

Analizzare i risultati del test di valutazione con il protocollo eGLU



Questo materiale didattico è stato realizzato da Formez PA nel *Progetto PerformancePA*, Ambito A Linea 1, in convenzione con il Dipartimento della Funzione Pubblica, organismo intermedio del Programma Operativo Nazionale Governance e Azioni di Sistema (PON GAS), Asse E Capacità istituzionale. Il PON GAS è cofinanziato dal Fondo Sociale Europeo ed è a titolarità del Ministero del Lavoro e delle Politiche Sociali.

L'opera è distribuita con Licenza [Creative Commons Attribuzione - Condividi allo stesso modo 4.0 Internazionale](https://creativecommons.org/licenses/by-sa/4.0/).



Autore: Maurizio Boscarol, Alessandra Cornero, Elvira Zollerano

Creatore: Formez PA

Diritti: Dipartimento della Funzione Pubblica

Data: Ottobre 2015

Analizzare i risultati del test di valutazione con il protocollo eGLU

In questa videolezione parleremo della fase di analisi dei risultati e di stesura del report di test di usabilità condotti con la metodologia eGLU 2.1.

I test di usabilità sono situazioni in cui si chiede a rappresentanti dei nostri utenti di provare a eseguire dei compiti sul nostro sito web e li si osserva senza intervenire, per rilevare eventuali difficoltà. Dopo la fase di preparazione e di esecuzione, abbiamo osservato le prestazioni dei partecipanti e vogliamo riassumere i risultati in un report utile per le modifiche.

I risultati che derivano dalle osservazioni delle prestazioni degli utenti ricadono in due categorie: i dati di prestazione e i questionari di valutazione. I dati di prestazione riguardano: da un lato il tasso di successo per ogni partecipante, e per tutti i partecipanti complessivamente; dall'altro i problemi incontrati. I dati sul tasso di successo dei task sono raccolti durante l'osservazione e vanno inseriti, dopo la fine dell'esecuzione della procedura, nell'Allegato 8, scaricabile sempre dal sito della Funzione Pubblica insieme al Protocollo eGLU.

Ogni task che l'utente completa sarà un successo, e ogni task non completato un fallimento. Per quanto possa sembrare logico, a volte non è sempre semplice capire quando considerare un task riuscito, specialmente se l'utente lo completa parzialmente o se il task non specifica, ad esempio, se un'informazione va solo trovata o anche capita. Per questo avere dei chiari criteri di successo, stabiliti durante la fase di preparazione, è così importante.

Per ogni partecipante, il tasso di successo sarà pari al numero di task riusciti rispetto a quelli tentati. Per l'intero test, si devono considerare i task riusciti da tutti i partecipanti, diviso quelli effettivamente tentati. In alcune condizioni è possibile che non tutti i partecipanti tentino di eseguire tutti i task, ad esempio perché vengono interrotti o decidono di terminare anzitempo il test, è dunque importante tenerne conto nel calcolo.

Con il modello già predisposto nell'Allegato 8 sarà più semplice inserire i dati ed effettuare i calcoli. Tale modello servirà anche a calcolare il tasso di successo complessivo del sito o del servizio web che corrisponde al numero di task riusciti rispetto ai task tentati da tutti gli utenti e a dare un dettaglio anche di quale task abbia avuto il tasso di successo più alto. Entrambi questi dati dovrebbero essere segnati nel report.

I problemi incontrati dagli utenti e osservati dal conduttore sono invece quelli nei quali incappa

il partecipante ad esempio quando sbaglia, si blocca o mostra esitazioni. Oppure quando prende una strada ma poi torna indietro, perché non è sicuro del risultato, o ancora quando invece è sicuro ma non si accorge che in realtà sta sbagliando. Non sempre è facile identificare un problema, perché talvolta gli utenti non parlano, non manifestano il proprio pensiero né esplicitano le ragioni di eventuali esitazioni. Per questo motivo è necessaria un po' di pratica da parte del conduttore.

Alla fine bisogna stilare un elenco dei problemi osservati e per farlo è utile preparare una griglia da utilizzare già nella fase di esecuzione, in cui sono elencati i problemi principali che di solito si incontrano, come:

- il partecipante si blocca;
- il partecipante dichiara di essere confuso da elementi di layout, immagini, video, eccetera;
- il partecipante dichiara di essere confuso dalla sovrabbondanza di opzioni;
- oppure il partecipante sceglie un percorso del tutto errato;
- o il partecipante non riconosce la funzione di testi, bottoni, inviti funzionali;
- il partecipante travisa il significato di testi, bottoni, inviti funzionali.

Non è utile solo tracciare i problemi riscontrati dai partecipanti, ma è possibile evidenziare eventuali apprezzamenti manifestati dai partecipanti. Anche per sottolineare agli sviluppatori aspetti da non modificare in caso di revisioni.

Alcuni esempi di apprezzamenti, che possiamo sempre sintetizzare nella specifica griglia, sono:

- il partecipante esprime di sua iniziativa apprezzamenti su un contenuto/servizio specifico;
- il partecipante esprime di sua iniziativa un apprezzamento rispetto alla ricchezza, alla completezza e all'utilità di un contenuto/servizio;
- il partecipante esprime di sua iniziativa la soddisfazione rispetto ad un task completato con successo e facilità.

Per ogni problema è utile indicare il numero di partecipanti che lo ha incontrato. In questo modo è possibile avere una stima dei problemi più frequenti. Anche se esula dallo scopo del Protocollo, può essere utile provare ad assegnare, ove possibile, un giudizio di gravità o di impatto per ciascun problema, a discrezione del conduttore e dell'eventuale assistente.

I problemi osservati andrebbero tutti affrontati e discussi dai responsabili del sito o del servizio web, che sono i principali candidati a indicare le modifiche da effettuare. Se si dovessero incontrare particolari difficoltà sia per l'interpretazione dei problemi che per l'identificazione di

soluzioni di miglioramento, è meglio avvalersi della consulenza di un esperto.

Oltre ai dati di prestazione (relativi ai tassi di successo e ai problemi) è utile raccogliere, analizzare e riportare nel report anche i dati di valutazione soggettiva dei partecipanti rispetto ai compiti e all'intero test, riferiti a ciò che hanno provato. Anche questi sono presenti negli allegati del Protocollo eGLU e scaricabili dal sito della funzione pubblica.

I dati soggettivi di intenzione d'uso (NPS) o di usabilità percepita (SUS e UMUX-LITE), espressi attraverso i questionari post-test, vanno elaborati inserendoli nei moduli per il calcolo automatico presenti nell'allegato 5a per il *Net Promoter Score* (NPS), nell'allegato 6a per il *System Usability Scale* (SUS) e nell'allegato 7a nel caso si sia usato lo *Usability Metric for User Experience* (UMUX-LITE).

Il risultato di questi questionari riportano a un punteggio che va interpretato. Per interpretare questi dati si considerano alcuni criteri riassuntivi. Il punteggio del *Net Promoter Score*, che può distribuirsi fra -100 e 100, dovrebbe essere quantomeno di segno positivo e il più possibile vicino al 100. Numeri positivi bassi (consideriamo da 0 a 30) indicano comunque risultati misti. Il *System Usability Scale* e lo *Usability Metric for User Experience* invece, utilizzano lo stesso criterio di valutazione, per cui su un punteggio che va da 0 a 100, si dovrebbe ottenere un numero maggiore almeno di 68, e idealmente più alto.

Alla fine, il report è il documento che contiene tutti i dati rilevanti durante il test e riassume indicazioni di problemi ed eventuali proposte di modifica. Per un modello di report semplice da utilizzare si può fare riferimento all'Allegato 9 del Protocollo. Anche se non è necessario usare questo modello prestabilito, l'importante è che il report contenga i seguenti dati minimi:

- Il numero dei partecipanti e dei task;
- la descrizione dei task e le pagine di completamento (o criterio di successo) del task;
- il tasso di successo del sito e del servizio;
- il tasso di successo per ciascun task e per ciascun partecipante;
- le Misure dirette dell'usabilità percepita (il SUS o l'UMUXLITE);
- le Misure di intenzione d'uso del sito o servizio web (il NPS);
- un elenco dei problemi riscontrati.

Un ulteriore livello di approfondimento del report può inoltre prevedere i suggerimenti per la risoluzione dei problemi, una valutazione della gravità dei problemi, oltre a una associazione dei problemi ai principali principi euristici violati dall'interfaccia.