

Rapporto conclusivo sui risultati della procedura di valutazione

Survey

Nel presente documento si riassumono i contenuti del deliverable "**D5.5_1 Rapporto conclusivo sui risultati della procedura di valutazione**" relativo all'attività *SP 5.5 Sperimentazione e Valutazione con utenti reali*, realizzata nell'ambito del quinto Obiettivo Realizzativo (OR 5) "Sperimentazione e valutazione dell'impatto della ricerca".

L'attività SP 5.5 prevede la sperimentazione dei dimostratori con utenti reali e la successiva valutazione dei risultati. Tale attività rappresenta la sintesi di quattro fasi temporalmente distinte ed operativamente diverse:

1. Identificazione dei criteri / metodologia di valutazione dei risultati attesi
2. Sperimentazione con utenti reali
3. Raccolta dei risultati
4. Elaborazione dei risultati

1. Identificazione dei criteri / metodologia di valutazione dei risultati attesi

In questa fase è stata proposta una metodologia ed una metrica che è possibile applicare per la valutazione dei prototipi realizzati in SAPI intesi come applicazioni web che debbono essere fruite da utenti disabili ed in particolare da un'utenza ipovedente.

Per gli scopi di ricerca del progetto SAPI, nel definire la metodologia ed i criteri per misurare l'accessibilità/usabilità dei prototipi, sono state utilizzate le linee guida dettate dal Decreto Ministeriale 8 luglio 2005 "Requisiti tecnici e i diversi livelli per l'accessibilità agli strumenti informatici" in cui è presente una procedura per la "Verifica tecnica e requisiti tecnici di accessibilità delle applicazioni basate su tecnologie internet".

La procedura di verifica generale è stata specializzata in base alle seguenti considerazioni:

- 1) Nella procedura i profili che nel decreto sono indicati come *esperto di interazione con persone disabili* ed *esperto di fattori umani* sono stati sostituiti con un'unica figura indicata come **esperto SAPI**;
- 2) In considerazione del fatto che i prototipi SAPI sono realizzati in tecnologia Silverlight, tutti i riferimenti ai requisiti tecnici che dettano l'uso di tag HTML o fanno riferimento a tecnologie che non vengono utilizzate dai prototipi non vengono presi in considerazione.
- 3) Nella fase di verifica soggettiva, *Fase b: Esecuzione dei task da parte del gruppo di valutazione*, tra le informazioni che l'**esperto SAPI** deve raccogliere durante l'uso del prototipo da parte del gruppo di valutazione, ci sono una serie di parametri che misurano l'usabilità dell'applicazione. La procedura originale prevede l'esecuzione dei task da casa e da un laboratorio attrezzato. Nel caso di SAPI la fruizione è avvenuta solo all'interno del laboratorio attrezzato.

- 4) Non vengono presi in considerazione i requisiti descritti negli allegati "Requisiti tecnici di accessibilità per i personal computer di tipo desktop e portatili" e "Requisiti tecnici di accessibilità per l'ambiente operativo, le applicazioni e i prodotti a scaffale".

Come previsto dalla metodologia di verifica soggettiva delle applicazioni basate su tecnologie internet, sono state tenute in conto le seguenti fasi per la valutazione del framework prototipale SAPI:

Fase a: Costituzione del gruppo di valutazione

La valutazione prevede la costituzione del gruppo di valutazione i cui componenti disabili utilizzano le tecnologie assistite fornite di prototipi SAPI.

Fase b: Esecuzione dei task da parte del gruppo di valutazione

L'esecuzione dei task da parte dei componenti del gruppo di valutazione avviene in contesti appositamente costituiti (ambiente di laboratorio).

Il gruppo di valutazione esegue una serie di prove basate sulla interazione con l'ambiente. Le prove vengono svolte sia in forma libera, cioè senza compiti specifici, sia per obiettivi, se eseguite secondo compiti specifici.

Nell'esecuzione delle prove, il gruppo di valutazione è guidato dall'esperto SAPI.

Nel corso della navigazione libera, l'esperto SAPI raccoglie i commenti dell'utente, anche verbali, e le osservazioni sul suo comportamento.

Nella prova su compiti specifici, l'esperto registra il tipo di compito e le seguenti informazioni:

- tempo impiegato per eseguire i compiti;
- numero di compiti completati in modo corretto e completo entro un tempo prestabilito;
- numero di errori commessi e loro gravità
- ricorrenza degli errori e loro tipologia;
- tempo speso per recuperare gli errori commessi;
- tempo speso in esplorazione del prodotto;
- tempo speso in inattività;
- numeri di accessi all'help o di ricorso all'assistenza;
- numero e tipo di comandi usati dall'utente per l'esecuzione del compito;
- numero e tipo di comandi mai usati dall'utente durante l'esecuzione del compito;
- numero di volte che l'utente esprime chiaramente frustrazione o soddisfazione;

Fase c: Valutazione dei risultati ed elaborazione del rapporto conclusivo

La verifica soggettiva si conclude con la predisposizione di un rapporto nel quale l'esperto SAPI indica la valutazione su scale soggettive ricavata dalla conoscenza del prototipo, le proprie considerazioni sulle caratteristiche qualitative del sito, i dati relativi alle prestazioni degli utenti in relazione ai compiti affidati: performance, commenti, osservazioni comportamentali, le risposte a questionari di valutazione compilati dagli utenti e la valutazione complessiva del livello di qualità raggiunto che, per il prototipo SAPI, è attribuita secondo il seguente schema:

1. Assolutamente d'accordo (voto massimo)
2. D'accordo (voto intermedio positivo)
3. In parte in disaccordo (voto intermedio negativo)
4. Completamente in disaccordo (voto minimo)

Per quel che riguarda i *risultati attesi* per ognuno dei dimostratori implementati in SAPI, dal punto di vista del sistema nel suo complesso, ci si aspetta, come minimo, che il suo impiego debba rispettare i requisiti tecnici e funzionali identificati nelle attività di ricerca industriale.

2. Sperimentazione con utenti reali

Nell'ambito delle attività del progetto SAPI è stata attivata una collaborazione con l'Istituto Professionale di Stato per l'Industria e Artigianato per Ciechi "Paolo Colosimo" di Napoli.

In tale ambito, è stato predisposto un questionario da sottoporre a soggetti ipovedenti studenti, operatori e insegnanti dell'istituto con lo scopo di poter desumere elementi di validazione degli orientamenti inizialmente previsti nel progetto.

L'analisi dei questionari compilati, tuttavia, non ha realizzato una base dati significativa, pertanto, per la costituzione di un adeguato panel di utenti da coinvolgere in fase di sperimentazione del prototipo si è scelto di contattare direttamente degli esperti di accessibilità e usabilità delle risorse web al fine di:

- validare l'attività di sviluppo effettuata nei due dimostratori SAPI;
- realizzare, grazie al contributo degli esperti, delle modifiche nel framework per garantirne un maggior livello di accessibilità/usabilità;

La fase di sperimentazione con utenti reali, pertanto, è stata condotta con la collaborazione dei membri dell'OSI (Osservatorio Siti Internet) dell'UIC (Unione Italiana Ciechi), che hanno messo a disposizione la propria esperienza e competenza in materia di accessibilità.

Gli utenti finali hanno permesso di tradurre efficacemente le esigenze di una larga fetta di utenza cui il framework intende rivolgersi potendo rilevare i problemi di accessibilità della piattaforma SAPI nei confronti di persone con particolari disabilità. La loro esperienza, tuttavia, si è rilevata proficua nell'evidenziare i bisogni e le problematiche di utilizzo del framework nei confronti di un'utenza ben più vasta, a prescindere dalla disabilità visiva.

3. Raccolta dei risultati

Nei giorni 18 e 19 Dicembre 2008 i membri dell'OSI hanno eseguito dei test presso il laboratorio di Napoli su una prima versione del prototipo SAPI, utilizzando i nuovi servizi "intelligenti" sia su un chiosco costruito appositamente per la sperimentazione, che su un PDA.

Le prove sono state svolte sia in forma libera, cioè senza compiti specifici, sia per obiettivi, ovvero secondo compiti specifici.

Nella esecuzione delle prove, il gruppo di valutazione è stato guidato dall'esperto SAPI.

Nel corso della navigazione libera, l'esperto SAPI ha raccolto i commenti dell'utente, anche verbali, e le osservazioni sul suo comportamento.

Nella prova su compiti specifici, l'esperto SAPI ha registrato, per ogni task, il tipo di compito e le seguenti informazioni:

- tempo impiegato per eseguire il task
- numero di compiti completati in modo corretto
- numero di errori commessi e loro gravità
- ricorrenza degli errori e loro tipologia
- numero di accessi all'help o di ricorso all'assistenza
- numero di volte che l'utente esprime chiaramente frustrazione o soddisfazione

Tali informazioni sono state organizzate secondo la Tabella 1.

Tipo di sperimentazione					
Utente:					
Sperimentazione N°:					
	Task 1	Task 2	Task 3	Task 4	Task 5
Tempo impiegato per eseguire il task					
N° di compiti completati in modo corretto					
N° di errori commessi e loro gravità					
Ricorrenza degli errori e loro tipologia					
N° di accessi all'help o di ricorso all'assistenza					
N° di volte che l'utente esprime chiaramente frustrazione o soddisfazione					

Tabella 1: Raccolta dati sperimentali

Nel corso delle prove gli utenti hanno testato le funzionalità (alcune delle quali risultavano in fase di sviluppo) dei due dimostratori implementati nel framework SAPI.

a) Test su chiosco

Trattandosi di persone esperte di tecnologie informatiche, hanno potuto familiarizzare in tempi brevi con i servizi.

Dopo una prima fase di esplorazione libera delle funzionalità del framework, è stato loro richiesto di eseguire sul chiosco multimediale i seguenti task:

1. compilazione di un bollettino bianco;
2. richiesta prestito Bancoposta;
3. compilazione di un bollettino premarcato;

Nella Figura 1 sono stati riportati i grafici dei tempi registrati rispettivamente per la tipologia di utente non vedente ed ipovedente per ciascuno dei tre task elencati ognuno dei quali ripetuto per 5 volte.

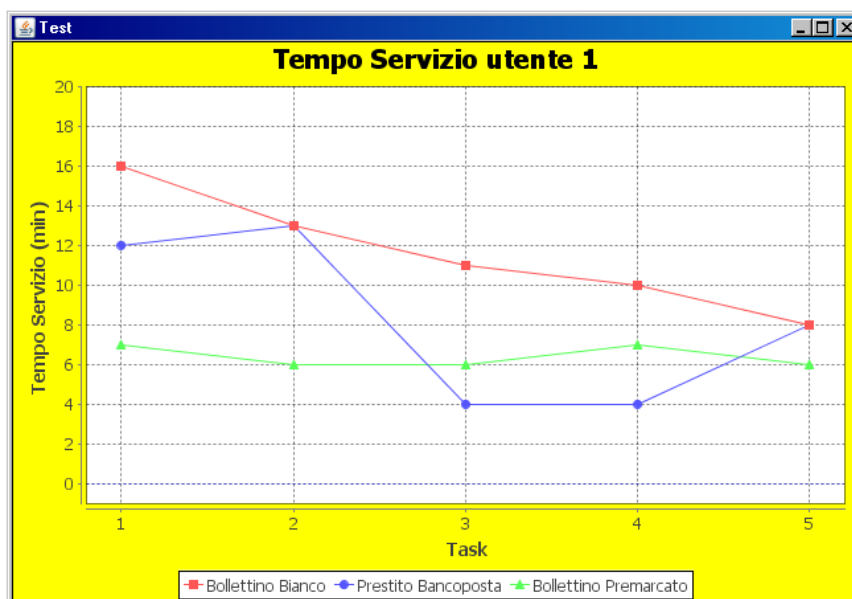


Figura 1: Grafico dei tempi di servizio registrati per l'utente non vedente

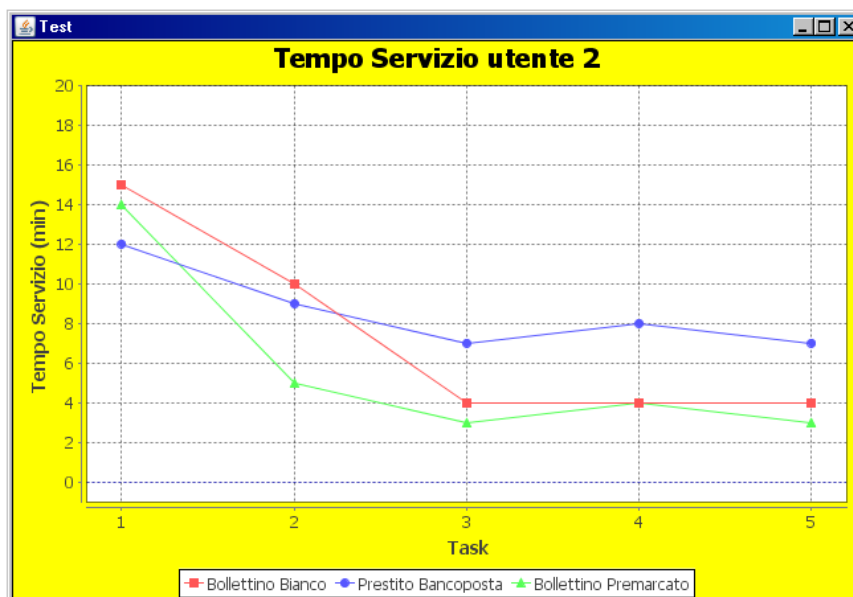


Figura 2: Grafico dei tempi di servizio registrati per l'utente ipovedente

Oltre ai tempi richiesti per l'esecuzione dei servizi sono stati registrati il numero, la tipologia e la ricorrenza degli errori che ciascun utente ha commesso nel corso delle sessioni. Le problematiche di utilizzo del chiosco e del framework, come dimostrano i grafici, hanno influenzato i tempi di esecuzione dei task, particolarmente nelle prime sessioni. Dopo un certo numero di volte che è stato eseguito lo stesso task, il tempo speso per completarlo è andato progressivamente a ridursi (l'utente cioè, nel corso delle prove, è diventato più esperto).

Nel caso di utente non vedente, gli errori e le difficoltà incontrati sono scaturiti principalmente dalla mancata conoscenza nel funzionamento del sistema di TTS (Text To Speech)/ASR (Automatic Speech Recognition) e dalla tastiera non proprio uguale a quella tradizionalmente utilizzata in commercio. Per quanto riguarda il primo aspetto, in particolare, diverse volte l'utente ha anticipato il canale Loquendo nel passaggio dalla modalità TTS ad ASR: tale operazione, infatti, richiede di norma alcuni secondi e il soggetto non vedente, dal canto suo, non avendo riferimenti temporali, si è trovato più volte a dover ripetere lo stesso comando prima di proseguire nelle operazioni. Nel servizio di pagamento bollettino premarcato, l'utente non vedente, avendo ormai compreso la logica della sintesi e del riconoscimento vocale non ha commesso alcun errore legato all'utilizzo della tecnologie TTS/ASR implementate nel dimostratore: i tempi di completamento del task sono pertanto risultati minori che negli altri due casi.

In particolare la tipologia di utente non vedente:

- ha espresso soddisfazione per il sistema di guida implementato in SAPI attraverso l'utilizzo di tecnologie TTS/ASR.
- ha espresso, inoltre, pieno consenso per il meccanismo di stampa differenziato per stereotipi che, nel caso specifico del soggetto non vedente, ha previsto la stampa in braille della ricevuta del bollettino oppure del report relativo alla richiesta prestito.

Per gli utenti ipovedenti, nelle prime due o tre interazioni, i tempi di esecuzione del task sono risultati decisamente maggiori rispetto alle successive interazioni. Ciò si è verificato a causa

di un'iniziale fase di esplorazione più dettagliata di ogni pagina del servizio in cui l'utente ipovedente ha dovuto leggere accuratamente tutte le istruzioni visualizzate. Nel corso del servizio di pagamento bollettino premarcato l'utente ipovedente ha compiuto alcuni errori: nella digitazione di comandi da tastiera, nella procedura di acquisizione del bollettino tramite scanner e soprattutto, nei diversi test, ha mostrato difficoltà nell'effettuare la procedura di pagamento tramite POS.

L'analisi dei tempi, per il bollettino bianco, ha mostrato una progressiva riduzione della durata in conseguenza di un rapido ed efficace apprendimento dell'utente per il servizio richiesto. Tale risultato in parte è da attribuirsi all'affinità con il servizio di compilazione bollettino premarcato rispetto al quale, il bollettino bianco prevede in più la compilazione "a mano" di alcuni campi.

Nel servizio di richiesta prestito l'eccessivo tempo impiegato per il primo test è da attribuirsi in parte ad un task più complesso in cui è stato chiesto di inoltrare la richiesta prestito. Ciò prevede la compilazione di diversi campi da digitare da parte dell'utente.

Infine per quanto riguarda il ricorso al sistema di help, non è stato testato dall'utente non vedente poiché non ancora vocalizzato mentre l'utente ipovedente, sebbene potesse farlo, non ha mai manifestato la necessità di richiederlo. I servizi, nel loro complesso, sono risultati fluidi, chiari e di facile utilizzo in ogni loro punto.

In particolare l'utente ipovedente ha espresso pieno consenso per la funzionalità di predizione implementata che permette di anticipare e quindi agevolare l'utente nella compilazione dei campi richiesti del bollettino bianco.

Gli utenti hanno tuttavia espresso chiaramente dissenso per la procedura di pagamento effettuata tramite POS, che è stata ritenuta non accessibile. Gli utenti, in particolare, hanno sottolineato nei confronti della gestione del POS delle mancanze dal punto vista hardware e software. A tal riguardo, la natura inaccessibile del dispositivo di pagamento presente sul chiosco era nota prima dei test in quanto erano già state individuate delle soluzioni hardware più accessibili.

b) Test su palmare

Questo test è stato eseguito esclusivamente dagli utenti ipovedenti utilizzando un palmare con sistema operativo Windows Mobile 5.0.

Condizione indispensabile per utilizzare i servizi di SAPI su PDA è l'aver installato ed attivo l'ambiente runtime per midlet Java; dopodiché l'utente si collega con i server di Poste Italiane per effettuare le operazioni come descritte per il terminale:

- Pagamento bollettino prestampato.
- Pagamento bollettino bianco.
- Richiesta prestito.

Il test è stato effettuato con la funzione di pagamento bollettino.

Purtroppo, per l'ipovisione sono emerse le limitazioni intrinseche nell'uso dei palmari, in quanto, un ipovedente, a causa dell'esiguo spazio a disposizione per la digitazione e compilazione dei campi, ha enormi difficoltà.

Però, nel caso specifico, nella sessione di test si è aggiunto un altro problema: il software, essendo provvisto di una personalizzazione dei caratteri che si sommava alla personalizzazione dei caratteri impostata dall'utente, a livello di sistema operativo, ha reso il testo pressoché illeggibile.

E' ovvio che, durante il test, è stato consigliato di annullare questa opzione perché, chi usa questi strumenti, tende sempre a personalizzare in maniera nativa il sistema operativo usato o utilizza un software ingrandente.

Il test è stato portato a compimento con notevoli difficoltà, ma con i suggerimenti dati, si confida in un forte miglioramento delle funzioni implementate.

4. Elaborazione dei risultati

Gli errori e le problematiche riscontrati dagli esperti nel corso dei test, hanno evidenziato l'esigenza di effettuare alcune modifiche hardware sul totem e software per migliorare il livello di accessibilità del framework SAPI. Tali accorgimenti, trattandosi di una prima versione del prototipo, si sono rivelati ampiamente pronosticabili nonché efficaci.

I punti discussi e i suggerimenti forniti, infatti, prescindono dalla specifica condizione di disabili visivi cui gli utenti fanno riferimento: la loro esperienza e competenza, infatti, è stata utile per rappresentare un panel di utenti ben più ampio, significativo ed eterogeneo. Dalla loro collaborazione sono emerse una serie di considerazioni che sintetizzano le necessità dell'utenza, intesa nella sua accezione più completa, cui rientrano persone con caratteristiche differenti in termini di età, disabilità, interessi, competenze, professione, background culturale, livello di alfabetizzazione informatica, etc.

Per tale motivo il contributo degli esperti è risultato estremamente efficace e ha consentito di testare la validità del prototipo simulando il funzionamento dei due dimostratori in diversi scenari e per differenti casi d'uso.

Rapporto conclusivo

Al termine delle due giornate di test è stato sottoposto agli utenti un questionario tramite cui i membri dell'OSI hanno fornito una valutazione qualitativa del prototipo.

All'interno del questionario si è fatto riferimento alla scala di valutazione strutturata su quattro livelli:

1. Assolutamente d'accordo (voto massimo)
2. D'accordo (voto intermedio positivo)
3. In parte in disaccordo (voto intermedio negativo)
4. Completamente in disaccordo (voto minimo)

Sulla base delle risposte al questionario, il livello di qualità del prototipo è risultato, allo stato dei test, pari al valore 2 (voto intermedio positivo).

Gli utenti, tuttavia, hanno lasciato intravedere chiaramente che, l'aggiunta delle modifiche suggerite, possa assicurare al chiosco multimediale e al framework un aspetto di assoluta accessibilità e di qui il conseguimento di una votazione massima in termini di qualità del prototipo.

Al di là delle risposte fornite, gli esperti hanno espresso pieno consenso e soddisfazione per il livello di innovazione e di usabilità del prototipo: l'eliminazione di tutti i punti di inaccessibilità

rilevati durante i test, renderà il chiosco multimediale in particolare ed il framework in generale un ottimo strumento per l'erogazione di servizi accessibili ed usabili.

Un'ulteriore metrica per la valutazione positiva del prototipo, infine, è stata fornita dall'aderenza del framework ai requisiti funzionali individuati nella fase di ricerca.

In conclusione, tutte le modifiche a livello software richieste dagli esperti ed annottate in fase di test, sono state discusse e quindi implementate dal gruppo di lavoro, compatibilmente alle tecnologie utilizzate per lo sviluppo dei due dimostratori.

In tal senso, i feedback forniti dai membri dell'OSI hanno consentito di perfezionare la qualità del framework evidenziando delle linee guida da tenere in conto per la progettazione di un'applicazione assolutamente accessibile.

L'intento è di seguire tali linee guida nella fase successiva di industrializzazione del prototipo durante la quale è previsto un nuovo incontro con gli esperti.

Va da se che per quanto concerne gli accorgimenti a livello hardware, dovranno sposarsi con l'attuale configurazione dei chioschi in dotazione all'interno degli uffici di Poste Italiane.