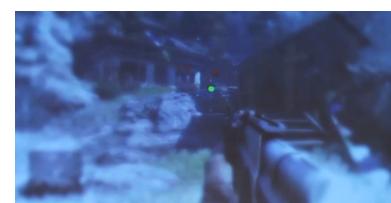
Esempio: Far Cry 4

Pallino verde: il giocatore sta guardando lì



Al buio si vede meno quindi bisogna illuminare di più

In periferia si vedo meglio i movimenti al buio



Lo user

Vista.

Centrale vs Visione periferica.

La visione centrale e periferica hanno proprietà diverse, dovute non solo alla retina ma anche alle diverse vie visive.

- Visione centrale:
 - ha una acuità visiva elevata
 - è ottimizzata per condizioni diurne luminose
 - è sensibile al colore
- Visione periferica:
 - è insensibile al colore
 - è più sensibile alla luce rispetto alla visione centrale in condizioni di oscurità
 - è meno sensibile alle lunghezze d'onda più lunghe (cioè rosso)
 - ha una risposta rapida e è più sensibile ai movimenti rapidi e lo sbarfallio
 - è meno sensibile ai movimenti lenti.

Esempio: Assassin's Creed: Parigi...



- Mappa con elementi
Ciclo base del gameplay

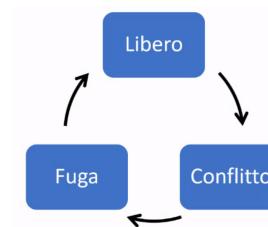
Visuale:

- Euristiche 7: tasti in alto a dx

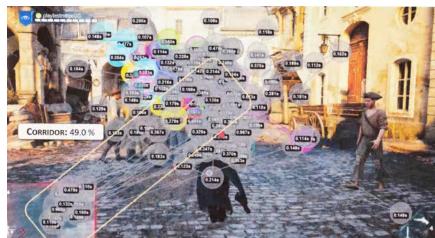
Mappa di temperatura: zona in cui le persone hanno visualizzato di più

Il 50% è al centro dello schermo

Ma solo l'8% alla mappa



0,1% linea di ???



Tondi = tempo di fissazioni

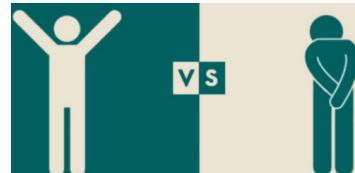
Righe collegano fissazioni, sono i movimenti dell'occhio

04/04/2022

Valutazione Quantitativa

Successo di un sistema interattivo:

- Gli utenti lo scelgono
- Gli utenti sono disposti a pagarlo bene
- La market share cresce a spese degli altri
- Il fatturato/ricavo cresce
- Viene comprato da una grande azienda
- Gli utenti ne parlano bene e lo consigliano a conosceni
- Gli utenti lo considerano parte della loro vita
- Gli utenti preferiscono usare il vostro sistema piuttosto che altri nuovi



Il **sistema di successo** è:

- **Utile:** fa quello che viene richiesto (es. riprodurre musica, formattare un documento, elaborare un'immagine) à se ne occupa il software engineering
- **Usabile:** deve consentire di fare quello che viene richiesto in modo naturale, senza aumentare i rischi di errore... à se ne occupa l'IUM in fase di pre-deployment

- **Usato:** deve rendere le persone desiderose di usarlo, quindi essere interessante, divertente, piacevole (User Experience – UX e Interaction Design – IxD) à se ne occupa la Data analytics e l'IUM nella fase di post-deployment

Bending Spuls: azienda molto importante in italia (sede a milano)

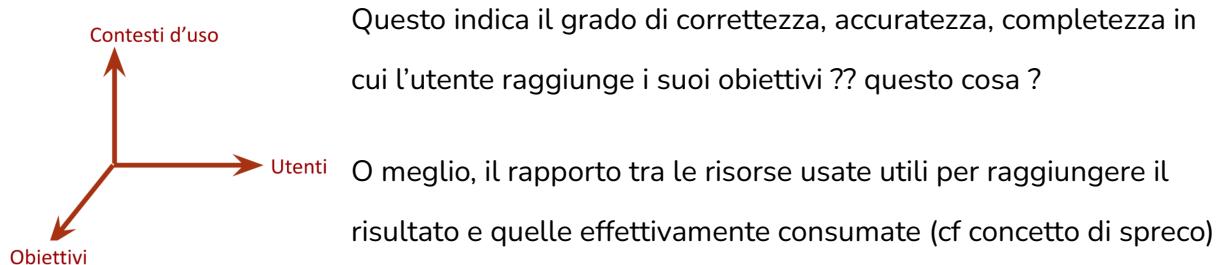
USABILITÀ

→ Il **grado** con cui un prodotto **può** essere usato da **determinati utenti** per raggiungere **determinati obiettivi** con **efficacia, efficienza, soddisfazione** in un determinato **contesto d'uso**. - Da sapere: ISO 9241

- **Grado:** possiamo definire delle metriche per misurare l'usabilità.
- **Efficacia:** grado di correttezza, accuratezza, completezza in cui l'utente **raggiunge i suoi obiettivi**.
- **Efficienza:** rapporto tra le **risorse usate** per raggiungere il risultato e quelle effettivamente consumate.

→ Efficacia è fare le cose giuste, efficienza è farle nel giusto modo.

- **Soddisfazione:** se l'esperienza (di raggiungimento del risultato) è stato “piacevole” (**UX**).
- **Contesto d'uso:** es. in casi di emergenza, distrazione, affaticamento



User experience (UX)

Per esperienza d'uso si intende ciò che una persona prova quando utilizza un prodotto, un sistema o un servizio. L'esperienza d'uso concerne gli aspetti esperienziali, affettivi, l'attribuzione di senso e di valore collegati al possesso di un prodotto e all'interazione con esso, ma include anche le percezioni personali su aspetti quali l'utilità, la semplicità d'utilizzo e l'efficienza del sistema.

L'esperienza d'uso ha una natura soggettiva perché riguarda i pensieri e le sensazioni di un individuo nei confronti di un sistema; inoltre è dinamica dal momento che si modifica nel tempo al variare delle circostanze.

Engagement è una particolare situazione di interazione uomo-sistema che è in grado di creare qualcosa a livello emozionale, cognitivo e comportamentale per l'utente.

Engagement riguarda quanto si trovano soddisfatti gli utenti durante l'uso dell'applicazione, quindi vengono proposte delle misurazioni proxy (di prossimità), ad esempio click-throughs (rapporto tra utenti che fanno clic su un collegamento specifico e numero di utenti totali che lo vedono), numero di download, tempo sullo schermo, numero di commenti su un blog post, Mi piace o retweet sui social media, utenti attivi giornalieri/settimanali/mensili (DAU,WAU, MAU).

→ Non ha senso parlare di usabilità di un sistema in sé, ma piuttosto in un determinato contesto d'uso, quindi ha senso parlare di **potenziale usabilità**.

Il prof tiene a questo concetto:

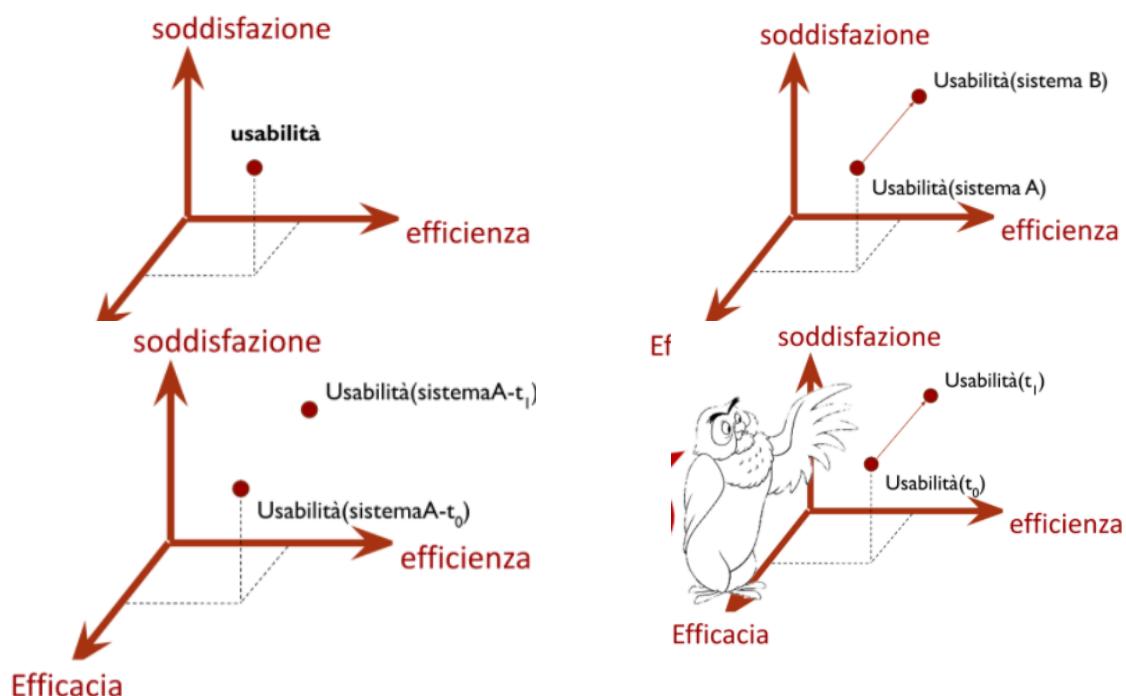
In ogni caso non ha senso parlare di usabilità di un sistema in sé, ma in un determinato contesto d'uso. Ha senso parlare di potenziale usabilità.

Differenza tra metrica e misura

- **Misura:**
 - Risultato della misurazione
 - Assegnazione empirica ed oggettiva di un valore (numerico o simbolico) ad un'entità, per caratterizzarne un attributo specifico (misurazione)

- **Metrica:** comprende il processo di misurazione quantitativa del grado di possesso di uno specifico attributo da parte di un sistema e definisce anche le procedure, modalità e regole di misurazione. Una metrica stabilisce:
 - Le entità/attributi da misurare
 - L'unità di misura
 - Una procedura di misura (per assegnare numeri e simboli)

→ Quando l'attributo non è “fisico” e difficilmente misurabile si parla di **indicatore**.



La **differenza di usabilità** può essere:

- **Praticamente significativa:** è quella che fa differenza.
- **Statisticamente significativa:** vuole semplicemente dire “non dovuta al caso” (la differenza c’è davvero).

Potete avere una differenza statisticamente significativa che NON è significativa ai fini pratici, mentre una differenza praticamente significativa è sempre statisticamente significativa.

Estensioni di usabilità

Noi non approfondiremo queste metriche

- **Robustezza** (simile all'efficacia): indurre un basso tasso di errori (o prevenire alcuni), facili da correggere e con basso impatto
- **Apprendibilità** (learnability): facilità nell'apprendere il comportamento del sistema (l'utente lo usa subito per attività utili) → più importante per gli utenti novizi
- **Ricordabilità** (memorability): facile da ricordare, anche quando un utente riusa il sistema dopo diverso tempo si ricorda come si fa → più importante per gli utenti occasionali

E' particolarmente importante per sistemi che vengono utilizzati di rado, ma che richiedono all'utente sicurezza ed efficienza nell'uso, anche senza consultare manuali.

Un chiaro esempio: sistemi di allarme

Trade-off tra learnability e soddisfazione:

- A volte conviene progettare sistemi "difficili da imparare" ma che danno soddisfazione all'utente esperto, perché si crea fidelizzazione (l'utente non cambia volentieri per il nuovo).
- Altre volte conviene invece fare sistemi che piacciono all'inizio anche se poi è più facile per l'utente passare ad altri, perché è la prima impressione quella che conta (e il rischio di essere accantonati dopo poche ore è alta).

Ovviamente un sistema che ha entrambi i profili non può esistere, ma bisogna sempre trovare un valido compromesso tra learnability nel breve e alta usabilità (soddisfazione) nel lungo periodo.

Trade-off tra robustezza ed efficienza:

- Ci può essere anche un trade-off tra robustezza ed efficienza
- Il desiderio di evitare all'utente errori (più o meno catastrofici) può portare a progettare un numero eccessivo di vincoli (vedi sopra) come ad es. Messaggi di conferma delle azioni dell'utente

→ I trade-off devono essere risolti definendo quali dimensioni di usabilità sono più importanti per un certo sistema.

Valutazione di **Qualità**- Usabilità

Attività di raccolta dati sull'usabilità di un sistema rispetto a uno specifico gruppo di utenti, nello svolgere una certa attività, in un certo ambiente, o in un certo contesto di lavoro.

È un'attività:

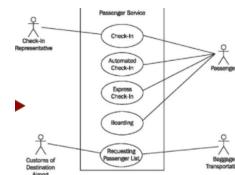
- Empirica: che coinvolge utenti veri (ovviamente non tutti, solo un campione)
- Concreta: eseguita su sistemi "reali" (non necessariamente finiti, anche prototipi)
- Sistematica: da eseguirsi con metodo

Attività di raccolta dati sull'usabilità di un sistema

Questo richiede di collegare la valutazione all'analisi (anzì si potrebbe pensare che la valutazione la completa!).

uno specifico gruppo di utenti, nello svolgere una certa attività, in un certo ambiente, o in un certo contesto di lavoro.

Si identificano gli obiettivi rilevanti (che gli utenti possono raggiungere usando il sistema), quindi i relativi casi d'uso cioè almeno il percorso/scenario principale (**happy flow** à passi principali per svolgere il compito).



Questo diviene il task/compito su cui valutare il sistema con utenti esperti. (questo è un collegamento con Analisi e Progettazione del Software, o Software Engineering).

Tipi di valutazione

Una valutazione può essere:

- **Formativa** (formative): aiuta nello sviluppo, aiuta a dare forma, all'inizio del progetto
Più rivolta a identificare problemi o opportunità di miglioramento che una "misura" dell'usabilità. Utile come strumento di miglioramento nel processo iterativo di sviluppo prototipale.
- **Riassuntiva** (summative): giudizio sul prodotto, dopo il deploy del prodotto

Svolta per avere una misura di usabilità.

Può essere volta a ottenere un valore assoluto oppure, più utilmente, relativo (ad altri sistemi, al benchmark), cioè comparativo.

- **Assoluta**

Ci dice ad es.: quanti problemi ci sono, la loro rilevanza/priorità, il parametro di tendenza centrale (media, mediana, moda) di una variabile di qualità (e.g., soddisfazione).

- **Comparativa**

Produce un confronto tra sistemi, tra il sistema e un benchmark, tra gruppi di utenti, tra release diverse dello stesso sistema.

Molto più utile perché a voi interessa sapere se il vostro sistema è migliore oppure se è migliorato.

- **Qualitativa**

Coinvolge esperti (con o senza utenti) e non è volta a misurare, bensì ad analizzare ciò che essi fanno, dicono, scrivono, le loro opinioni, i loro comportamenti (e.g., protocol analysis, valutazione euristica, questionari).

- **Quantitativa**

Basata su test utente (basati su task o scenario) e volta a misurare degli attributi di qualità, ad esempio tempi di esecuzione (efficienza), rapporto tra interazioni corrette ed errori (efficacia), e anche l'opinione degli utenti/experti “quantificata” su scale psicométriche.

QUALSIASI VALUTAZIONE È MEGLIO DI NESSUNA

Perché la valutazione?

- **Capire il mondo reale**

→ Come lavora l'utente, come si può migliorare la sua attività, se i suoi sistemi sono di qualità

- **Verifica di conformità agli standard**

→ Es. la leggibilità dello schermo è accettabile? Gli enti di standardizzazione emettono procedure che permettono di verificare rigorosamente il rispetto degli standard

- **Confrontare progetti/sistemi**

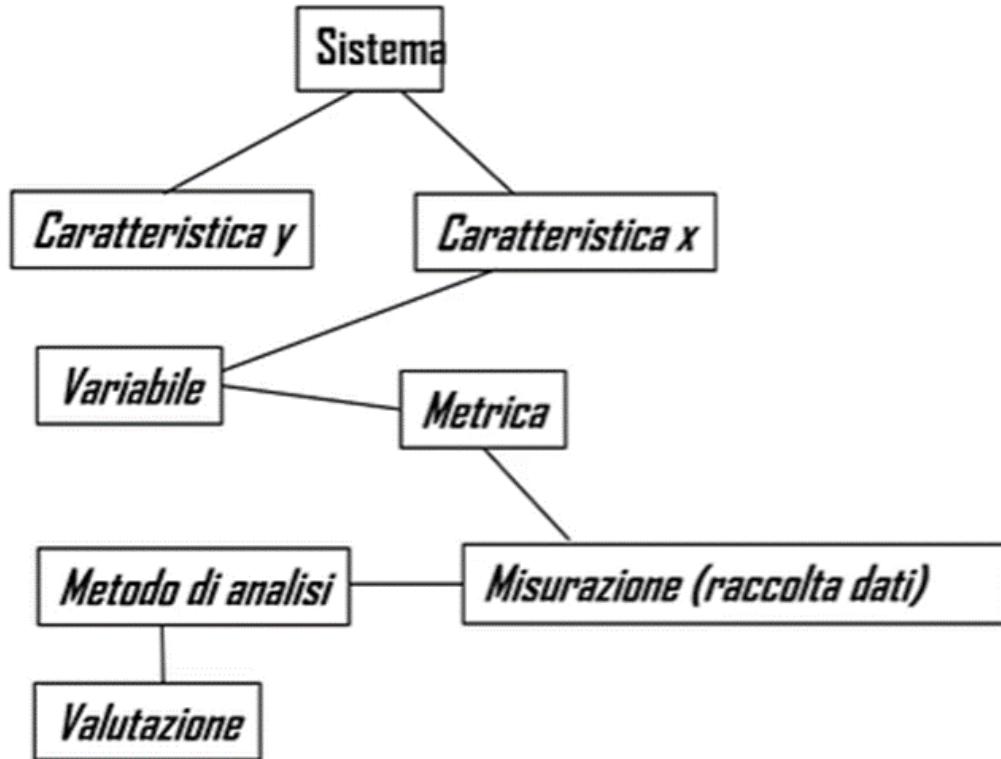
→ Durante l'attività di progetto confrontando prototipi, o dopo la realizzazione per capire se il sistema è migliore di altri (concorrenti)
Confronto trasversale (tra sistemi diversi) o longitudinale (di un sistema nel tempo)

05/04/2022

Come fare la valutazione di usabilità?

- 1) **Metodo “scientifico”** (positivista, quantitativo): adotteremo un metodo di analisi di natura statistica campionaria: analizzeremo i dati relativi a un campione estratto dalla

popolazione, e faremo un'analisi di "inferenza statistica inversa".



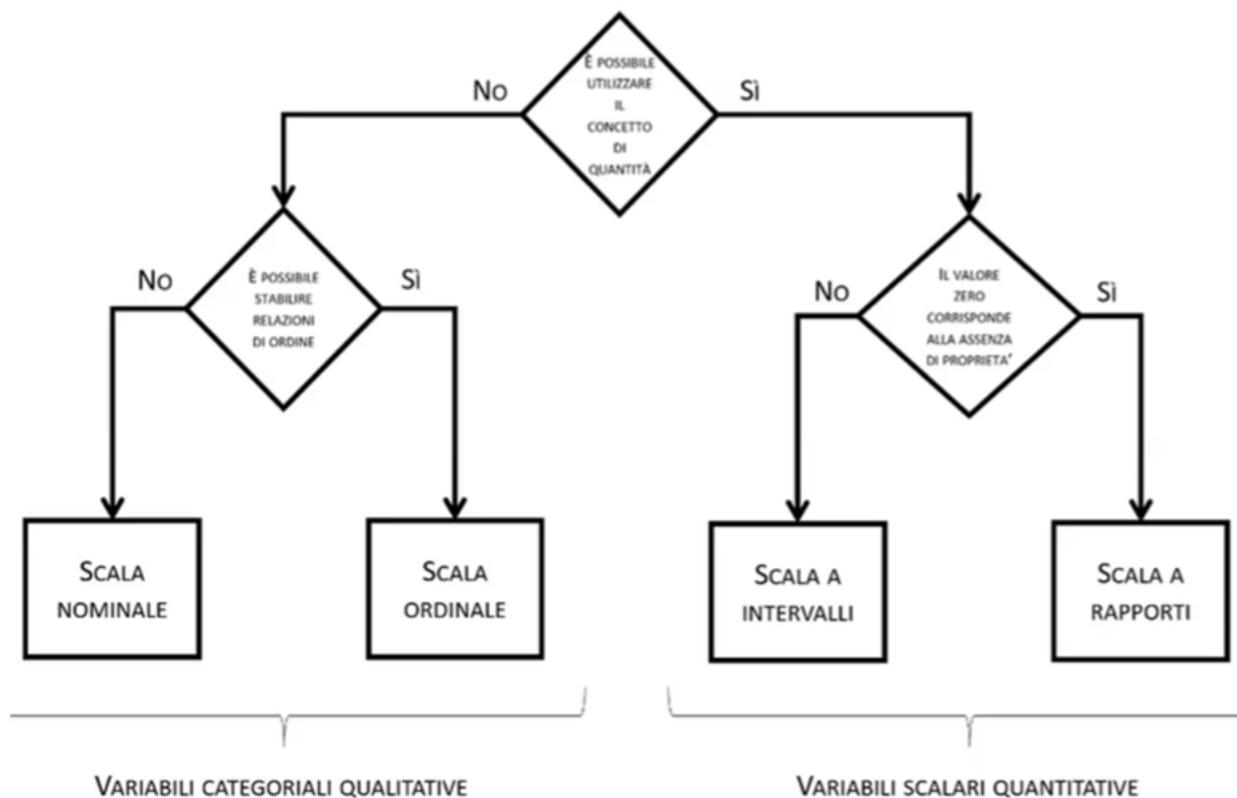
2) Metodi sociologici-antropologici (etnografici, etnometodologici, qualitativi)

- Osservazione e monitoraggio: tecniche che derivano dalla sociologia per capire l'interazione nell'ambiente di lavoro
- Raccolta opinioni degli utenti: metodi strutturati e semi-strutturati per chiedere all'utente cosa pensa e registrare le sue risposte
- Interpretativi: dati raccolti con tecniche naturalistiche per non disturbare l'utente e per capire come il sistema si integra con le altre attività

Come si confronta?

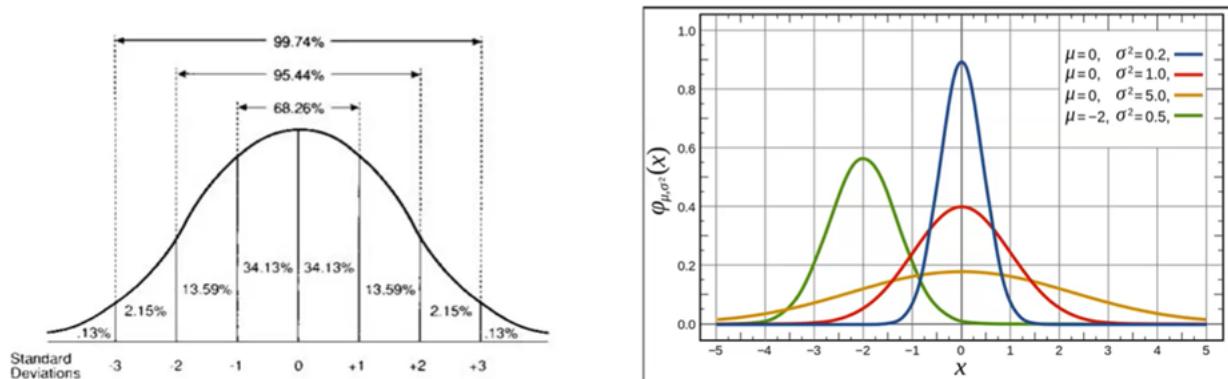
1. Si raccolgono dati (numerici, ordinali o categorici) da un campione di utenti del sistema
 - **Variabili scalari** (o numeriche): (es. altezza, tempo, numero di errori)
 - **Variabili ordinali**: è possibile stabilire un ordine totale sui possibili valori/attributi (es. titoli di studio, classi di età, giudizi di merito)

- **Variabili categoriche** (o nominali): i valori sono modalità, cioè parole (es. sì/no, maschio/femmina)



Tipo di variabile	Nominale Qualitativa categorica	Ordinale Qualitativa rettilinea o ciclica	A intervalli Quantitativa	A rapporti Quantitativa
Finalità	Classificazione	Ordinamento	Misura attraverso applicazione di una unità di misura, conservando le relazioni tra distanze e adottando zero arbitrario.	Misura attraverso applicazione di una unità di misura, conservando il rapporto tra i valori e adottando uno zero assoluto.
Operazioni matematiche di base	=, ≠.	=, ≠, <, >.	=, ≠, <, >, +, -, ×, /.	=, ≠, <, >, +, -, ×, /.
Indicatori statistici (esempi)	Numero di casi (N), Moda.	Mediana, Percentili, Quartili, scarto interquartile, Correlazione per ranghi di Spearman.	Media, Varianza, Deviazione Standard, Correlazione di Pearson.	Media, Varianza, Deviazione Standard, Correlazione di Pearson, concentrazione.
Rappresentazione dei dati raccolti	Aerogrammi, Istogrammi	Istogrammi, Diagramma a scatola e bassi (box and whiskers plots).	Diagramma a scatola e bassi; Distribuzione continue; Grafico di dispersione.	Diagramma a scatola e bassi; Distribuzione continue; Grafico di dispersione.
Test di ipotesi inferenziale (esempi)	Test Binomiale, Test del Chi Quadro.	Test Binomiale, Test del Chi Quadro, Test di Mann-Whitney, Test di Kruskall-Wallis.	Test T di Student, ANOVA	Test T di Student, ANOVA

2. Si raccolgono le statistiche per descrivere il campione (es. mediana, media)



Non si può conoscere il parametro della popolazione, ma solo la “statistica” di un campione (es. media). Non si può essere precisi (si commette sempre un **errore standard – ES**), ma si può stimare l’intervallo che è ragionevole pensare contenga il parametro “vero”, questo è **l’intervallo di confidenza**.

Di solito si stima un intervallo di confidenza del 95%, accettando quindi sbagliarsi 1 volta su 20.

⚠️ Un intervallo di confidenza al 95% NON significa che esiste una probabilità del 95% che il parametro della popolazione sia compreso tra i due estremi dell’intervallo.

→ L’intervallo di confidenza è l’intervallo di valori che siete abbastanza confidenti che contenga il valore vero del parametro. Indica l’incertezza intrinseca nella vostra stima del parametro.

Intervallo di confidenza del 95% à t vale 1,96 (99% à 2,58)

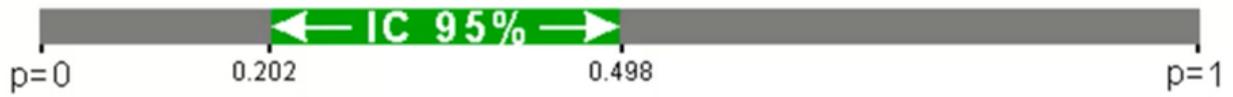
$$ES = \frac{\sigma}{\sqrt{N}}$$

Per la media:

$$ES = \sqrt{\frac{p*(1-p)}{N}}$$

Per le proporzioni:

Più è alto il livello di confidenza, maggiore è l’intervallo di confidenza (~t). maggiore è la numerosità del campione, maggiore è la precisione (cioè minore è l’errore standard – ES)



3. Si cerca di stimare i parametri dell'intera popolazione

Siti per calcolare:

- Medie: http://vassarstats.net/conf_mean.html

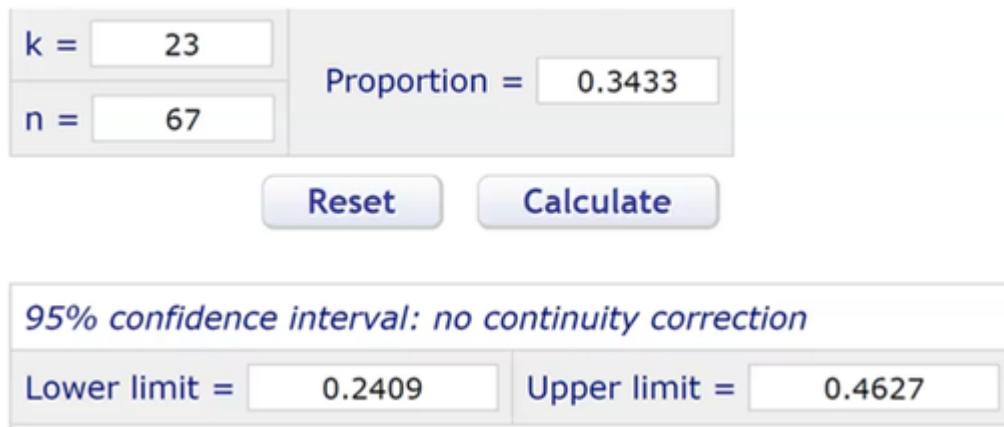
Data Entry:

Please remember to perform the data-check procedure.

Reset

Calculate

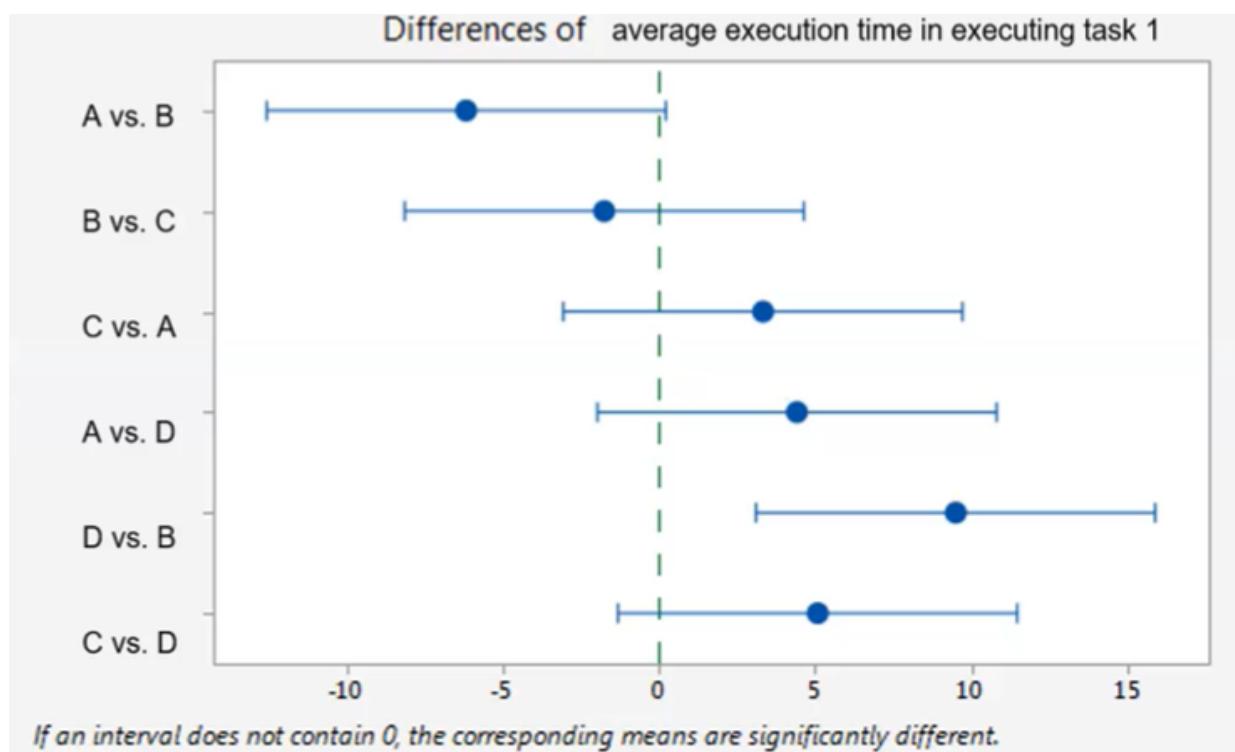
- Proporzioni: <http://vassarstats.net/prop1.html>



APPROCCIO DI NEYMAN

→ Si confrontano gli intervalli di confidenza

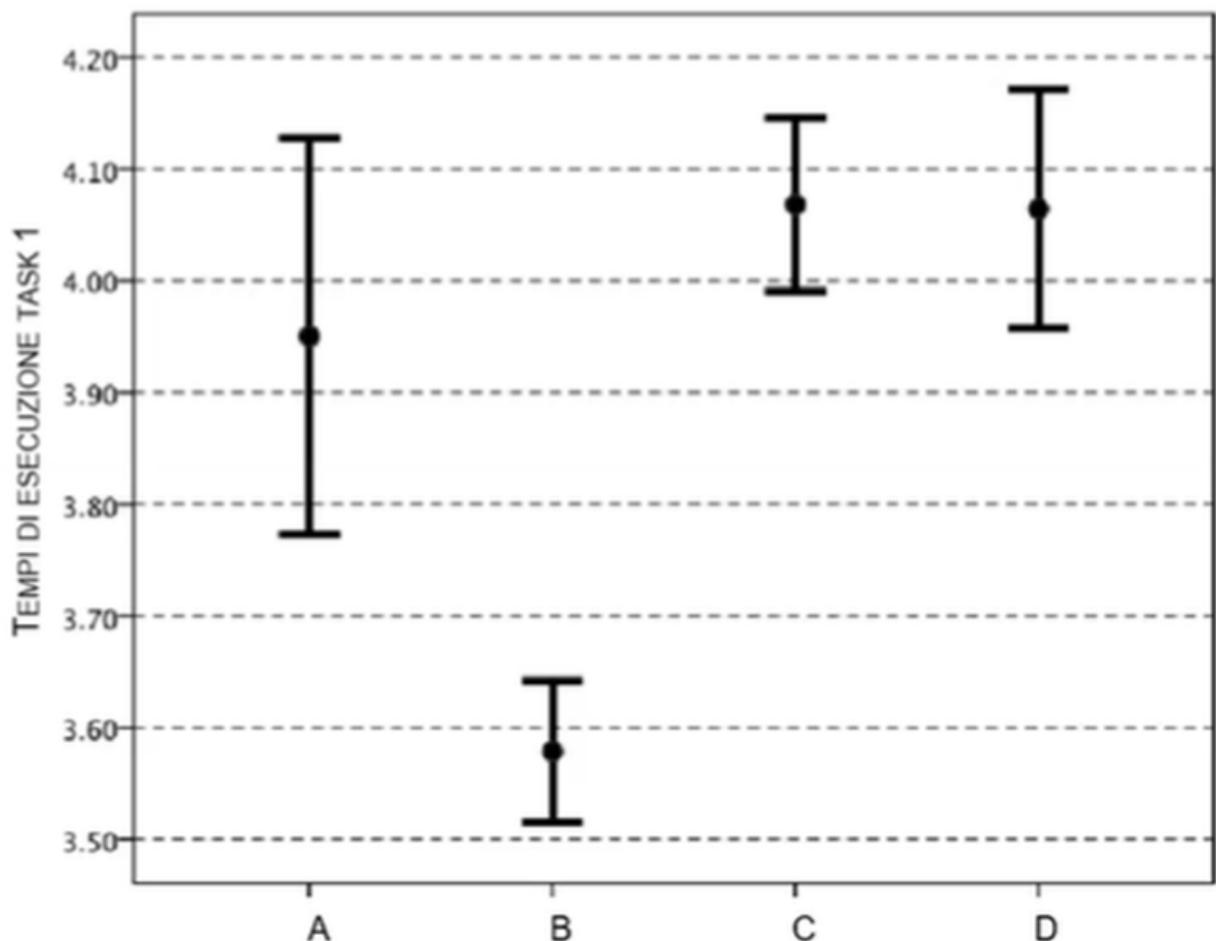
- Intervallo di confidenza della differenza (se passa per la linea 0 non c'è differenza statisticamente significativa).



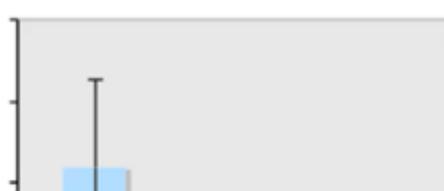
- Lo si capisce anche visualizzando gli intervalli di confidenza dei tempi medi di esecuzione (non più della media della differenza), se gli intervalli si sovrappongono non c'è una differenza statisticamente significativa.

Nonostante dalla media dei tempi degli esperimenti si potesse stabilire che alcuni hanno media più bassa/alta, sovrapponendo gli intervalli di confidenza notiamo che non c'è una differenza statisticamente significativa tra A, C e D, mentre B ha una differenza di tempo statisticamente significativa rispetto agli altri.

⚠ Tutti i valori dell'intervallo di confidenza sono equiprobabili.



→ Gli intervalli di confidenza al 95% appaiono ugualmente ampi a parità di varianza e numero di utenti considerati, altrimenti appaiono di ampiezze diverse.



La differenza percentuale di utenti soddisfatti si può intuire solo dagli intervalli di confidenza (quelli degli anziani non si sovrappongono).

Senza gli intervalli di confidenza non si può dire nulla a livello di significatività né di intera popolazione.

Nel caso della mediana l'intervallo di confidenza si mostra con i boxplot (notch – strozzatura):

- Gli estremi sono i valori massimi e minimi
- La scatola rappresenta il range interquartile ($Q_1 - Q_3$), nel quale ricadono il 50% delle osservazioni
- La linea centrale rappresenta il secondo interquartile (o mediana)
- La strozzatura rappresenta l'intervallo di confidenza della mediana

APPROCCIO DI FISHER

→ Si formulano test di ipotesi e si verificano con essi le ipotesi “nulle” (H_0)

L'ipotesi nulla è assunta vera e prevede che non ci sia differenza tra:

- I sistemi s_1 e s_2 (per la medesima caratteristica), o
- Per il sistema s tra la versione a tempo t_0 e quella a t_1 , o
- Tra strati (gruppi) di utenti (ad es. maschi vs femmine, giovani vs senior, esperti vs non esperti)

→ Per noi l'ipotesi nulla sarà sempre H_0 : non ci sono differenze statisticamente significative tra A e B (da scrivere nella relazione)

Quindi assumiamo che A non sia meglio di B per quanto riguarda: i tempi di esecuzione, gli errori indotti, ...

Noi saremo soddisfatti se raccogliamo dati campionari tali da poter rigettare l'ipotesi nulla. L'ipotesi è assunta vera fino a prova contraria.

“L'assenza di prove non è prova dell'assenza.”

Procedura di statistica inferenziale:

- Si sceglie il test giusto per l'ipotesi e il tipo di dato che si raccoglie
- Si esegue il test e si calcola il corrispondente “livello di significatività osservato” dei dati (detto anche p-value)
- Si riportano i risultati in uno stile chiaro e “confrontabile” (es standard APA)
- Ipotesi nulla: non ci sono differenze tra A e B
- Ipotesi alternativa: ci sono differenze tra A e B

Le ipotesi vengono assunte vere

→ P-value: la probabilità di raccogliere dati come quelli raccolti (o ancora più estremi) assumendo l'ipotesi nulla H_0 come vera.

Se questa probabilità è bassa, è improbabile che la differenza sia solo dovuta al caso (cioè che scegliendo un altro campione non le troveremmo).

Se il p-value è inferiore al livello di significatività impostato a priori (ad es 5%) allora si può rigettare H_0 e congetturare che il suo contrario (H_1) sia vero (cioè che le differenze siano reali).

Altrimenti, se $p > \alpha$ il confronto è stato utile lo stesso, ma non si sono raccolte prove sufficienti che ci siano reali differenze.

⚠ Se non riusciamo a rigettare H_0 , questo non vuol dire che sia veramente vera.

Il p-value è la $P(D|H_0)$, cioè non è né $P(D)$ né $P(H_0|D)$, è questa ultima probabilità modulata per le probabilità a priori (formula di Bayes).

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)}$$

Formula di Bayes:

- $P(A|B)$: likelihood
- $P(B)/P(A)$: rapporto di probabilità a priori (evidence)

Esempio:

Domanda 1: Qual è la probabilità che una persona venga inserita nell'elenco di Facebook, dato che non è un terrorista?

$$p(\text{f} | \text{t})$$

Domanda 2: Qual è la probabilità che una persona non sia un terrorista, dato che è nell'elenco di Facebook?

$$p(\text{t} | \text{f})$$

Sembra che Facebook su una persona raccoglie talmente tanti dati che la probabilità che si sbagli è 1 su 2000 allora è 99,99%.

$$p(\text{f} | \text{t}) \neq p(\text{t} | \text{f})$$

Se anche possiamo rigettare \$H_0\$ non è detto che le differenze siano praticamente significative, oltre che statisticamente significative.

Ecco perché è stata anche criticato l'approccio Fisher, i dati a volte identificavano delle differenze che però nel mondo reale non sono significative.

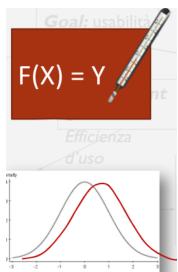
In altre parole: dovete **sempre** interpretare i risultati!

Recap

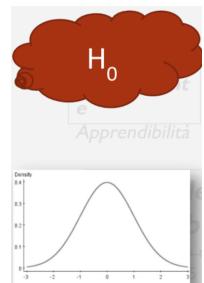
Statistica: ottenere dei risultati statisticamente significative. Un'indagine campionaria, con un campione molto piccolo (30 persone)

⚠ Che cos'è il p-value? - chiesto in esame Il p-value è un concetto fondamentale del mondo moderno, perché si usa un approccio statistico perché si usa un approccio campionario.

Pensate ad una ipotesi (nulla), questa si porta dietro un modello su come dovrebbero essere i dati che raccogliamo, come la distribuzione delle **medie (campionarie)** dei tempi di esecuzione.



Poi si sceglie uno strumento di misura (e relativa metrica). Raccogliete dati sperimentali (Y); questi hanno loro distribuzione



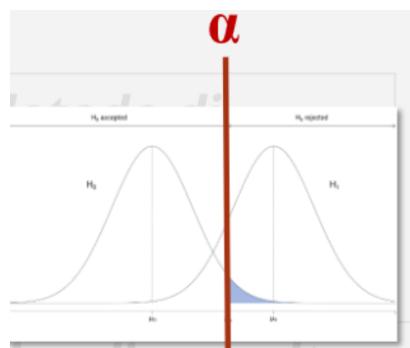
Questi sono i miei dati, che potrebbero avere una distribuzione un po' diversa rispetto a quella immaginata.

$$T(Y) = p(Y | H_0)$$

Poi scegliete un test statistico, questo test “confronta” i dati osservati con quelli ipotizzabili

Calcola la probabilità di avere qui dati (o peggiori) se H_0 è vera

Probabilità dei dati che ho raccolto, data l' H_0 vera



Fissate una soglia tale per cui ritenete i dati troppo strani per essere veri (o se i dati sono veri l’ipotesi infondata)

Posso rigettare H_0 e pensare che la sua alternativa sia più plausibile?

Sì se il P-value è sufficientemente basso (α). E riporto i risultati in maniera chiara e semplice!

Per riportare i risultati al committente, usiamo lo standard APA (American Psychology Association). È un approccio semplice anche per chi non si occupa di statistica. Non è uno standard riconosciuto ufficialmente, ma è diffuso.

La convenzione del gorilla

<https://www.graphpad.com/quickcalcs/ttest1/>

Un t-test per campioni indipendenti è stato eseguito per confrontare [la variabile di interesse, es: il tempo di esecuzione del task 1] usando il [sistema A] e il [sistema B].

-> Che test avete fatto

In base a tale test, si è trovata una differenza [non] significativa per [la variabile di interesse] tra il [sistema A] ($M=[il\ valore\ della\ media]$ +/- [intervallo di confidenza]) e il [sistema B] ($M=[il\ valore\ della\ media]$ +/- [intervallo di confidenza]); $t([gradi\ di\ libertà,\ se\ noti,\ altrimenti\ non\ mettere\ nulla])=[il\ valore\ di\ t,\ solo\ se\ noto]$, $p=[il\ valore\ di\ P,\ esatto]$. !!

-> Che risultato avete ottenuto

Questo risultato suggerisce che gli utenti impiegano un tempo significativamente minore per eseguire il task 1 usando il [sistema A|B] rispetto a quando usano il sistema [sistema B|A].

-> Detto in modo tale che anche il vostro capo lo capisca

Confronta tasso di errori

<http://www.socscistatistics.com/tests/ztest/Default2.aspx>

Uno zeta score test è stato eseguito per confrontare il tasso di errori nel compiere il task 1 usando il [sistema A] e il [sistema B].

In base a tale test, si è trovata una differenza [non] significativa riguardo al tasso di errore tra il [sistema A] (proporzioneA+/- [intervallo di confidenza]) e il [sistema B] (proporzioneB+/- [intervallo di confidenza]); $z\ score=[0]$, $p=[il\ valore\ di\ P\ esatto]$.

Questo risultato suggerisce che gli utenti [non] commettono più errori per eseguire il task 1 usando il [sistema A|B] rispetto a quando usano il sistema [sistema B|A].

Gerd Gigerenzer:

Se riportiamo i P value assoluti, anche in caso di non significatività, e riportiamo gli intervalli di confidenza dei parametri misurati, adottiamo due suggerimenti di Gigerenzer per rendere il "null ritual" più informativo e utile a prendere decisioni.

Esempi di variabili misurabili

Le più comuni:

Table 19.3 Possible measurement criteria (Tyldesley, 1988).

(1)	Time to complete task.
(2)	Percentage of task completed.
(3)	Percentage of task completed per unit time (speed metric).
(4)	Ratio of successes to failures.
(5)	Time spent on errors.
(6)	Percentage number of errors.
(7)	Percentage or number of competitors that do this better than current product.
(8)	Number of commands used.
(9)	Frequency of help or documentation use.
(10)	Time spent using help or documentation.
(11)	Percentage of favourable/unfavourable user comments.
(12)	Number of repetitions of failed commands.
(13)	Number of runs of successes and of failures.
(14)	Number of times the interface misleads the user.
(15)	Number of good and bad features recalled by users.
(16)	Number of available commands not invoked.
(17)	Number of regressive behaviours.
(18)	Number of users preferring your system.
(19)	Number of times users need to work around a problem.
(20)	Number of times the user is disrupted from a work task.
(21)	Number of times the user loses control of the system.
(22)	Number of times the user expresses frustration or satisfaction.

Quali riguardano
Efficienza,
Efficacia,
Soddisfazione?

(Preece, 1994)

Table 19.3 Possible measurement criteria (Tyldesley, 1988).

(1) Time to complete task.	
(2) Percentage of task completed.	
(3) Percentage of task completed per unit time (speed metric).	
(4) Ratio of successes to failures.	
(5) Time spent on errors.	
(6) Percentage number of errors.	
(7) Percentage or number of competitors that do this better than current product.	
(8) Number of commands used.	
(9) Frequency of help or documentation use.	
(10) Time spent using help or documentation.	
(11) Percentage of favourable:unfavourable user comments.	
(12) Number of repetitions of failed commands.	
(13) Number of runs of successes and of failures.	
(14) Number of times the interface misleads the user.	
(15) Number of good and bad features recalled by users.	
(16) Number of available commands not invoked.	
(17) Number of regressive behaviours.	
(18) Number of users preferring your system.	
(19) Number of times users need to work around a problem.	
(20) Number of times the user is disrupted from a work task.	
(21) Number of times the user loses control of the system.	
(22) Number of times the user expresses frustration or satisfaction.	

(Preece, 1994)

Ingegneria dell'usabilità

Usability objective	Effectiveness measures	Efficiency measures	Satisfaction measures
Suitability for the task	Percentage of goals achieved	Time to complete a task	Rating scale for satisfaction
Appropriate for trained users	Number of power features used	Relative efficiency compared with an expert user	Rating scale for satisfaction with power features
Learnability	Percentage of functions learned	Time to learn criterion	Rating scale for ease of learning
Error tolerance	Percentage of errors corrected successfully	Time spent on correcting errors	Rating scale for error handling

E' solo un modo per valutare efficienza, efficacia e soddisfazione.

1. Necessario definire dei task (cf. obiettivo).
2. Equivalente a contare gli errori e farne il rapporto sui tentativi totali (error rate)

Efficienza d'uso

Il sistema deve essere efficiente da usare cosicché, quando l'utente lo ha appreso, sia possibile un alto livello di produttività.

Un modo di misurare l'efficienza è calcolare il tempo necessario ad eseguire il compito correttamente

Gosset 1908: Il tempo è una variabile continua (solitamente con distribuzione normale) e quindi si può applicare la tecnica nota con il lo pseudonimo che ho usato nel 1908 "test T di Student" (grazie a Fisher). Con il mio test potete confrontare i tempi medi di esecuzione

usando il sistema A e quelli ottenuti usando il sistema B (confronto trasversale); o del sistema A a tempo t, e del sistema A' a tempo t, (confronto longitudinale).

Ambito: confronti quantitativi ->

Test di normalità per i più bravi:

<https://goo.gl/wA4Ejc>

Ipotesi nulla

L'**ipotesi nulla** del test del T di Student: non c'è differenza significativa tra le medie di due sotto-campioni (gruppi estratti da un campione di osservazioni indipendenti), assumendo la loro varianza uguale.

Per eseguire il Test T di Student: <http://vassarstats.net/tu.html> (independent/correlated samples) <http://graphpad.com/quickcalcs/ttestl.cfm> (paired/unpaired t test)

Come decido tra paired/unpaired?

Facciamola semplice!

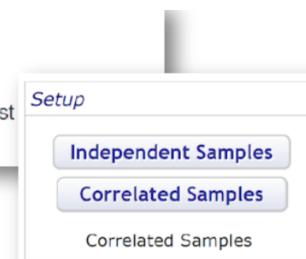
Se usate gli **stessi soggetti**, i campioni sono dipendenti/accoppiati/correlati (es.: confronto tra applicazioni diverse con gli stessi esecutori).

Se confrontate **soggetti diversi** per la stessa applicazione i campioni sono indipendenti/disaccoppiati (confronti tra sottogruppi, es.: femmine vs. maschi per la stessa applicazione).

A intervalli Quantitativa	A rapporti Quantitativa
Misura attraverso applicazione di una unità di misura, conservando le relazioni tra distanze e adottando zero arbitrario. $=, \neq, <, >, +, -$.	Misura attraverso applicazione di una unità di misura, conservando il rapporto tra i valori e adottando uno zero assoluto. $=, \neq, <, >, +, -, \times, /$.
Media, Varianza, Deviazione Standard, Correlazione di Pearson.	Media, Varianza, Deviazione Standard, Correlazione di Pearson, concentrazione.
Diagramma a scatola e bassi; Distribuzione continue; Grafico di dispersione. Test T di Student, ANOVA	Diagramma a scatola e bassi; Distribuzione continue; Grafico di dispersione. Test T di Student, ANOVA

3. Choose a test

- Unpaired t test.
- Welch's unpaired t test
- Paired t test.



 Fare vedere sempre prima un'applicazione che l'altra mette in svantaggio la prima, quindi vanno invertite ogni tanto. - order lies

E per calcolare l'efficienza assoluta del sistema A?

Controntate la performance degli utenti con l'optimal execution time (OET)

Per calcolare l'OET, due di voi faranno 5 o 6 prove di esecuzione del task, in modo corretto e a varie velocità (senza esagerare) riportando i tempi di esecuzione. Calcolerete la media (M) e

deviazione standard (SD). Poi confrontate i tempi di esecuzione del campione di utenti, opportunamente formati. Fate un confronto tra la "vostra" media cronometrica e quella degli utenti.

E se voglio confrontare più di 2 sistemi tra loro?

Allora bisogna applicare un test ANOVA, che generalizza il test t di Student

<http://vassarstats.net/anovalu.html>

Robustezza (Efficacia)

(cf. Efficacia ISO 9241). Il sistema deve indurre un basso tasso di errori (o prevenirne proattivamente alcuni), facili da correggere, e con basso impatto;

Un modo di misurare l'efficacia è calcolare il tass di errore medio dei soggetti del campione quando questi usano il sistema A e confrontarlo con il tasso medio quando gli stessi utenti usano il sistema B (o A').

Il tasso di errore di un soggetto (del campione) è il numero di errori (volte che non raggiunge l'obiettivo del task) messo in rapporto con il numero di tentativi (che può anche essere uno solo per soggetto, "pass or fail").

Il concetto di "errore" è più sottile di quello che si creda. Ad esempio:

1. E' un errore raggiungere l'obiettivo ma attraverso il procedimento sbagliato oppure quello non ottimale?
2. Se l'esaminatore "aiuta" e "corregge" il soggetto, è ancora un errore?

Decidete voi. Se l'utente ha ricevuto una formazione adeguata si può essere severi, altrimenti meglio essere un po' laschi... Per gestire questa ambiguità, si registrano i task eseguiti perfettamente o quasi (verde), i task eseguiti correttamente in maniera non ottimale o con alcuni suggerimenti (giallo) e i task falliti (rosso), cioè gli errori veri e propri. Poi si disegna una matrice task – utenti.

Attenzione alla convenzione colore, è meglio cambiarla per i daltonici

	T1	T2	T3
U1	Red	Yellow	Red
U2	Green	Red	Green
U3	Yellow	Yellow	Green
U4	Red	Green	Yellow
U5	Yellow	Red	Yellow
U6	Green	Yellow	Red

⚠️ Esplicitare la scelta nella relazione, anche se lui consiglia la opzione più morbida visto che abbiamo pochi utenti

Così avete una idea qualitativa di quanto è difficile usare un sistema oltre il mero numero dell'error rate (e capite anche se ci sono utenti particolarmente "difficili"). A proposito, dalla tabella sottostante si evince che l'error rate è uguale per tutti e tre i task (33%).

Ricordate: il confronto si fa tra i sistemi A e B a parità di task, per ciascun task (che vale ovviamente per entrambi)

	T1-SA	T2-SB
U1	Red	Red
U2	Green	Green
U3	Yellow	Green
U4	Red	Yellow
U5	Yellow	Yellow
U6	Green	Red

Tipo variabile di Finalità	Nominale Qualitativa categorica	Ordinale Qualitativa rettilinea o ciclica	A intervalli Quantitativa	A rapporti Quantitativa
	Classificazione	Ordinamento	Misura attraverso applicazione di una unità di misura, conservando le relazioni tra distanze e adottando zero arbitrario.	Misura attraverso applicazione di una unità di misura, conservando il rapporto tra i valori e adottando uno zero assoluto.
Operazioni matematiche di base	=, ≠.	=, ≠, <, >.	=, ≠, <, >, +, ×, ÷.	=, ≠, <, >, +, ×, ÷.
Indicatori statistici (esempi)	Numero di casi (N), Moda.	Media, Quartili, interquartile, Correlazione per ranghi di Spearman	Media, Varianza, Deviazione Standard, Correlazione di Pearson, concentrazione.	Media, Varianza, Deviazione Standard, Correlazione di Pearson, concentrazione.
Rappresentazione dei dati raccolti Test di ipotesi inferenziale (esempi)	Aerogrammi, Istogrammi Test Binomiale, Test del Chi Quadro.	Istogrammi, Diagramma a scatola e bassi; whiskers plot; Test Binomiale, Test del Chi Quadro, Test di Mann-Whitney, Test di Kruskall-Wallis.	Diagramma a scatola e bassi; Distribuzione continua; Grafico di dispersione. Test T di Student, ANOVA	Diagramma a scatola e bassi; Distribuzione continua; Grafico di dispersione. Test T di Student, ANOVA

Se volessi misurare l'efficacia assoluta di A?

Confrontate per ogni task il numero di errori osservato nel campione di utenti con un tasso ottimale (per convenzione il 5% o l'1% - gli errori possono sempre capitare anche a utenti esperti!)

Tasso di errori = booleano: si o no

Il numero di errori è una variabile discreta ed è possibile applicare un test delle proporzioni (o una tabella di contingenza 2x2) per stabilire se il numero di errori di due applicazioni è diverso oltre l'effetto dovuto al caso.

Ipotesi nella del test del Chi-quadro (nella tabella di contingenza 2x2): non c'è differenza tra i valori osservati (e.g., numero di errori effettivamente osservati) e i valori previsti (o in base ad una teoria o in base all'idea che i sistemi siano equivalenti)

Per eseguire il Test Chi Quadro: <http://vassarstats.net/tab2x2.html>

<http://graphpad.com/quickcalcs/contingencyl.cfm>

Il test delle proporzioni https://www.medcalc.org/calc/test_one_proportion.php

Ecco un calcolatore del test del Chi Quadro (Chi squared) un po ' più generale :
<http://www.socscistatistics.com/tests/chisquare2/Default2.aspx>

Per confrontare l'efficacia di più sistemi

Per registrare tempi, codificare errori (ed eventualmente le espressioni del volto) è consigliabile videoriprendere la sessione di test (previa autorizzazione ai sensi della Legge sulla privacy n 196 del 2003) e fare le dovute misurazioni in fase di post produzione



E' consigliabile perchè non vorreste sembrare degli investigatori, ma mettere a vostro agio gli utenti coinvolti. Le misure si fanno dopo con calma visionando le registrazioni. Ricordate: sotto esame NON sono gli utenti, ma le applicazioni informatiche!!

 **Open Broadcaster Software**

Un software GRATUITO per la registrazione video da pc: OBS STUDIO
 Link: <https://obsproject.com/>

Piccola guida per il setup:

1. In SOURCES, click destro -> ADD -> DISPLAY CAPTURE -> selezionare il display, accertarsi che sia spuntata la registrazione del mouse
2. AGGIUNGERE anche la webcam con click destro -> add -> Video Capture Device.
 E' possibile ridimensionare le finestre direttamente dalla schermata di anteprima.
3. Per far partire la registrazione, a destra START RECORDING

☐ Impostazioni importanti:

- ☐ Path di output: Output -> Recording -> Recording path. Meglio se selezionate come formato mp4!
- ☐ Qualità: Output -> Video Bitrate, tenetelo almeno a 1500 (più aumentate, migliore è la qualità e la grandezza del file)
- ☐ Vantaggio: registrare sia quello che avviene sullo schermo che l'utente con la ebcam (e la sua espressione)

Per la lode: con le riprese video è possibile coinvolgere due o tre membri del gruppo che valutino autonomamente quali sono errori e quali no. Poi potete riportare l'Alpha di

Krippendorff (test utile anziché contare le volte perché considera la possibilità che qualcosa siano dovuto al caso). E mettetevi d'accordo!

<https://nlp-ml.io/jg/software/ira/>

Segnare sempre le espressioni facciali avendo più pareri (di 3 persone)

Se i valutatori raggiungono senza mettersi d'accordo una alpha uguale o superiore a 80%, allora avete dati sperimentali di ottima qualità. Sopra il 67% l'accordo (agreement) è comunque buona!

Non confondete l'alpha di Krippendorff con il livello di significatività/probabilità di errore del tipo $\$\\alpha\$$. Questa alpha indica la quantità di "accordo" che non è dovuta al caso!

Soddisfazione.

Il sistema deve essere gratificante da usare.

Misura oggettiva: si fanno misurazioni psicofisiologiche per misurare livelli di stress o di benessere (es. EEG, dilatazione pupilla, battito cardiaco, pressione arteriosa, livello di adrenalina nel sangue)

Un po' difficile....

Misura soggettiva di un osservatore esterno: si conta il numero di espressioni del viso legate a frustrazione, incomprensione, perplessità e quelle invece legate a soddisfazione, piacere, sorpresa positiva

Per semplicità alcuni codificano semplicemente le espressioni positive (verde); espressioni neutre (giallo) ed espressioni negative (rosso). Per fare le cose bene però l'analisi andrebbe fatta da più esaminatori, indipendentemente, e si dovrebbero considerare solo le espressioni codificate allo stesso modo da tutti (o almeno il 75% degli esaminatori), scartando le altre o, se sono tante, valutandole collegialmente

E si ragiona come per gli errori. Si calcola la "smile rate".

A me piace (e vi suggerisco) l'approccio di John Stuart Mill.

Misura soggettiva dell'utente stesso attraverso uno strumento psicométrico, ad esempio un questionario con item a scale ordinali (tipo Likert oppure a differenziale semantico) e un uso preparato e responsabile delle tecniche statistiche opportune

“...la sensibilità e il giudizio di coloro che ne hanno fatto esperienza”

Scala Likert

Scala Likert, di Rensis Likert: Una scala ordinale su cui il rispondente mappa il suo grado di accordo con una affermazione proposta dallo sperimentatore (possibilmente termini neutri).

1. fortemente contrario
2. contrario
3. parzialmente contrario
4. parzialmente d'accordo
5. d'accordo
6. molto d'accordo
99. Non saprei

Website User Survey

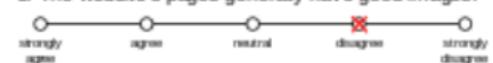
1. The website has a user friendly interface.



2. The website is easy to navigate.



3. The website's pages generally have good images.



4. The website allows users to upload pictures easily.



5. The website has a pleasing color scheme.



Una scala Likert non è per forza da 6, è molto diffusa per esempio quella a 7 valori.

Il numero 99 è un numero usato come convenzione per una non-opinione.

Priming: il tipo di sfumatura a cui sono stati esposti gli utenti con la affermazione di riferimento Esempio: priming positivo “Le macchine Fiat sono belle” priming negativo “Le macchine Fiat sono brutte”

Il priming può influenzare un po' la opinione, perché imposta la idea di partenza.

Come affermazioni iniziali usare il priming che ci si aspetta, ma non polarizzare troppo le affermazioni (nè renderle neutrali). Esempio sbagliato “Le macchine Fiat sono le migliori al mondo”.

Pro/Contro

Contro:

1. Misura l'accordo del soggetto con affermazioni dello sperimento: ci può essere un forte l'acquiescence bias

“tendenze all'accordiscendenza”: bias dovuto alla tendenza di certi rispondenti a preferire la connotazione positiva delle opzioni fornite, ad essere d'accordo con le domande dei ricercatori e ad “aiutarli” (simile al social desirability bias)

2. Se si adotta un numero **dispari** di opzioni ci può essere anche un forte central tendency bias

“Tendenza alla risposta centrale o neutra”: bias dovuto alla tendenza di certi rispondenti a preferire la opzione di mezzo, perchè meno impegnativa, più veloce (non ci si deve pensare), e non richiede di prendere una posizione

E' un rifugio per chi non si vuole schierare

Pro:

1. Molto nota, scala familiare a molti potenziali soggetti.
2. Semplice da capire, veloce da rispondere

Proposta del prof:

Numero pari così chiediamo di schierarsi sempre, ma c'è un “non saprei” “non so” al di fuori della scala.

Si può considerare le opzioni centrali con unica, per avere una compatibilità con l'opzione dispari. Quindi una scala pari è compatibile con una dispari, ma non viceversa. Questo rende la scala con opzioni pari un po' più ricca di informazioni e quindi migliore.

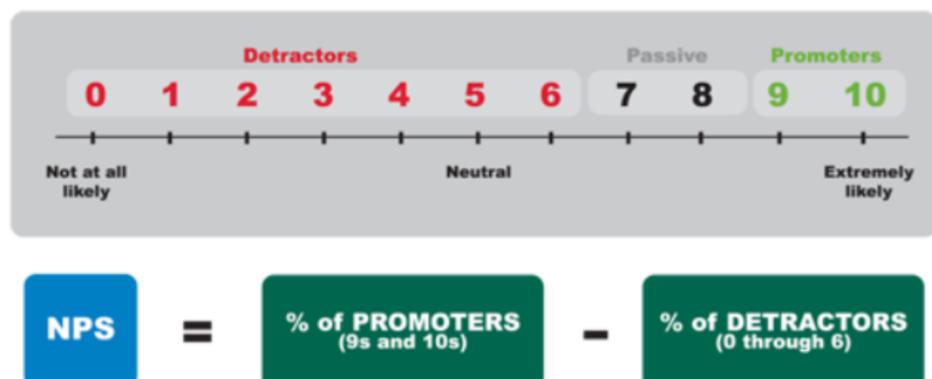
Net promote score (NPS)

Tratta la domanda giusta da fare quando non si ha tempo per un questionario lungo e più o meno complesso. I risultati rispecchiano spesso le tendenze totali di questionari anche complessi, quindi è diventato uno strumento importante e utilizzato.

Promotori: rispondenti che scelgono 9 o 10. Sono le persone entusiaste, che consigliano candidamente il vostro prodotto agli altri.

Passivi: rispondenti che scelgono 7 o 8. Sono persone soddisfatte ma non eccessivamente entusiaste che potrebbero essere tentate da altri fornitori migliori di voi e difficilmente potrebbero cantare le lodi del vostro prodotto ad amici e conoscenti

Detrattori: chi risponde scegliendo un valore tra 0 e 6 (incluso!). Sono clienti che non manifestano soddisfazione per la vostra soluzione o prodotto e anzi potrebbe spargere la voce in modo negativo, sconsigliando il vostro prodotto



Il suo grafico preferito è lo stile contachilometri perché più immediato, con la barra sotto che dà le informazioni in più

Pro/Contro

Contro:

1. Misura solo la raccomandabilità
2. Può essere confuso con un voto da 0 a 10 (in Italia)

<http://www.datarevelations.com>



<https://thermostat.io/>

Pro:

1. La raccomandabilità è molto importante!
2. Considera l'acquiscence bias ("6" è con i detrattori)
3. Non considera come cardinali delle categorie ordinali (normalizza una differenza si
numerosità)
4. È semplice e permette confronti nel tempo

02/05/2022

Esempi di dark patterns:

1. Bait and switch: l'utente si abitua ad una certa conseguenza relativa ad una certa interazione, ma ad un certo punto la stessa interazione cambia conseguenza
2. Confirmshaming: far sentire in colpa l'utente per aver scelto un'opzione
3. Disguised ad: le pubblicità sono nascoste nella pagina, come se fossero parte del "normale" contenuto
4. Forced continuity: quando il periodo di prova gratuito finisce, l'abbonamento si rinnova automaticamente addebitando il prezzo
5. Friend spam: il sistema spamma i contatti o gli amici dell'utente "fingendo" di essere l'utente stesso
6. Hidden costs: solo all'ultimo step per l'acquisto vengono aggiunti alcuni costi
7. Misdirection: porta l'attenzione dell'utente in uno specifico posto per non fargli notare altro
8. Price comparison prevention: il venditore rende difficile per l'utente comparare il prezzo con quello di un altro prodotto, in modo da non fargli fare una decisione informata
9. Privacy zuckering: sei "ingannato" nel pubblicare più informazioni di te stesso di quante avessi davvero intenzione di condividerne
10. Roach motel: il design rende semplice finire in certe situazioni, ma difficile uscirne (es: iscrizione)

11. Sneak into basket: quando aggiungi un prodotto al carrello, ne vengono in automatico aggiunti altri "correlati"
12. Trick questions: domande "sbagliate", in cui non si capisce bene quale opzione scegliere per ottenere un certo risultato

03/05/2022

QUESTIONARI PSICOMETRICI

Quale scegliere è una questione di preferenze e di finalità. Per il progetto d'esame si può chiedere al prof quale questionario sia meglio somministrare per il nostro campione.

Usati più frequentemente nella valutazione della soddisfazione percepita:

- SUMI
- AttrakDiff
- UEQ à preferito del prof
- SUS à non piace al prof

SUMI (Software Usability Measurement Inventory): metodo per misurare la qualità di sistemi interattivi "dal punto di vista dell'utente", cioè l'"quality of use" (ovvero l'usabilità).

Misura la QoU in 5 sottodimensioni:

- Efficiency (efficienza)
- Affect (affezione)
- Helpfulness (utilità)
- Controllability (controllabilità)
- Learnability (apprendibilità)

Il SUMI è molto lungo, quindi l'utente è soggetto al bias di affaticamento (solitamente dopo 12 minuti).

Variabili di stratificazione

- Genere
- Età: si può chiedere l'età precisa
- Familiarità con il software: si può chiedere con una scala da 1 a 6 e dividere in due (quelli da 1 a 3 e quelli da 4 a 6)

Questi item possono essere considerate variabili indipendenti per creare gruppi tra gli utenti.

Potete calcolare le correlazioni tra questi item e altri singoli item. Oppure dicotomizzare le risposte, creare due gruppi e poi applicare un test di Mann-Whitney, che è un test non parametrico per confrontare la mediana delle risposte date da 2 gruppi diversi di rispondenti.

Pro:

- Facile da usare (solo tre attributi ordinali)
- Risultati facili da comunicare (proporzioni, non medie e test binomiali per le statistiche inferenziali)
- Usato in molti contesti (grande comparabilità)
- Ha spazio per domande aperte e un po' di feedback dall'utente
- Suggerito dall'ISO 9241 per testare la "user satisfaction"

Contro:

- Non facile trovare la traduzione italiana validata
- Ci vuole una licenza per poterlo usare in contesti istituzionali o aziendali
- Lunghetto (ai limiti della soglia del fatigue bias)
- Richiede un uso abbastanza prolungato da parte dell'utente

AttrakDiff: questionario per valutare la UX, diviso in 4 ambiti:

- Qualità pragmatica (PQ): riguarda l'usabilità in termini di utilità ed efficacia
- Qualità edonica (HQ): riguarda più l'esperienza estetica, relativo al bello e a ciò che piace, divisa in

Identità (HQ-I): riguarda l'identificazione dell'utente con il prodotto (o la marca) in un dato contesto sociale

Stimolazione (HQ-S): riguarda la soddisfazione di bisogni in termini di voglia di nuovo (anche nuove informazioni), desiderio di padroneggiare uno strumento, e di interazione stimolante

- Costrutti di valutazione: riguardano più esplicitamente “il bello e il buono”

Pro:

- Misura dimensioni UX difficili da catturare
- Usa scale a differenziale semantico
- Ha una rappresentazione grafica appetibile (anche detto snake diagram)

Contro:

- Non minimizza il central tendency bias (ha 7 posizioni)
- Si basa su concetti “semantici” non di ovvia interpretazione
- Manca traduzione italiana validata
- Come indicatore quantitativo adotta la media (per singola dimensione) delle medie (dei singoli rispondenti) trasformando posizioni sulla scala in numeri)

UEQ (User Experience Questionnaire): questionario di 26 item (differenziale semantico) e 6 sottoscale:

- Attractiveness: attraente, piacevole
- Perspicuity: facile da capire e imparare
- Efficiency: efficienza, velocità
- Dependability: ci si può affidare, predicibile
- Stimulation: eccitante, interessante
- Novelty: creatività, innovatività

Correlazione con il tempo di completamento del task:

Negativa se tempo e proprietà sono inversamente proporzionali

Gli asterischi sono messi quindi a fianco di risultati che potrebbero variare in base al campione misurato (quindi in base al p-value):

*: p-value tra 1% e 5%

**: p-value tra 1‰ e 1%

***: p-value inferiore a 1‰

Pro:

- Breve e con parole più semplici del AttrakDiff
- Copre aree semantiche importanti
- Ha traduzioni validate in molte lingue
- Ottima documentazione: <http://www.ueq-online.org/>
- Fornisce anche un tool di analisi dei dati (MS Excel)
- Fornisce un data set per fare benchmark

Contro:

- Non minimizza il central tendency bias (7) à possiamo farle diventare 8 o 6 (se facciamo 8 possiamo contare i 2 valori centrali come se fossero un solo valore, per renderla compatibile con quella originale, ma potendo comunque capire la tendenza positiva o negativa)
- Come indicatore quantitativo adotta la media (per singola dimensione) delle medie (dei singoli rispondenti) trasformando posizioni sulla scala

SUS (System Usability Scale): ha solo 10 item

→ Nel 10 la traduzione “processi” non è troppo azzeccata, potrebbe essere procedure

→ La 5 e la 6 sono domande di controllo, se queste due sono in contrasto il questionario viene scartato, in quanto le risposte sono in contrasto tra loro

Per calcolare lo score SUS:

1. Per gli item dispari (1, 2, 3, 5, 7 e 9) il contributo allo score è dato dal valore scelto dall’utente meno 1. Per gli item pari (2, 4, 6, 8 e 10), il contributo è dato dalla differenza tra 5 e il valore scelto dall’utente. Questo converte tutte le risposte in un range da 0 a 4.
2. Somma tutte le risposte di un singolo utente e moltiplica il risultato per 2.5 per ottenere il valore totale per quell’utente (il risultato varia tra 0 e 100).

3. Fai poi la media tra tutti gli score individuali.
4. Interpreta: uno score SUS medio sopra il 68 sarebbe da considerarsi sopra la media (risultato buono).

⚠ I valori degli item vanno trasformati (i dispari decrementati, i pari anche “girati”) **subito**, prima di riportare i dati in formato tabellare, visualizzarli come boxplot e fare il calcolo dello score!

Lo score SUS è un “numero” un po’ discutibile, però il questionario può essere utile per confrontare i sistemi, anche considerando singolarmente gli item (soddisfazione – 1, facilità d’uso – 3, efficienza – 8, learnability – 7). Per i singoli item si possono fare test binomiali e di Mann-Whitney.

Pro:

- Breve e veloce da compilare
- Da un numero tra 0 e 100 (piace al cliente)
- Molto noto e utilizzato
- Minimizza l’option order bias e l’acquiescence bias

Contro:

- Somma valori ordinali, normalizza la somma per ciascun rispondente in centesimi e poi fa la media tra tutti i rispondenti
- Mescola usabilità e apprendibilità
- Non si possono fare confronti statisticamente significativi (a meno di non isolare singoli item)

Euristiche del form

à Gli slider possono essere utilizzati se si vogliono gestire valori continui

à L’effetto “prime” può valere al contrario se le opzioni non sono numeriche, perché l’utente (se le legge tutte) si ricorda l’ultimo valore letto

La distribuzione di risposte su scala ordinale può essere caratterizzata da moda, mediana e somma di ranghi (non dalla media!). La visualizzazione migliore è, per ogni item, quella dei boxplot “con i baffi”.

<http://shiny.chemgrid.org/boxplotr/>

Questo sistema permette anche di rappresentare i violinplot, usati per i dati continui (es: per la distribuzione dei tempi di risposta).

Per confrontare le mediane tra due gruppi di risposte:

- Test di Mann-Whitney: se campioni indipendenti

<http://vassarstats.net/utest.html>

<http://www.socscistatistics.com/tests/mannwhitney/default2.aspx>

- Wilcoxon test: se campioni dipendenti

<http://vassarstats.net/wilcoxon.html>

<http://www.socscistatistics.com/tests/signedranks/Default2.aspx>

Ipotesi nulla dei test Mann-Whitney/Wilcoxon: le distribuzioni dei punteggi (i mean ranks) raccolti nel campione per la variabile dipendente (es: soddisfazione) nei due gruppi generati a partire da due valori della variabile indipendente (es: maschi vs femmine, sistema_tipo_1 vs sistema_tipo_2) hanno la medesima forma (cioè provengono dalla medesima popolazione) o, assumendo che la forma sia veramente la stessa, la medesima mediana (non parliamo di media!).

: Un **test Mann-Whitney U test** è stato eseguito su N partecipanti per determinare se ci fossero differenze significative per quanto riguarda [la dimensione X] percepita usando il [sistema A] e il [sistema B] (oppure tra [gruppo 1] (N_1 = numero partecipanti gruppo 1) e [gruppo 2] (N_2 = numero partecipanti gruppo 2) usando un solo sistema).

In base a tale test, si è trovata una differenza [non] statisticamente significativa tra le percezioni mediane relativamente alla [dimensione X] per il [sistema A] (o il gruppo 1) (MedianA) e il [sistema B] (o il gruppo 2) (MedianB), U = [valore di U] p = [valore esatto di P].

Questo risultato suggerisce che gli utenti percepiscono una [la dimensione X] significativamente minore usando il [sistema A/B] rispetto a quando usano il [sistema B/A] (oppure: suggerisce che [i membri del gruppo 1] percepiscono una [la dimensione X] significativamente più bassa dei [membri del gruppo 2] usando il [sistema A]).

→ Se il p-value è il 0.9%, sarebbe statisticamente significativo anche con un livello di confidenza del 99% (sì, perché comunque rigetto l'ipotesi nulla, dato che aplha è <)

→ Per convenzione 0.4 si riporta .4

 Un **test** (dei ranghi con segno) **di Wilcoxon** è stato eseguito su N partecipanti per determinare se ci fossero differenze significative per quanto riguarda [la dimensione X] percepita quando si usa il [sistema A] e il [sistema B].

In base a tale test, si è trovata una differenza [non] statisticamente significativa tra le percezioni mediane relativamente alla [dimensione X] per il [sistema A] (o il gruppo 1) (MedianA) e il [sistema B] (o il gruppo 2) (MedianB), Z = [valore di Z] p = [valore esatto di P].

Questo risultato suggerisce che gli utenti percepiscono una [la dimensione X] significativamente minore usando il [sistema A/B] rispetto a quando usano il [sistema B/A] (oppure: suggerisce che [i membri del gruppo 1] percepiscono una [la dimensione X] significativamente più bassa dei [membri del gruppo 2] usando il [sistema A]).

Recap:

Per la sola dimensione della soddisfazione:

1. Selezionate un questionario che misuri la “soddisfazione”, motivando la scelta (semplicità, completezza, diffusione, benchmarking) à mettere quindi i vantaggi del questionario scelto
2. Implementatelo su una piattaforma online (Google forms, Limesurvey, SurveyMonkey), aggiungendo 3 domande di profilazione: genere, classe di età e familiarità con tecnologie Web
3. Fatelo compilare ai rispondenti (almeno 24)
4. Riportate tutte le risposte in formato tabellare
5. Visualizzatele sotto forma di diagramma Boxplot o Violin Plot
6. Riportate gli score, se previsti dal questionario, per ciascun sistema ed eventualmente sottogruppi di rispondenti (es: junior vs senior)

7. Selezionate 2 o 3 item rilevanti dal questionario e, per ciascuno di questi, eseguite test Mann-Whitney (o Kruskall Wallis) per identificare eventuali differenze statisticamente significative
8. Riportate il risultato del test in maniera chiara e semplice

06/05/2022

CONCETTI AVANZATI

SISTEMI E INTERFACCE D'USO

Sviluppatore:

- Oggettivista: ritiene che la realtà possa essere descritta in un certo modo, e che queste descrizioni informano la progettazione dei sistemi informatici in maniera diretta
- Soggettivista / costruttivista: ritiene che il significato delle rappresentazioni non possono essere definite a priori, o non esistono in maniera oggettiva, ma sono coprodotte da utenti, programmatore e uso (contesto, modo d'uso...)

Sommario dei concetti avanzati:

1. Gli esseri umani sono esseri sociali (**psicologia**) e tendono a considerare le macchine come degli attori sociali (teoria **CASA**, Computers as Social Actors)

In particolare, le persone non pensano di essere messi in comunicazione con gli sviluppatori tramite esse

2. Invece i programmatore dovrebbero ragionare diversamente: le macchine sono il loro strumento di comunicazione con gli utenti (**Semiotic Engineering**)

Ma perché gli sviluppatori dovrebbero comunicare con gli utenti?

3. Per convincerli a fare qualcosa! **Captology!** (Computers as Persuasive Technology)

PSICOLOGIA

Gerarchia dei bisogni:

8 meccanismi di cui essere consapevoli:

1. **Riflesso pavloviano**: gli esseri umani nei loro comportamenti esprimono qualcosa di simile agli altri animali.

Dark patterns: sfruttano gli automatismi che si instaurano nella nostra mente.

2. **Effetto placebo** (cf feedback): alcune cose ci piacciono automaticamente.

Associamo un determinato gesto ad un feedback positivo (di rinforzo).

3. **Effetto camaleonte/specchio**: mettersi nei panni dell'utente e fargli da "specchio".

Un'interfaccia empatica viene apprezzata di più.

4. **Effetto alone** (halo effect): effetto che condiziona l'opinione di qualcuno sulla base di caratteristiche superficiali.

L'estetica dell'interfaccia influenza di molto l'opinione dell'utente.

5. **Effetto isolamento** (Von Restorff effect): è importante che le notifiche attirino l'attenzione dell'utente.

6. **Priming** (legato al bias di ordine e con l'effetto della posizione seriale): la probabilità di scegliere le opzioni centrali è bassa.

7. **Effetto di familiarità** (o di semplice esposizione): la splash screen crea un ambiente "familiare" per l'utente, ed è per questo che le applicazioni tendono a copiare tra di loro gli elementi.

8. **Gratificazioni e incentivi**: importante alimentare l'utente con premi e punizioni.

CASA (Computer as Social Actors)

È molto comune parlare con le macchine.

Alcuni esperimenti tra gli anni 90 e 2000 hanno dimostrato che gli utenti sono indotti a comportarsi come se le macchine fossero umane, anche se sanno benissimo che non hanno coscienza o non agiscono mosse da motivazioni tipicamente umane.

L'attribuzione di attitudini, interazioni e disposizioni umane alle macchine (e agli oggetti in particolare) si chiama **ethopoeia** (attribuzione di comportamenti umani alle cose, diversa dall'antropomorfizzazione, che invece guarda la forma, l'aspetto estetico).

Basta poco per indurre l'ethopoeia:

1. Messaggi in linguaggio testuale (words for outputs)
2. Risposte basate su input predefiniti multipli (interactivity)
3. Attribuzione di ruoli tipicamente umani (es: assistente)
4. Utilizzare voci umanoidi
5. Utilizzare volti umanoidi

Alcuni comportamenti sociali:

- **Reciprocità nella cortesia:**

La valutazione di un'applicazione non andrebbe mai fatta nel contesto della medesima applicazione (perché l'utente tende ad essere più "gentile" con l'applicazione che l'ha aiutato a raggiungere il suo obiettivo).

Gli utenti si aspettano educazione e cortesia: i messaggi di errore devono essere educati e non troppo bruschi.

Attenzione alla semplice traduzione, altrimenti all'utente può sembrare che l'applicazione sia "stupida".

- **Reazione a personalità diverse** (dominante/remissivo):

Gli utenti sono in grado di identificare messaggi che indicano dominanza o remissività.

Gli utenti sono in grado di avvertire somiglianze tra la personalità dell'applicazione e la loro.

Gli utenti preferiscono interagire con computer che esibiscono la stessa tipologia di personalità (il simile apprezza il simile).

Programmi denotati come di informazione sono stati considerati più “informativi” degli stessi programmi denotati come “di informazione e intrattenimento”.

- **Adulazione:**

→ Molti sistemi sono normalmente severi o neutri nel loro feedback. Potrebbe essere una buona idea addolcire il feedback, anche sfiorando l’adulazione, soprattutto nel caso di applicazioni tutorial o per la formazione, in quelle per aumentare la creatività o che impongono compiti ripetitivi e noiosi (se l’utente viene adulato (anche essendone consapevole) si sente più soddisfatto di aver utilizzato l’applicazione).

- **Stereotipi legati al genere:**

Applicazioni che hanno usato voci femminili per impartire istruzioni in tono perentorio (dominanza) sono state valutate più negativamente delle stesse applicazioni che usavano voci maschili.

Applicazioni di tutoring che impiegavano voci maschili sono state valutate come più “competenti” delle stesse applicazioni con voci femminili.

Applicazioni che davano suggerimenti in tema di amore e rapporti umani con voci femminili sono state considerate migliori delle stesse applicazioni con voci maschili.

→ Purtroppo, i pregiudizi sono meccanismi umani potentissimi e non si limitano solo al genere (anche età, accento straniero o fortemente connotato regionalmente): bisogna tenerne conto o sfidarli apertamente?

- **Senso di appartenenza**

→ Il comportamento sociale è rivolto al programmatore? Gli utenti considerano i programmi come parte del computer, e non come prodotto di programmatori, e questo rende agli occhi dell’utente il sistema più “usabile”.

10/05/2022

SEMEIOTICA

→ Studio della comunicazione umana

Decodifica aberrante (studiata da Umberto Eco): a prescindere dalla chiarezza del massaggio, c'è comunque un certo grado di fraintendimento, che non può essere completamente eliminato.

Semeiotica = scienza dei **segni**

I segni devono essere studiati nel contesto in cui sono scambiati, ovvero nei "sistemi segnici". Studiare i segni significa studiare le modalità con cui essi sono costruiti.

→ La semeiotica è lo studio di come viene costruito il significato e dell'efficacia della comunicazione.

La semeiotica può essere divisa in 3 ambiti:

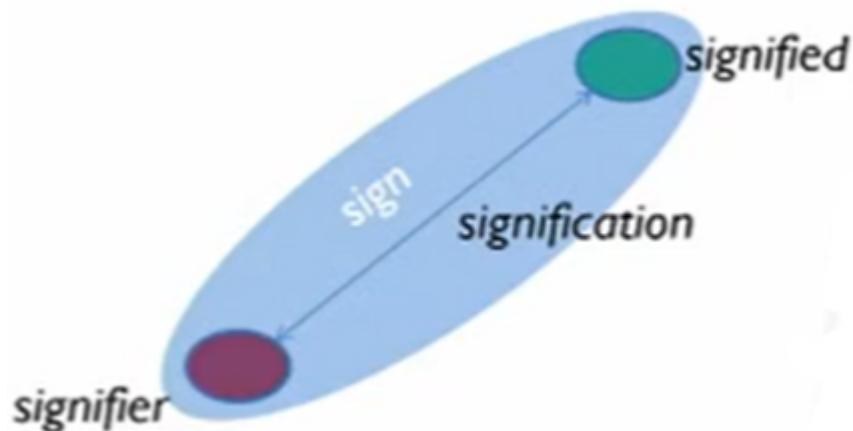
- **Sintassi:** studio delle relazioni che sussistono tra i segni (spesso sotto forma di strutture formali e regole logiche) à forme
- **Semantica:** studia le relazioni tra i segni e ciò a cui si riferiscono à significati
- **Pragmatica:** relazioni che sussistono tra i segni e i loro effetti (ciò che effettivamente avviene nella pratica umana) à azioni intraprese dall'uomo

→ **Segno:** qualunque cosa percepibile che ha un significato per qualcuno in un certo contesto sociale.

I segni sono dappertutto (simboli, parole, suoni, linguaggio del corpo, ma anche odori, sapori, atteggiamenti, posizione degli oggetti nello spazio). Un segno è tale solo quando ha un significato per qualcuno.

Una persona che ragiona semioticamente cerca di capire qual è il modo migliore di usare i segni in maniera efficace.

Saussure:



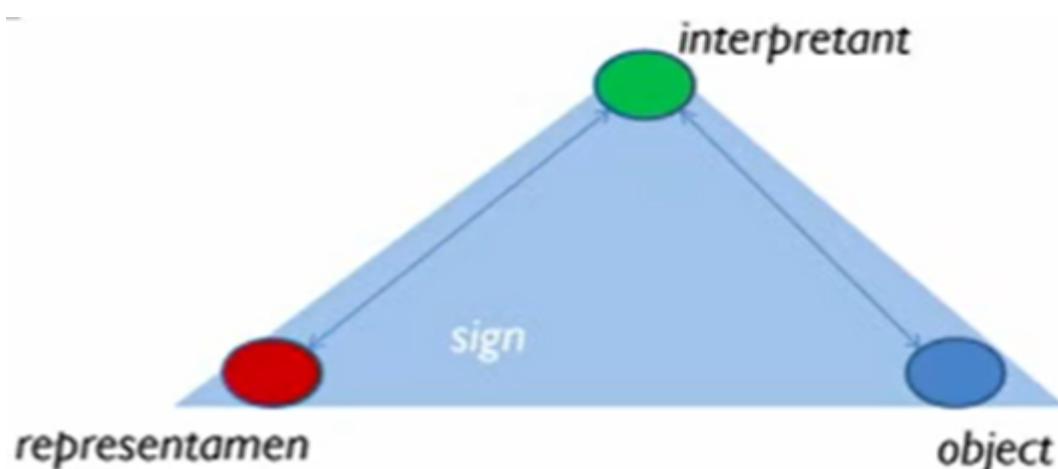
Nella semeiotica un segno emerge dalla correlazione tra un significante (signifier) e un significato (signified).

Il significante è interpretato come qualcosa di percepibile (o fisico) con uno dei 5 sensi.

I segni hanno relazione se sono messi in relazione tra di loro.

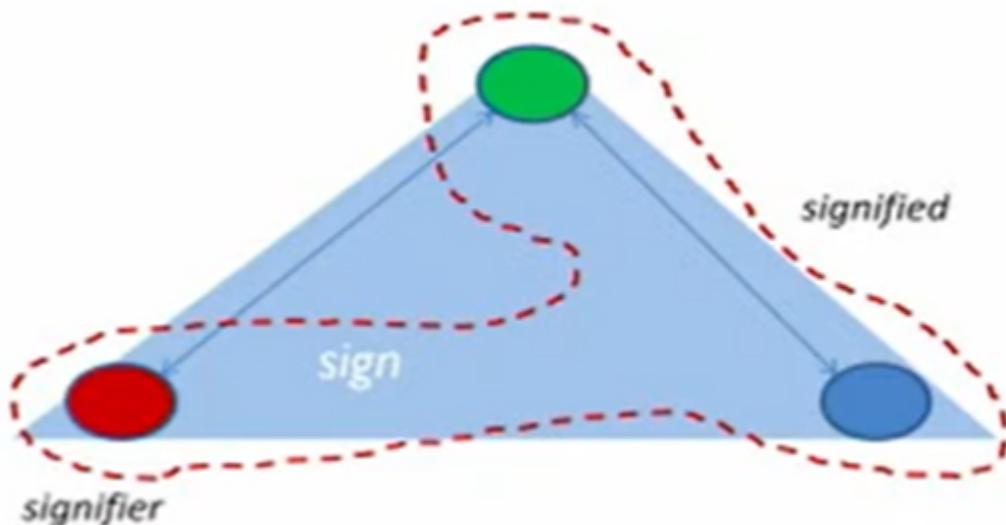
Peirce:

→ Un segno è qualcosa che sta per qualcuno sta per qualcos'altro rispetto a qualche aspetto o capacità.



- Significante → rappresentante
- Significato → interpretant
- Oggetto → cosa a cui ci si riferisce con il segno

Per Peirce il segno è un'unità di ciò che viene rappresentato (oggetto), di come viene rappresentato (rappresentante) e di come è interpretato (interpretant).



Il significato di una rappresentazione non può essere altro che un'altra rappresentazione SSSS(semeiotica infinita).

Semeiotica è il processo in cui noi ragioniamo **abduittivamente**: noi generiamo ipotesi alla luce delle quali i segni hanno un certo significato e se le cose funzionano prendiamo queste ipotesi come principi generali per interpretare tutto ciò che segue, se invece queste ipotesi iniziali non funzionano formuliamo nuove ipotesi.

L'abduzione spiega il “**paradosso dell'utente attivo**”: gli utenti preferiscono dedicare più tempo in prove d'errori piuttosto che leggersi il manuale.

→ I progettisti devono progettare semioticamente.

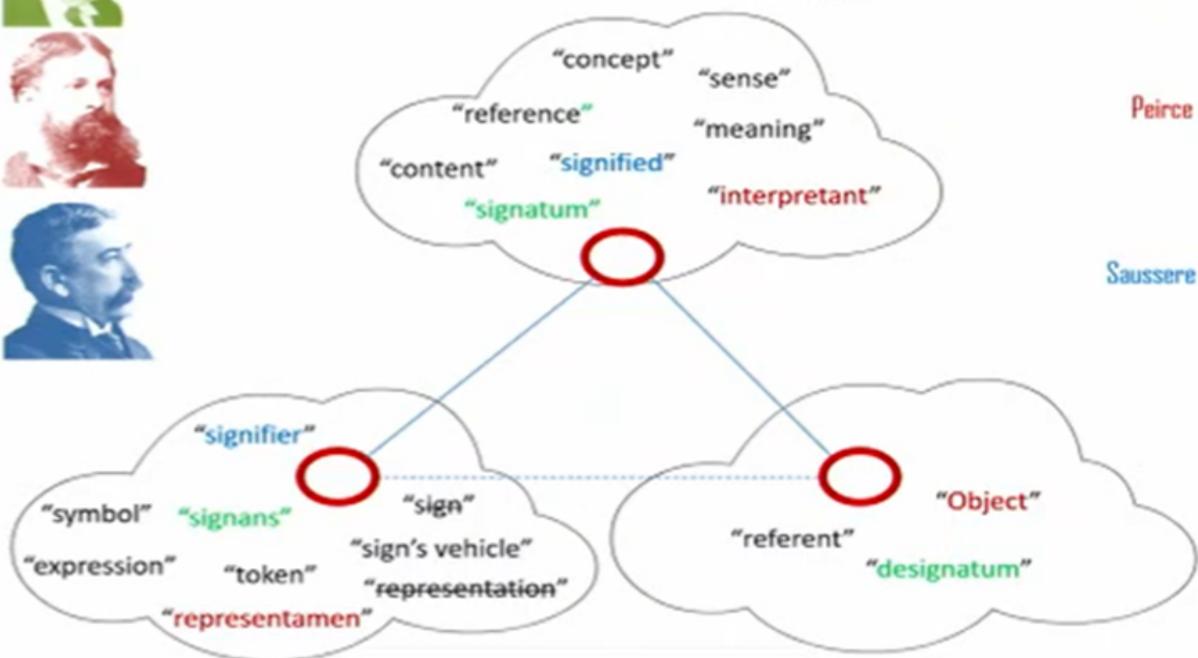


Semiotic Triangle

Jakobson

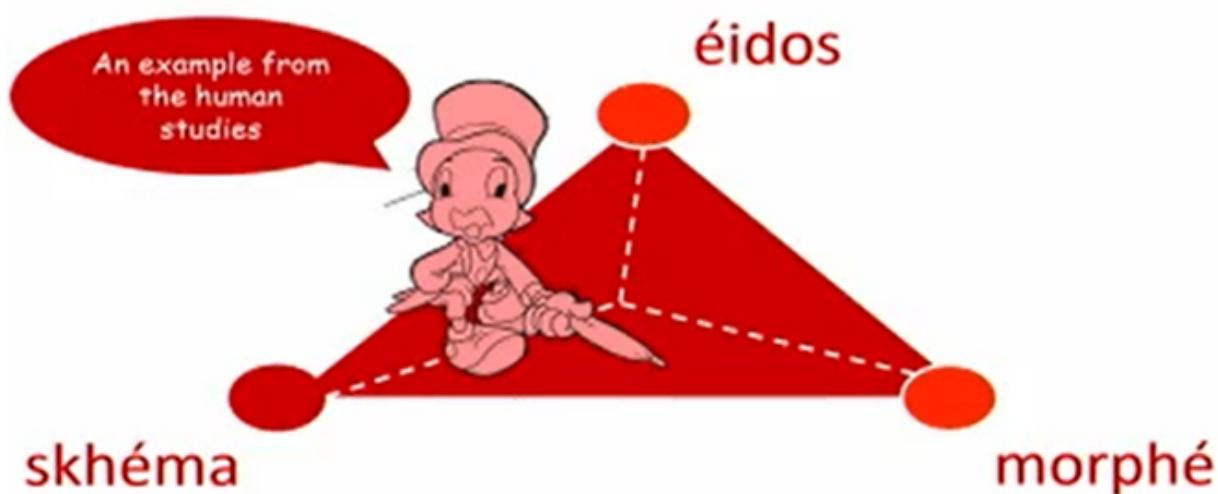
Peirce

Saussure

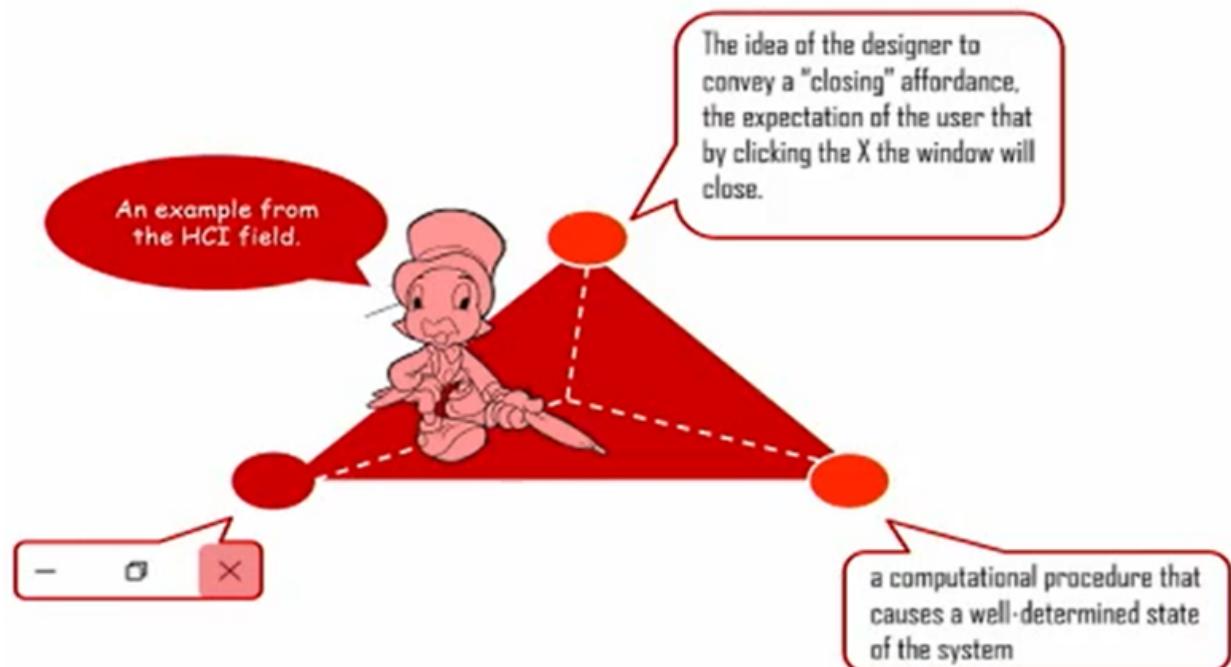


Signifier (o rappresentante): forma sotto la quale il segno appare (es: parola scritta o pronunciata), dove con segno si intende l'intero insieme che ha un suo significato.

Triangolo semeiotico



Éidos (interpretant), morphé e skhéma significano forma in greco.



- Simbolo: DOG
- Icona: foto del cane
- Indice: guardiamo dove guarda il cane, lui ci sta “indicando” dove guardare



- **Simbolo** (o segno simbolico): modalità in cui il significante non assomiglia al significato, ma in cui il significante è fondamentalmente arbitrario o del tutto convenzionale, quindi la relazione tra significante e significato dev'essere appresa, concordata all'interno di un certo contesto sociale.

Tutto ciò che è linguaggio possiamo dire che è segno simbolico, quindi le lettere dell'alfabeto, segni per funzioni, parole, frasi, numeri, codici morse, semafori, bandiere nazionali.

- **Icona:** modalità in cui il significante è percepito come somigliante o che imita il significato.

Un'icona può essere una onomatopea, qualcosa che appare simile, che ha lo stesso odore o sapore, che condivide qualche caratteristica (come una caricatura), un modello in scala.

- **Indice** (o indicale): modalità in cui un significante, non in modo arbitrario, ma collegato direttamente (fisicamente o causalmente) al significato.

Sono indici i segni naturali (fumo per il fuoco, le impronte, il fulmine e il tuono), i puntatori (la mano che indica, il cartello segnaletico), il suono del cellulare.

L'indicalità è basata su una connessione diretta o sulla vicinanza.

Lo stesso segno può essere una combinazione qualunque di simbolo, icona e indice.

Il semaforo sotto, ad esempio, è sia simbolico che iconico.

