

□ Lesezeit: 3 min.

Wir stellen die neue Plattform donbosco.info vor, eine salesianische Suchmaschine, die die Abfrage von Dokumenten im Zusammenhang mit dem Charisma Don Boscos erleichtern soll. Sie wurde entwickelt, um das Salesianische Online-Bulletin zu unterstützen, und überwindet die Grenzen traditioneller Archivierungssysteme, die oft nicht in der Lage sind, alle Vorkommnisse von Wörtern (Treffer) zu erfassen. Diese Lösung integriert eine Ad-hoc-Hardware und eine speziell entwickelte Software und bietet auch eine Lesefunktion. Die bewusst einfach gehaltene Weboberfläche ermöglicht die Navigation durch Tausende von Dokumenten in verschiedenen Sprachen, mit der Möglichkeit, die Ergebnisse nach Ordner, Titel, Autor oder Jahr zu filtern. Dank der OCR-Scans der PDF-Dateien erkennt das System den Text auch dann, wenn er nicht perfekt ist, und wendet Strategien an, um Satzzeichen und Sonderzeichen zu ignorieren. Die Inhalte, die reich an historischem und informativem Material sind, zielen darauf ab, die salesianische Botschaft flächendeckend zu verbreiten. Durch das freie Hochladen von Dokumenten wird die ständige Bereicherung der Plattform gefördert und die Suche verbessert.

Im Rahmen der Arbeiten zur Erstellung des Salesianischen Online-Bulletins war die Entwicklung verschiedener Hilfsmittel erforderlich, darunter eine spezielle Suchmaschine.

Diese Suchmaschine wurde unter Berücksichtigung der derzeitigen Einschränkungen der verschiedenen salesianischen Ressourcen im Internet konzipiert. Viele Websites bieten Archivierungssysteme mit Suchfunktionen an, aber oft können sie aufgrund technischer Einschränkungen oder Beschränkungen, die zur Vermeidung von Serverüberlastung eingeführt wurden, nicht alle Vorkommnisse von Wörtern (Treffer) finden.

Um diese Schwierigkeiten zu überwinden, haben wir, anstatt ein einfaches Dokumentenarchiv mit einer Suchfunktion zu erstellen, eine echte Suchmaschine entwickelt, die auch über eine Lesefunktion verfügt. Es handelt sich um eine Komplettlösung, die auf einer Ad-hoc-Hardware und einer speziell entwickelten Software basiert.

In der Planungsphase haben wir zwei Optionen geprüft: eine lokal zu installierende Software oder eine *serverseitige* Anwendung, die über das Web zugänglich ist. Da

der Auftrag des Salesianischen Bulletins Online darin besteht, das salesianische Charisma einer möglichst großen Anzahl von Menschen zu vermitteln, haben wir uns für die Weblösung entschieden, um es jedem zu ermöglichen, salesianische Dokumente zu suchen und einzusehen.

Die Suchmaschine ist unter www.donbosco.info verfügbar. Die Weboberfläche ist bewusst einfach und „spartanisch“ gehalten, um eine höhere Ladegeschwindigkeit zu gewährleisten. Auf der „Homepage“ sind die vorhandenen Dateien und Ordner aufgelistet, um die Einsicht zu erleichtern. Die Dokumente sind nicht nur auf Italienisch, sondern auch in anderen Sprachen verfügbar, die über das entsprechende Symbol oben links ausgewählt werden können.

Die meisten hochgeladenen Dateien liegen im PDF-Format vor, das aus Scans mit optischer Zeichenerkennung (OCR) gewonnen wurde. Da die OCR nicht immer perfekt ist, werden manchmal nicht alle gesuchten Wörter erkannt. Um dem abzuhelpen, wurden verschiedene Strategien umgesetzt: Satzzeichen und Akzent- oder Sonderzeichen ignorieren und die Suche auch bei fehlenden oder falschen Zeichen ermöglichen. Weitere Details finden Sie im FAQ-Bereich, der über die Fußzeile zugänglich ist.

Aufgrund der Tausenden von Dokumenten kann die Suche eine sehr große Anzahl von Ergebnissen liefern. Aus diesem Grund ist es möglich, den Suchbereich nach Ordern, Titeln, Autoren oder Jahren einzugrenzen: Die Kriterien sind kumulativ und helfen, das Gesuchte schneller zu finden. Die Ergebnisse werden auf der Grundlage einer Relevanzbewertung aufgelistet, die derzeit hauptsächlich die Dichte der Schlüsselwörter im Text und ihre Nähe berücksichtigt.

Idealerweise wären Dokumente im Vektorformat anstelle von gescannten Dokumenten vorzuziehen, da die Suche immer genau wäre und die Dateien leichter wären, was zu Geschwindigkeitsvorteilen führen würde.

Wenn Sie Dokumente im Vektorformat oder in besserer Qualität als die bereits in der Suchmaschine vorhandenen besitzen, können Sie diese über den Upload-Service auf www.donbosco.space hochladen. Sie können auch andere Dokumente hinzufügen, die nicht in der Suchmaschine vorhanden sind. Um die Zugangsdaten (Benutzername und Passwort) zu erhalten, senden Sie eine E-Mail an bsol@sdb.org.