

donbosco.info: un motore di ricerca salesiano

Presentiamo la nuova piattaforma donbosco.info che è un motore di ricerca salesiano pensato per rendere più agevole la consultazione dei documenti legati al carisma di Don Bosco. Nato per supportare il Bollettino Salesiano Online, supera i limiti dei tradizionali sistemi di archiviazione, spesso incapaci di intercettare tutte le occorrenze delle parole. Questa soluzione integra un hardware dedicato e un software sviluppato appositamente, offrendo anche una funzione di lettura. L'interfaccia web, volutamente semplice, permette di navigare tra migliaia di documenti in diverse lingue, con la possibilità di filtrare i risultati per cartella, titolo, autore o anno. Grazie alla scansione OCR dei documenti PDF, il sistema individua il testo anche quando non perfetto, e adotta strategie per ignorare punteggiatura e caratteri speciali. I contenuti, ricchi di materiale storico e formativo, mirano a diffondere il messaggio salesiano in modo capillare. Con l'upload libero per documenti, si incoraggia l'arricchimento continuo della piattaforma, migliorando la ricerca.

Nell'ambito dei lavori per la redazione del Bollettino Salesiano OnLine, si è resa necessaria la creazione di vari strumenti di supporto, fra cui un motore di ricerca dedicato.

Questo motore di ricerca è stato concepito tenendo conto dei limiti attualmente presenti nelle varie risorse salesiane disponibili in rete. Molti siti offrono sistemi di archiviazione con funzionalità di ricerca, ma spesso non riescono a individuare tutte le occorrenze delle parole, a causa di limiti tecnici o restrizioni introdotte per evitare il sovraccarico dei server.

Per superare tali difficoltà, anziché costruire un semplice archivio di documenti con una funzione di ricerca, abbiamo

realizzato un vero e proprio motore di ricerca, dotato anche di una funzione di lettura. Si tratta di una soluzione completa, basata su hardware dedicato e su un software sviluppato appositamente.

In fase di progettazione, abbiamo valutato due opzioni: un software da installare localmente oppure un'applicazione *server-side* accessibile via web. Poiché la missione del Bollettino Salesiano Online è diffondere il carisma salesiano al maggior numero di persone, si è deciso di optare per la soluzione web, così da consentire a chiunque di cercare e consultare documenti salesiani.

Il motore di ricerca è disponibile all'indirizzo www.donbosco.info. L'interfaccia web è volutamente essenziale e "spartana", per garantire maggiore velocità di caricamento. Nella "*home page*" sono elencati i file e le cartelle presenti, con lo scopo di agevolarne la consultazione. I documenti non sono soltanto in italiano, ma disponibili anche in altre lingue, selezionabili tramite l'apposita icona in alto a sinistra.

La maggior parte dei file caricati è in formato PDF ricavato da scansioni con OCR (riconoscimento ottico dei caratteri). Poiché l'OCR non è sempre perfetto, a volte non tutte le parole cercate vengono rilevate. Per ovviare a ciò, sono state implementate diverse strategie: ignorare la punteggiatura e i caratteri accentati o speciali, e consentire la ricerca anche in presenza di caratteri mancanti o errati. Ulteriori dettagli sono consultabili nella sezione FAQ, accessibile dal piè di pagina.

Data la presenza di migliaia di documenti, la ricerca può restituire un numero molto elevato di risultati. Per questo è possibile restringere l'ambito della ricerca per cartelle, per titolo, autore o anno: i criteri sono cumulativi e aiutano a trovare più rapidamente ciò che serve. I risultati vengono elencati in base a un punteggio di pertinenza, che attualmente

tiene conto principalmente della densità delle parole chiave all'interno del testo e della loro vicinanza.

Idealmente, sarebbe preferibile disporre dei documenti in formato vettoriale invece che scannerizzati, poiché la ricerca risulterebbe sempre accurata e i file sarebbero più leggeri, con conseguenti vantaggi in termini di velocità.

Se possedete documenti in formato vettoriale o di qualità migliore rispetto a quelli già presenti nel motore di ricerca, potete caricarli tramite il servizio di upload disponibile su www.donbosco.space. Potete anche aggiungere altri documenti non presenti nel motore di ricerca. Per ottenere le credenziali di accesso (nome utente e password), inviate una richiesta via e-mail a bsol@sdb.org.