

## White Paper: Apache Solr Suchserver

### Zusammenfassung

Dieses White Paper vermittelt Unternehmensentscheidern einen schnellen und präzisen Überblick über die wesentlichen Funktionen und Möglichkeiten der Open-Source-Suchlösung „Apache Solr“ für Webseiten. Wichtige Funktionen sowie Vorteile und Nutzen beim Einsatz dieser Lösung werden aufgezeigt. Im Anschluss wird anhand einer Fallstudie kurz erläutert, wie die Firma initOS GmbH & Co. KG im Rahmen eines Kundenprojekts Solr-Suchfunktionen für einen Magento Online Shop erfolgreich implementiert, konfiguriert und in Betrieb genommen hat.



## Apache Solr Enterprise Suchlösung

Mit Solr stellt die Apache Foundation eine mächtige Open-Source-Plattform für die effiziente Suche von Informationen auf Webseiten, insbesondere mit hohem Informationsaufkommen, zur Verfügung. Hauptargumente für Solr sind dabei Geschwindigkeit und stufenlose Skalierbarkeit seines Suchindexes. In Zahlen ausgedrückt: Mit einer Antwortzeit von 150 Millisekunden (0,15 Sekunden) erhält man sein Suchergebnis in nahezu Echtzeit. Der Suchindex ist dabei in der Lage bis zu 10 Millionen Items (z. B. Produkte eines Online-Shops) zu verwalten. Typische Zielgruppe sind daher Online-Shops mit sehr großem Produktsortiment. Aber auch Webverzeichnisse, Preissuchmaschinen oder kommerzielle Newsprovider sind potentielle Nutzer. Die wichtigsten Suchfunktionen, die Solr bereitstellt, sind:

- Volltextsuche inklusive
  - Ähnlichkeitssuche (ähnliche Begriffe, Schreib- und Tippfehler)
  - Auto-Vervollständigung eingegebener Suchbegriffe (Auto Completion)
  - Berücksichtigung von Stoppwörtern
- Grafische Hervorhebung von Treffern (Hit Highlighting)
- Schrittweise Eingrenzung des Suchergebnisses per Filternavigation (Facettierung, Faceted Search)
- Thematisches Clustern von Suchergebnissen
- Suche in Dokumenten (Word, PDF, ...)

Die Veröffentlichung von Apache Solr basiert auf OpenSource. Das bedeutet, dass die gesamte Architektur inklusive aller Schnittstellen und APIs auf offenen, dokumentierten Standards beruht. So ist es möglich, Solr in Verbindung mit nahezu jeder anderen Webapplikation zu betreiben. Solr ist explizit für den Betrieb auf stark frequentierten Webseiten ausgelegt. Darüber hinaus lässt es sich an eine Vielzahl individueller Kundenanforderungen anpassen.

### Vorteile und Nutzen

Durch den erfolgreichen Einsatz eines Apache-Solr-Servers ergeben sich folgende Vorteile und Nutzen für den Seitenbetreiber:

- Solr glänzt durch seine Geschwindigkeit bei der Präsentation relevanter Suchergebnisse.
- Durch die Auslagerung der Suche wird die eigentliche Webapplikation (z. B. die Shop-Software) nicht durch eingehende Suchanfragen ausgebremst.

- Solr bietet eine Vielzahl ausgereifter Suchmöglichkeiten, welche die integrierte Suche einer Webapplikation in der Regel nicht bietet.
- Die Software lässt sich aufgrund seines offenen Standards relativ flexibel an Kundenwünsche anpassen.
- Auf Wunsch bietet Solr ein hohes Maß an Ausfallsicherheit.

Sämtliche Produktvorteile zielen darauf ab, den Nutzen eines Webseitenbesuchers bzw. das Kundenerlebnis durch schnelle und intelligente Suchfunktionen nachhaltig aufzuwerten. Dadurch steigt die Wahrscheinlichkeit, einen potentiellen Kunden schneller in einen tatsächlichen Kunden zu wandeln. Eine optimierte, schnelle und effiziente Suche trägt dazu bei, die Konversionsrate (und damit den Umsatz) signifikant zu steigern.

### Problemstellung eines Kunden

Ein durch initOS betreuter Online-Shop wird auf Basis der Open-Source-Lösung Magento betrieben. Magento konnte den Kundenbedürfnissen entsprechend umfassend angepasst werden. Nicht zufrieden waren die Betreiber jedoch mit der integrierten Suchfunktion. Der Shop bietet einerseits ein relativ breites (Bedarf für Hunde, Katzen und Pferde, Fischnahrung sowie entsprechendes Zubehör), wie auch ein tiefes Sortiment (hohe Anzahl verschiedener Nass- und Trockenfuttermarken und Zubehör). Es wurde festgestellt, dass ca. 35 % der Conversions durch Nutzung der Produktsuche erfolgen. Die Suchfunktion ist daher von hoher Bedeutung und integraler Bestandteil der Navigation. Aufgrund der hohen Artikeldichte ermöglicht sie es dem Nutzer, einfach und effektiv ans gewünschte Ziel zu kommen. Trotz der bereits hohen Konversionsrate funktionierte die integrierte Suche nicht zufriedenstellend. Messungen ergaben, dass die Suche im Durchschnitt 3,5 Sekunden benötigte, um das gewünschte Suchergebnis zu liefern. Dies dauerte aus Kundensicht deutlich zu lange. Als weiteres Problem erwies sich ein zeitlich hohes Suchaufkommen. Eine flüssige Webseitennavigation war nicht mehr möglich, sobald der Server mehr als 20 Suchanfragen zeitgleich abuarbeiten hatte. Eine weitere Anforderung war Facettierung, also die Realisierung einer schrittweisen Eingrenzung des Suchergebnisses per Filternavigation. Diese Funktion ist mit der integrierten Suchfunktion in Magento nur eingeschränkt umzusetzen. Kunden des Shops suchen trotz der Fülle des Angebots sehr spezifische Produkte (in diesem Fall z. B. Hundenassfutter für Welpen mit Allergien oder Katzentrockenfutter einer bestimmten Marke und Fleischsorte).

## Projektbeschreibung

Das Kundenprojekt wurde in folgende Phasen unterteilt

### Konzeptphase:

Nachdem die Anforderungen des Kunden eindeutig definiert waren, wurde als erstes eine geeignete Lösung evaluiert. Als Alternativen wurden die Open-Source-Lösungen Elastic Search und Apache Solr sowie der Einkauf einer optimierten Magento-Suche als SaaS (Software as a Service) ermittelt. Die SaaS-Lösung schied als erstes aus, da der Kunde nicht gewillt war, seine Suchinformationen bei einem Drittanbieter auszulagern. Elastic Search bietet im Vergleich zu Solr keine wesentlichen Vorteile, solange sich ein Suchprojekt unter 10 Millionen Items bewegt. Um die Machbarkeit der Umsetzung mit Solr nachzuweisen, wurde mit dem Kunden ein Proof of Concept vereinbart. Hierfür wurden sämtliche Produkte des Online-Shops (ca. 10.000) in die Solr Datenbank und somit in den Suchindex übernommen. Anschließend wurde mittels Ajax eine Bedienoberfläche erstellt, um damit Produkte gemäß Anforderungen suchen und filtern zu können.



### Implementierung der Lösung

Nach erfolgreichem Test wurde die Lösung beim Kunden implementiert. Folgende Schritte wurden realisiert, um Solr sowie die im Proof of Concept gestaltete Suchoberfläche mit Ajax erfolgreich in den Magento-Shop zu integrieren.

### 1. Pre-Rendering der Produktdarstellung und Implementierung des Suchindexes

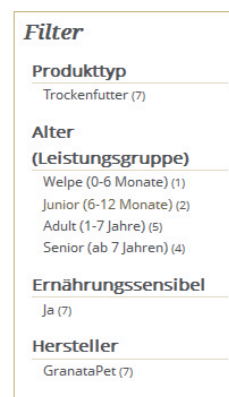
Aus Geschwindigkeitsgründen musste sicher gestellt werden, dass die für die Darstellung der Produktergebnisse notwendigen Informationsblöcke (HTML-Blöcke) bereits zum Zeitpunkt des eigentlichen Indexaufbaus (auf Basis der Suche) berechnet werden. Ein zusätzliches Modul wurde entwickelt, um alle benötigten Informationen von Magento an Solr zu übermitteln.

### 2. Weiteres Ajax-Interface zur Darstellung der Suchergebnisse

Dieses Modul übernimmt die Suchergebnisse vom Solr-Server und ist für die korrekte Darstellung innerhalb der Webseite verantwortlich. Der Browser kommuniziert direkt mit der Suche, Ressourcen des Shops werden geschont.

### Test und Freigabe

Die abschließende Testprozedur zeigte, dass die durchschnittliche Geschwindigkeit von Berechnung und Darstellung der Suchergebnisse von ehemals 3,5 Sekunden auf 0,15 Sekunden reduziert werden konnte. Die facet- tierte Suche wurde frei gegeben und befindet sich nun ebenfalls im produktiven Einsatz.



## Über initOS

Seit über 15 Jahren beschäftigt sich die initOS GmbH & Co. KG mit moderner Informationstechnologie. Das Hauptunterscheidungsmerkmal im Vergleich zu anderen IT-Dienstleistern liegt dabei auf einem sehr starken Fokus bezüglich der Anpassung und Implementierung von Open-Source-Lösungen. Diese werden zielgerichtet dort eingesetzt, wo sie Kunden einen insgesamt höheren Nutzen stiften, eine verbesserte Nutzung von Ressourcen gewährleisten, oder wo Open Source eine schnellere Erreichbarkeit strategischer Ziele ermöglicht als vergleichbare proprietäre Produkte.

Unser Ziel ist es nachhaltige und vertrauensvolle Geschäftsbeziehungen mit den Kunden einzugehen. Wir möchten unsere Kunden nicht durch den Aufbau einseitiger Abhängigkeiten sondern durch die bedingungslose Qualität der geleisteten Arbeit an uns binden. Durch den konsequenten Einsatz von Open-Source-Software haben unsere Kunden jederzeit die volle Kontrolle über ihre Informationssysteme.

Quellen:

Foto Labyrinth: 3DAgentur / aboutpixel.de  
Logo Apache Solr: Logo/Trademark der Apache Software Foundation / apache.org



initOS GmbH & Co. KG

An der Eisenbahn 1  
D-21224 Rosengarten

e-Mail: [info@initos.com](mailto:info@initos.com)  
Web: [www.initos.com](http://www.initos.com)