

DaCENA

Documento di usabilità e testing

Nome gruppo: SOKE

Numero assegnato: 24

Componenti del gruppo:

761678 Angiolillo Andrea

761702 Belingheri Omar

761763 Khayam Adam

Scopo del documento

Questo documento serve per mostrare nel dettaglio come sono stati fatti i nostri studi sugli utenti che interagiscono con l'applicazione, oltre che per giustificare le scelte di design da essi derivanti. Come in ogni caso di user-centered design, i test sono stati realizzati durante tutto il processo di sviluppo, e quelli che effettuiamo in questa prima iterazione serviranno sicuramente a raccogliere del feedback per migliorare l'interfaccia per la seconda consegna.

Obiettivi di usabilità

Gli obiettivi dei nostri test sono principalmente i seguenti:

1. valutare se gli utenti siano in grado di muoversi con facilità all'interno dell'applicazione raggiungendo tutte le funzioni principali, senza bisogno di aiuto e in maniera autonoma;
2. individuare quali ostacoli si pongono tra gli utenti e il raggiungimento dei loro obiettivi (ad esempio nello svolgere una determinata funzione); eventuali ostacoli saranno affrontati e risolti nella seconda iterazione;
3. misurare le prestazioni degli utenti durante l'uso dell'applicazione;
4. assicurarsi che l'estetica dell'applicazione soddisfi gli utenti.

Per effettuare dei test che coprano in maniera esaustiva le varie funzionalità del software, si escogiteranno degli scenari di situazioni comuni che gli utenti si trovano ad affrontare durante l'utilizzo dell'applicazione. Gli utenti verranno quindi osservati mentre svolgono queste operazioni e si raccoglieranno le loro opinioni.

Errori di progettazione

Le diverse difficoltà che possono sorgere per gli utenti durante i test possono essere riassunte come segue:

1. **errori di navigazione:** è impossibile trovare e accedere alle funzioni del sistema, inclusi casi di eccessiva lunghezza del flusso (per esempio, una determinata funzione è troppo difficile da raggiungere);
2. **errori di presentazione:** è impossibile trovare le informazioni desiderate, oppure le informazioni sono ambigue o non chiare;
3. **problemi di utilizzo dei controlli:** pulsanti, caselle di testo, o altri elementi dell'interfaccia non funzionano come dovrebbero.

Metodologia

Nella prima iterazione utilizzeremo un solo utente per i test di usabilità, con l'obiettivo di aggiungerne altri per la consegna finale. Secondo uno studio di Virzi¹ si possono individuare l'80% dei problemi di usabilità usando solamente 4 o 5 utenti. Questo è possibile scegliendo utenti che appartengano a categorie eterogenee, ovvero utenti con background diversi, con diverse competenze informatiche, ecc.

DaCena è un software che si rivolge sicuramente a un pubblico di giornalisti e/o blogger, i quali possono usare l'applicazione per reperire più (e migliori) informazioni, ma in generale può essere utilizzato da chiunque legga news online per facilitare la comprensione dell'articolo che sta leggendo. Per questo abbiamo individuato le seguenti categorie di utenti:

1. utenti interessati al mondo del giornalismo/blogging, con precedenti esperienze nella cura di un blog/sito;
2. utenti che leggono spesso (tutti i giorni) news online, senza pubblicarne;
3. utenti che leggono saltuariamente (due volte a settimana o meno) news online, senza pubblicarne.

Ci aspettiamo che un utente nel gruppo 1 abbia meno problemi ad interagire con l'interfaccia, essendo un utente più "esperto". Gli utenti del gruppo 3 sono invece quelli dai quali ci si aspetta più feedback utile, in quanto dovrebbero essere quelli con meno esperienza e quindi quelli che potrebbero incontrare più difficoltà.

Procedura

L'interfaccia grafica è stata realizzata da noi scrivendo codice HTML e CSS caricato in electron (<http://electron.atom.io/>).

I test di usabilità si suddivideranno in 3 fasi:

1. **Preparazione:** definizione degli obiettivi del test, dei profili degli utenti e degli scenari d'uso

¹ Virzi, Robert A. "Refining the test phase of usability evaluation: How many subjects is enough?." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 34.4 (1992): 457-468.

2. **Condizione:** inizialmente si spiegherà ai partecipanti la metodologia del test e quali siano i loro compiti; poi si lascerà che essi usino il software per alcuni minuti sotto la nostra osservazione

3. **Analisi e report:** valutazione dei feedback degli utenti, dei loro problemi e dei loro suggerimenti per migliorare l'interfaccia. Nel caso in cui gli utenti non siano riusciti a completare qualche task cui sono stati sottoposti, si mostrerà loro la soluzione e si chiederà che cosa abbiano trovato poco intuitivo e quali siano stati gli ostacoli più grandi.

Task e scenari

Gli scenari sono delle descrizioni di attività tipiche che gli utenti si troveranno a svolgere durante l'uso della nostra applicazione. In ogni scenario si definirà un obiettivo che l'utente dovrà raggiungere attraverso diversi task: lo scenario si riterrà concluso solo una volta che tutti i task saranno stati completati.

È possibile che un utente dichiari di aver completato uno scenario nonostante non abbia effettivamente completato ogni task correttamente: in questi casi si mostrerà all'utente la soluzione corretta e gli si faranno domande specifiche circa la ragione per cui non ha completato lo scenario correttamente (l'interfaccia potrebbe essere ambigua o non chiara, e questo potrebbe confondere l'utente).

Prima iterazione

I risultati di questo primo scenario saranno parziali. Per massimizzare l'utilità del feedback ricevuto in questa fase, abbiamo scelto di testare l'applicazione con un utente del terzo gruppo, cioè un utente inesperto non familiare con il mondo del giornalismo. Grazie alle difficoltà incontrate dall'utente durante l'esecuzione dei task di questo scenario, miglioreremo l'interfaccia e verificheremo i miglioramenti nella seconda iterazione.

Scenario 1

La protagonista del primo scenario è Ambra, una ragazza di 17 anni che legge saltuariamente news online, appartenente quindi alla terza tipologia di utente.

Le chiediamo di ricavare, partendo dall'articolo di Hillary Clinton, alcune informazioni:

1. Dove è nata Hillary Clinton?
2. Di chi è stato predecessore Bernie Sanders?
3. Quanti collegamenti ha Donald Trump nel grafo?

Queste tre domande rispecchiano i task da effettuare (più un task iniziale per trovare l'articolo).


Esecuzione ottimale

L'utente sceglie l'articolo di Hillary sulla destra della pagina iniziale contenente gli articoli

☰


DaCENA

Articles




Bernie Sanders Endorsment

Sen. Bernie Sanders delivers his final endorsement of Hillary Clinton at the 2016 Democratic Convention. I understand that many people here in this convention . . .



Pushes for a Place at Europe's Power Table

Italy has given Europe many things, but rarely leadership. / But with the European Union fragmenting politically



I will work with Hillary to stop Trump

Bernie Sanders has urged his supporters to look beyond the Democratic presidential nomination in a speech that stopped short of fully endorsing Hillary Clinton . . .

L'utente guarda la lista delle associazioni rilevate dal software, e trova la risposta alle prime due domande: 1. New York (prima riga), 2. Peter Welch (riga 6).



User 12 ▾

User 12

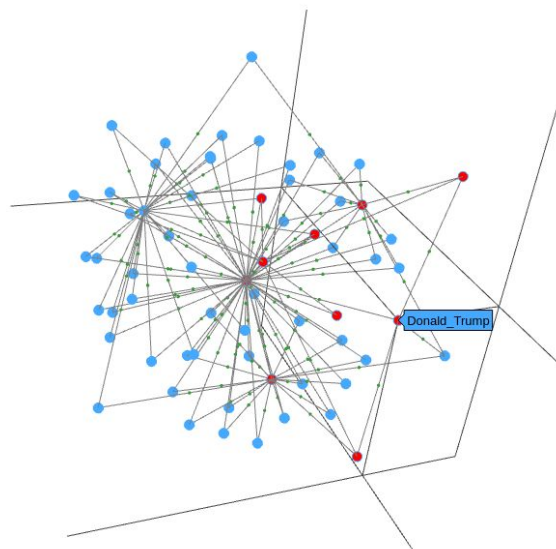
User 8

I will work with Hillary to stop Trump

Bernie Sanders has urged his supporters to look beyond the Democratic presidential nomination in a speech that stopped short of fully endorsing Hillary Clinton but made clear he was no longer actively challenging her candidacy. In an anticlimactic speech that signalled the effective end of a 14-month campaign odyssey, the Vermont senator insisted his "political revolution continues" despite Clinton's effective victory in the delegate race. 'History made': Hillary Clinton claims Democratic nomination [Read more](#) But crucially, he implied he would soon be working with her campaign to help defeat Donald Trump. "The major political task that we face in the next five months is to make certain that Donald Trump is defeated and defeated badly," Sanders told supporters in a live-stream video. "And I personally intend to begin my role in that process in a very short period of time."

N	Source	First Property	Middle one	Second Property	Destination	
1	Hillary_Clinton	R:region	New_York	L:birthPlace	Donald_Trump	5
2	Bernie_Sanders	R:otherParty	Democratic_Party_(United_States)	L:party	Hillary_Clinton	6
3	Bernie_Sanders	R:otherParty	Democratic_Party_(United_States)	L:party	Donald_Trump	6
4	Bernie_Sanders	R:predecessor	Patty_Murray	R:party	Democratic_Party_(United_States)	3
5	Bernie_Sanders	R:associate	Patrick_Leahy	R:party	Democratic_Party_(United_States)	4
6	Bernie_Sanders	L:predecessor	Peter_Welch	R:party	Democratic_Party_(United_States)	5
7	Bernie_Sanders	L:associate	Patrick_Leahy	R:party	Democratic_Party_(United_States)	5
8	Bernie_Sanders	R:party	Independent_politician	L:politicalPartyInLegislature	United_States_Senate	4
9	Bernie_Sanders	R:otherParty	Democratic_Party_(United_States)	L:politicalPartyOfLeader	United_States_Senate	5

Infine, l'utente guarda il grafo, e cerca tra i nodi rossi (quelli più importanti) il nodo di Donald Trump.



Tempo impiegato: 1 minuto circa.

Da notare è che l'ultimo task da solo (cioè quello di contare i collegamenti di Donald Trump) è stato quello che ha richiesto più tempo, circa 40 secondi. Infatti non è chiaro che i nodi rossi siano i più importanti, e inoltre la mancanza di etichette adeguate sui nodi riduce la facilità di navigazione del grafo.

Questo importante feedback sarà considerato per la seconda consegna, in cui uno degli obiettivi sarà migliorare il grafo.

Seconda iterazione

Ipotesi di design

Il design è stato progettato tenendo a mente due obiettivi: **semplicità d'uso** e **immediatezza dell'informazione**. La navigazione tra una pagina e l'altra deve essere semplice, i passaggi devono essere intuitivi e l'utente non deve mai trovarsi nella situazione di voler fare qualcosa senza capire come farla. Per questo si è cercato di mantenere l'interfaccia minimale, senza sacrificare però le funzionalità.

Allo stesso tempo vogliamo che gli utenti utilizzino l'applicazione in modo naturale, e che questa sia solo un mezzo che essi usano per accedere alle informazioni degli articoli.

Vogliamo che l'interfaccia non sia mai un ostacolo per l'utente, come purtroppo invece spesso accade, ma che sia un supporto al raggiungimento dei suoi obiettivi; l'utente deve essere messo in condizione di agire con essa in maniera spontanea. Inoltre, l'interfaccia deve essere in grado di comunicare in maniera immediata le informazioni che l'utente sta cercando: questo è un punto fondamentale nella visualizzazione del grafo, che però non siamo riusciti a raggiungere durante la prima fase. In questa seconda iterazione abbiamo quindi insistito su questo aspetto, portando a un sostanziale miglioramento.

Come abbiamo concretizzato questi due principi nella nostra interfaccia? Di seguito alcuni punti fondamentali che abbiamo tenuto a mente durante la sua implementazione:

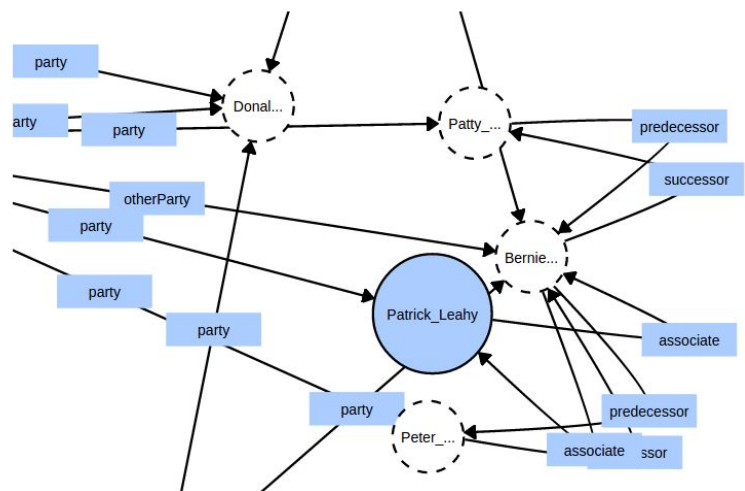
1. La pagina principale mette in mostra le foto degli articoli, i titoli e gli incipit. Le foto sono grandi in modo da essere immediatamente riconoscibili. Viene mostrato l'incipit di ogni articolo in modo da aiutare l'utente a scegliere l'articolo giusto nei casi in cui le figure non siano sufficienti (talvolta gli articoli coinvolgono personaggi simili);

2. C'è un menù sulla sinistra che l'utente può aprire in qualsiasi momento, per aiutarsi nella navigazione. È contraddistinto da un'icona standard per i menù (le tre barrette orizzontali);
3. La pagina che visualizza gli articoli è divisa in tre parti: l'articolo, il grafo, e l'elenco di entità da valutare. L'articolo e il grafo sono uno di fianco all'altro, in modo che si possano confrontare velocemente, mentre la tabella per le valutazioni è più in basso, utile solo nel caso in cui il grafo non rappresenti le entità in modo soddisfacente;

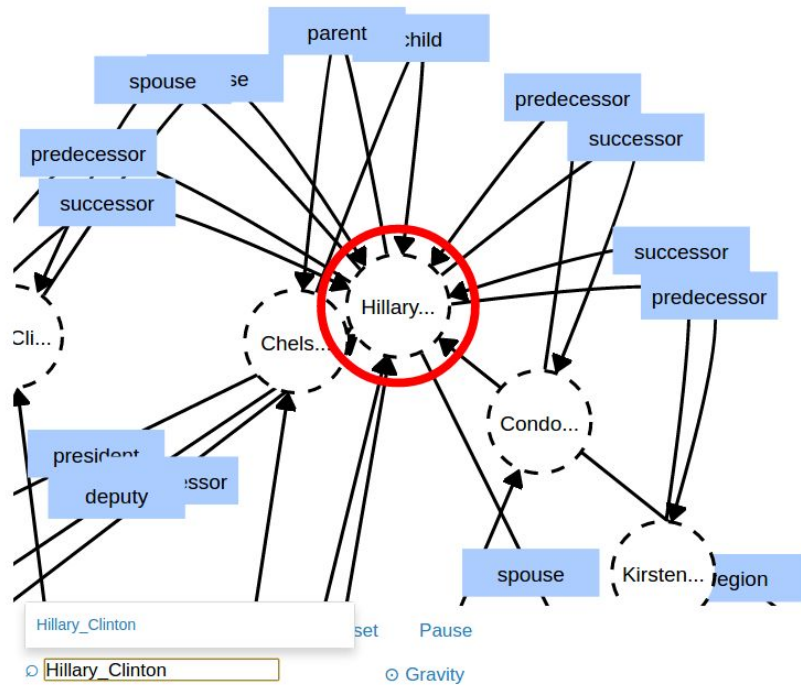
Bernie Sanders Endorsment

Sen. Bernie Sanders delivers his final endorsement of Hillary Clinton at the 2016 Democratic Convention. "I understand that many people here in this convention hall and around the country are disappointed about the final results of the nominating process," he said. "I think it's fair to say that no one is more disappointed than I am." "This election is not about, and has never been about, Hillary Clinton, or Donald Trump, or Bernie Sanders or any of the other candidates who sought the presidency. This election is not about political gossip. It's not about polls. It's not about campaign strategy. It's not about fundraising. It's not about all the things the media spends so much time discussing."

Knowledge Graph



4. Il grafo incorpora delle funzioni di ricerca (con suggerimenti durante la digitazione), e tutti i suoi nodi e archi sono etichettati. Questo è un requisito che è emerso dalla prima iterazione: inizialmente infatti non avevamo pensato di aggiungere una funzione che cercasse le entità nel grafo, e questo ha rallentato notevolmente gli utenti nei loro task. Inoltre abbiamo etichettato tutti i nodi e gli archi, mancanza di cui eravamo consapevoli anche prima di valutare i risultati dei task.



Siamo convinti che questi semplici accorgimenti saranno sufficienti per fornire agli utenti un'esperienza intuitiva.

Protocollo di validazione

Descriviamo di seguito l'impostazione del metodo di validazione che abbiamo seguito.

Task da eseguire

Task 1

Per il primo task chiederemo all'utente di rieseguire il task che avevamo proposto nella prima iterazione. In questo modo potremo fare un confronto e verificare se i risultati siano migliorati. Ricordiamo che il task consisteva nel ricavare, a partire dall'articolo di Hillary Clinton, le seguenti informazioni:

1. Dove è nata Hillary Clinton?
2. Di chi è stato predecessore Bernie Sanders?
3. Quanti collegamenti ha Donald Trump nel grafo?

Task 2

In questo task vogliamo verificare che la procedura di valutazione delle associazioni sia chiara. L'utente verrà messo a conoscenza (in maniera breve e semplice) del funzionamento dell'algoritmo di online learning che abbiamo incorporato nell'applicazione: in altre parole, si dirà all'utente che attribuendo a delle associazioni un punteggio alto, queste avranno maggiore peso nel grafo, che verrà quindi ridisegnato in modo da mostrare più in grande le entità coinvolte nelle associazioni.

Chiediamo che l'utente scelga un articolo a piacere, e:

1. Scelga un'entità che per lui è importante, ma che nel grafo non è segnata come una delle entità principali;
2. Valuti le associazioni in modo da dare punteggi alti a quelle che coinvolgono l'entità scelta, finché questa non diventa una delle entità più grandi del grafo;

Task 3

Con questo task vogliamo valutare la semplicità di navigazione della nostra applicazione.

Per farlo, si prenderanno degli utenti principalmente dal terzo gruppo (i meno esperti) e:

1. Al loro arrivo, troveranno l'applicazione già aperta sulla pagina di un articolo;
2. Sarà loro chiesto di navigare a un altro articolo. Per farlo, sarà necessario
 - a. Aprire il menu laterale
 - b. Tornare alla pagina iniziale
 - c. Selezionare l'articolo da noi indicato

Questionari

A differenza dei task, in cui gli utenti vengono misurati da noi, producendo quindi dati "passivamente", i questionari sono un modo che gli utenti hanno per lasciare un feedback più "attivo" ed esplicito. Il questionario comprende poche affermazioni che gli utenti votano utilizzando una scala Likert, esprimendo un giudizio che varia tra "assolutamente vero" a "assolutamente falso".

Il questionario si compone dei seguenti punti:

1. Non hai avuto difficoltà nel navigare tra le varie pagine dell'applicazione.
2. È stato semplice trovare le entità all'interno del grafo.
3. Non hai avuto problemi nel valutare le associazioni.
4. L'interfaccia grafica è di tuo gradimento.

5. Trovare gli articoli che stai cercando è un'operazione quasi immediata.

Ognuna di queste affermazioni va valutata con una delle seguenti risposte:

“assolutamente vero”, “più vero che falso”, “neutrale”, “più falso che vero”, “assolutamente falso”.

Metriche di usabilità

Per misurare quanto sia usabile l'interfaccia, si sono usate tre metriche:

1. **Tasso di completamento:** percentuale di compiti completati con successo, cioè senza commettere errori critici. Se un utente chiede aiuto nell'esecuzione di un task, questo dovrà essere considerato un errore critico;
2. **Tasso di completamento senza errori:** percentuale di partecipanti che portano a termine il task senza errori (sia critici che non);
3. **Tempo:** misureremo il tempo necessario agli utenti per completare i task, dal momento in cui gli viene detto di cominciare fino al momento in cui dichiarano di aver finito.

Criticità

Errori non critici

Nella categoria degli errori non critici rientrano tutti quegli errori che non compromettono la corretta esecuzione del task. Un errore critico ostacola l'utente nel completamento *ottimale* del task, per esempio facendogli perdere tempo, o confondendolo. Per esempio, l'utente può perdere tempo nel cercare una funzione, può selezionare quella sbagliata, o può navigare andando su pagine che a lui non servono. Queste cose ostacolano l'utente nell'esecuzione “perfetta” del task, ma non ne intaccano il risultato.

Errori critici

Gli errori critici sono definiti come un mancato raggiungimento degli obiettivi del task da parte dell'utente. Se un utente non riesce a completare in maniera autonoma il task, e chiede quindi aiuto, si dovrà considerare l'accaduto come un errore critico. Gli errori critici sono ovviamente più importanti di quelli non critici, in quanto si vuole mettere l'utente nella condizione di svolgere ogni attività senza bisogno di supporto da parte nostra.

Somministrazione e raccolta dei dati

Per la somministrazione dei test procederemo in maniera simile a quanto fatto nella prima iterazione. In una fase iniziale spiegheremo agli utenti come funziona la procedura, spiegheremo il tipo di applicazione che andranno ad usare, e faremo loro capire che dovranno cercare di eseguire i task in maniera autonoma.

Misureremo poi i tempi di completamento per ogni utente, fatta eccezione per coloro che si troveranno assegnati il task 2: il tempo di completamento di questo task infatti dipende non solo dalla velocità dell'utente nel muoversi all'interno dell'interfaccia, ma anche da altri fattori come l'efficienza dell'algoritmo di online learning che viene sfruttato nell'aggiornamento del grafo, che dipende dall'entità scelta dall'utente, e che influenza le possibili associazioni visualizzate in tabella. In altre parole, ci sono troppe variabili che influenzano le tempistiche di questo task, e calcolarne i tempi di esecuzione avrebbe poco senso.

Cercheremo inoltre di avere un gruppo eterogeneo di utenti. Per i primi due task sceglieremo utenti da tutti i gruppi definiti precedentemente (esperti, medi, inesperti), mentre per il terzo task (quello in cui valutiamo la facilità di navigazione con un'operazione veramente molto semplice) sceglieremo solo utenti inesperti. Sono loro infatti che hanno una probabilità più alta di far emergere grossi problemi di usabilità.

Infine sottoporremo i questionari agli utenti, dando loro un minuto per rispondere.

Analisi dei dati

I risultati che abbiamo ottenuto somministrando i nostri test sono soddisfacenti, in quanto le nostre ipotesi di design sono state confermate (complice il fatto che avevamo già ricevuto feedback nella prima iterazione).

Il tempo medio per svolgere il primo task si è notevolmente ridotto, passando da 1 minuto a 40 secondi circa: la funzione di ricerca ha fatto risparmiare 20 secondi di tempo agli utenti.

Anche il terzo task è andato bene, confermando l'intuitività dell'interfaccia: il tempo di completamento medio è circa 8 secondi. Un utente ci ha anche segnalato un interessante suggerimento per migliorare la navigabilità: aggiungere un pulsante "indietro" con relativa freccia per tornare alla pagina principale contenente gli articoli.

L'unico task con cui abbiamo riscontrato qualche problema è il secondo. Qui infatti gli utenti dovevano continuare a valutare le associazioni indefinitamente: questo task è risultato leggermente difficile per gli utenti inesperti, che nonostante siano riusciti a portarlo a termine hanno impiegato più tempo degli altri. La loro "difficoltà" stava nel fatto che non capivano esattamente cosa stesse succedendo nel grafo, cioè come questo cambiasse con le loro

valutazioni. Se ci fosse un'altra iterazione, potremmo lavorare su questo punto e aiutare l'utente a riconoscere le entità nel grafo valutazione dopo valutazione, magari utilizzando dei colori o visualizzando il punteggio di importanza all'interno di ogni nodo: così gli utenti dovrebbero essere più consapevoli degli effetti che i loro voti hanno sul grafo.