

donbosco.info: salezjańska wyszukiwarka

Przedstawiamy nową platformę donbosco.info, która jest salezjańską wyszukiwarką stworzoną, aby ułatwić przeglądanie dokumentów związanych z charyzmatem Księdza Bosko. Powstała, aby wspierać Biuletyn Salezjański Online, przekracza ograniczenia tradycyjnych systemów archiwizacji, często niezdolnych do wychwycenia wszystkich odmian słów. To rozwiązanie umożliwia hardware i specjalnie opracowany software, oferując również funkcję czytania. Interfejs, celowo prosty, pozwala na przeglądanie tysięcy dokumentów w różnych językach, z możliwością filtrowania wyników według folderu, tytułu, autora lub roku. Dzięki skanowaniu OCR dokumentów PDF, system identyfikuje tekst nawet, gdy nie jest idealny, i stosuje strategie ignorowania interpunkcji i znaków specjalnych. Treści, bogate w materiały historyczne i formacyjne, mają na celu szerzenie przesłania salezjańskiego w sposób kompleksowy. Dzięki swobodnemu przesyłaniu dokumentów zachęca się do ciągłego wzbogacania platformy, poprawiając wyszukiwanie.

W ramach prac nad redakcją *Biuletynu Salezjańskiego Online* konieczne stało się stworzenie różnych narzędzi wsparcia, w tym specjalnej wyszukiwarki.

Ta wyszukiwarka została zaprojektowana z uwzględnieniem ograniczeń obecnych w różnych zasobach salezjańskich dostępnych w sieci. Wiele stron oferuje systemy archiwizacji z funkcjami wyszukiwania, ale często nie są one w stanie zidentyfikować wszystkich wystąpień słów z powodu ograniczeń technicznych lub ograniczeń wprowadzonych w celu uniknięcia przeciążenia serwerów.

Aby przezwyciężyć te trudności, zamiast budować proste archiwum dokumentów z funkcją wyszukiwania, stworzyliśmy

prawdziwą wyszukiwarkę, wyposażoną również w funkcję czytania. Jest to kompletne rozwiązanie, oparte na dedykowanym sprzęcie i specjalnie opracowanym oprogramowaniu.

W fazie projektowania oceniliśmy dwie opcje: oprogramowanie do zainstalowania lokalnie lub aplikację po stronie serwera dostępną przez Internet. Ponieważ misją *Biuletynu Salezjańskiego Online* jest szerzenie charyzmatu salezjańskiego wśród jak największej liczby osób, zdecydowano się na rozwiązanie internetowe, aby umożliwić każdemu wyszukiwanie i przeglądanie dokumentów salezjańskich.

Wyszukiwarka jest dostępna pod adresem www.donbosco.info. Interfejs internetowy jest celowo prosty i „spartański”, aby zapewnić większą szybkość ładowania. Na stronie głównej znajduje się lista plików i folderów, aby ułatwić przeglądanie. Dokumenty są dostępne nie tylko w języku włoskim, ale także w innych językach, które można wybrać za pomocą ikony w lewym górnym rogu.

Większość przesłanych plików jest w formacie PDF uzyskanym ze skanów z OCR (optyczne rozpoznawanie znaków). Ponieważ OCR nie zawsze jest idealny, czasami nie wszystkie szukane słowa są wykrywane. Aby temu zapobiec, wdrożono różne strategie: ignorowanie interpunkcji i znaków z akcentem lub specjalnych oraz umożliwienie wyszukiwania nawet w przypadku brakujących lub błędnych znaków. Dalsze szczegóły można znaleźć w sekcji FAQ, dostępnej w stopce.

Ze względu na obecność tysięcy dokumentów, wyszukiwanie może wykazać bardzo dużą liczbę wyników. Z tego powodu można zawęzić zakres wyszukiwania według folderów, tytułu, autora lub roku: kryteria są kumulatywne i pomagają szybciej znaleźć to, czego potrzeba. Wyniki są wyświetlane na podstawie wyniku trafności, który obecnie uwzględnia głównie zagęszczenie słów kluczowych w tekście i ich bliskość.

Idealnie byłoby dysponować dokumentami w formacie wektorowym

zamiast skanami, ponieważ wyszukiwanie byłoby zawsze dokładne, a pliki lżejsze, co przyniosłoby korzyści pod względem szybkości.

Jeśli posiadasz dokumenty w formacie wektorowym lub lepszej jakości niż te, które są już dostępne w wyszukiwarce, możesz je przesłać za pomocą usługi przesyłania dostępnej na stronie www.donbosco.space. Możesz również dodać inne dokumenty, których nie ma w wyszukiwarce. Aby uzyskać dane dostępowe (nazwę użytkownika i hasło), wyślij prośbę e-mailem na adres bsol@sdb.org.