



# ecm4u-SEARCH

Suchen, wie Sie es erwarten.

Suchergebnisse in Alfresco werden häufig als unzuverlässig erlebt. Jetzt kann die Indizierung und Suche einfach konfiguriert und im Sinne der Anwender verbessert werden.

Anwender von Alfresco sind oft verwundert oder negativ überrascht, wenn ihre Suchanfrage nicht zu den erwarteten Ergebnissen führt. Es kommt vor, dass Dokumente nicht gefunden werden, obwohl genau bekannt ist, dass sie im System abgelegt sind und die abgefragten Worte enthalten. Auch die Suche nach Dateinamen ist häufig nicht erfolgreich. Es gibt Fälle, bei denen mehrere Anwender, die über die gleichen Rechte verfügen, unterschiedliche Suchergebnisse präsentiert bekommen. Oder dass Suchanfragen, die kurz hintereinander durchgeführt werden, zu unterschiedlichen Ergebnislisten führen. Kurzum: die Alfresco-Suche wird oft als unberechenbar erlebt und Anwender misstrauen den Ergebnissen.

Die Gründe für diese Probleme liegen nicht etwa in einem fehlerhaften Programmcode von Alfresco. Vielmehr deckt sich die Art und Weise, wie Alfresco Inhalte ablegt, verarbeitet und wieder danach sucht, oft nicht mit den Erwartungen der meisten Anwender. In der Praxis wird die Suche also anders verwendet als es die Alfresco-Konfiguration vorsieht. Hinzu kommt, dass ein Suchergebnis stark davon abhängen kann, auf welcher Sprachoberfläche von Alfresco sich der Anwender bewegt. Die fest in Alfresco integrierte Suchmaschine Lucene ist so konfiguriert, dass sie abhängig von der Oberfläche Dokumente nach bestimmten Regeln verarbeitet und Dokumente daher mit der gleichen Abhängigkeit wieder gefunden werden - oder eben nicht.

Die praktische Konsequenz all dessen: eine Suchabfrage führt nicht zu den erwarteten Ergebnissen und wird somit von Nutzern als nicht 100%-ig zuverlässig erlebt.

## Die Alfresco-Suche mit *ecm4u-Search*:



Administratoren könnten zwar theoretisch die Suchfunktion konfigurieren und nach den Vorstellungen der Mitarbeiter anpassen. Praktisch passiert dies aber wohl nur selten, da es mit einem erheblichen Aufwand verbunden ist und nur unter der Voraussetzung angegangen werden kann, dass man über einen tiefgreifenden Einblick in die Arbeitsweise von Alfresco verfügt.

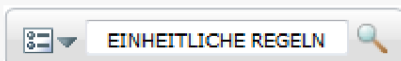
*„Mit ecm4u-Search erhält man nun erstmals eine einfache, praktische Lösung, um das Such- und Indizierungsverhalten von Alfresco nach eigenen Vorstellungen und ohne Programmierung den Erwartungen der Anwender anzupassen. Mit Hilfe dieses Moduls kann man sicher sein, dass die Suche nach nachvollziehbaren Regeln abläuft und immer feststellen, wie ein Suchergebnis zustande kommt.“*

## Wie macht *ecm4u-Search* die Suche besser?

Vereinfacht gesagt, holt das Modul die Sucheinstellungen aus den Tiefen des Alfresco-Codes heraus und lagert sie in eine leicht zugängliche Konfigurationsdatei aus. Das hat zwei Vorteile:

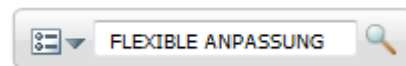
1. Man erhält einen **einfachen Zugriff** auf die "Stellschrauben", mit deren Hilfe man eine erfolgreiche Suche entscheidend beeinflussen kann.
2. Man macht die Dokumentsuche, die sonst über verschiedene Sprachoberflächen hinweg stur nach dem gleichen Konzept abläuft, **flexibel** und kann sie so bestmöglich für das jeweilige Unternehmen oder Projekt konfigurieren.

Um zu zeigen, wie das Modul funktioniert, ist es notwendig, einige technische Aspekte der Alfresco-Suche herauszugreifen. An diesen Beispielen wird sichtbar, wo *ecm4u-Search* ansetzt und worin seine Leistungsfähigkeit besteht.

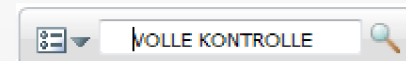


Stoppwörter sind häufig gebrauchte Wörter, die bei der Indizierung ignoriert werden können (zum Beispiel "und", "aber", etc.). Dazu hält Alfresco eine Liste bereit, anhand derer überprüft wird, ob ein Wort indiziert werden soll oder nicht. Auch hier kann es in einem internationalen Arbeitsumfeld zu Problemen kommen, da die aufgerufene Liste der Stoppwörter wiederum von der Alfresco-Oberflächensprache abhängt. Sie gilt also nicht für alle Oberflächen, sondern ist sprachspezifisch. Dies scheint auch plausibel, da es schließlich je nach Sprache unterschiedliche Füllwörter gibt. In der Praxis führt es aber zu Problemen: je nachdem, unter welcher Oberfläche ein Dokument indiziert bzw. gesucht wird, ergibt sich unter Umständen eine unterschiedliche Trefferliste. Zum Beispiel wird das Wort "war" in einem deutschen Index ignoriert, da es als Stoppwort aufgeführt ist. In der englischen Oberfläche wird es allerdings - als Substantiv für "Krieg" - indiziert.

Um dieses Problem zu umgehen und eine einheitliche, verlässliche Suche zu gewährleisten, kann man mit *ecm4u-Search* die sprachabhängige Stoppwortliste von Alfresco durch eine globale Liste ersetzen und frei anpassen. Damit man bei der Erstellung der Liste nicht "von Null" anfangen muss, wird eine - in unseren Augen bereits ausreichende - Liste mit Stoppwörtern mitgeliefert, die sich nach Belieben ändern oder ausbauen lässt.



Alfresco sucht Wörter durch so genannte Tokenizer nach Trennmöglichkeiten ab und zerlegt diese dann in ihre Bestandteile. So wird zum Beispiel der Ausdruck "Vorlage\_Rechnung" standardmäßig in die Bestandteile "Vorlage" und "Rechnung" zerlegt und so indiziert. Das hat zur Folge, dass eine Suche nach "Vorlage" oder "Rechnung" erfolgreich ist, eine Suche nach "Vorlage\_Rechnung" aber nicht! In Alfresco lässt sich diese Funktion an- und abschalten. Dies gilt dann allerdings für alle Dokumente in Alfresco. *ecm4u-Search* macht dahingegen eine feinere Justierung möglich: man kann nun mit einer einfachen Einstellung festlegen, ob z.B. Dokumente eines bestimmten Typs (Rechnungen, Lieferscheine, etc.) oder mit bestimmten Aspekten (z.B. Versionsnummern, Tags, etc.) mit dem Tokenizer bearbeitet werden oder nicht.



Bisher lässt sich in Alfresco nur sehr schwer herauszufinden, nach welchen Regeln ein Dokument indiziert wird. Mit *ecm4u-Search* lässt sich nun die Indizierung "loggen". Dies ist dann hilfreich, wenn ein Dokument fälschlicherweise nicht in der Suchliste erscheint und man feststellen will, warum dies der Fall ist. Der Log gibt Aufschluss darüber, welche Suchwörter für das Dokument gebildet wurden und ermöglicht es zu verstehen, wie Alfresco die erstellten Index- und Suchregeln umsetzt. Damit hat ein Administrator nun die Möglichkeit, Suchergebnisse nachzuvollziehen und den Erwartungen der Anwender entsprechend zu verbessern.

### *ecm4u-SEARCH* ist

- nach der Installation **sofort einsatzbereit** und kommt **ohne zusätzliche Programmierung** aus.
- **upgrade-safe** für ihre Alfresco-Installation, da nur Standarderweiterungspunkte des Systems benutzt werden.
- kompatibel mit Alfresco Community und Enterprise; jeweils 3.4.x und 4.0.x (Lucene-Suche)