

Elasticsearch

Was ist die Elasticsearch - App?

Die kostenpflichtige App stellt ein Webinterface zu einem Elasticsearch-System her und wird als Alternative zur Standard Shop Suche eingesetzt. Die Suchmaschine basiert auf der Lucene-Engine und ist neben Solr, der am weitesten verbreitete Suchserver.

Die Verwendung der App setzt einen vom Shopbetreiber bereitgestellten Server voraus. Elasticsearch fällt unter die Apache-Lizenz und kann serverseitig kostenlos verwendet werden.

Die Verwendung von Elasticsearch ersetzt folgende Frontend-Bereiche im Shop:

- Suche
- Suchvorschau
- Schnellerfassung
- After Search - Navigation
- Produkt Finder-Navigation

Einrichtung und allgemeine Einstellungen

ALLGEMEINE EINSTELLUNGEN	INDEX EINSTELLUNGEN	SORTIMENTE ÜBERTRAGEN
EINSTELLUNGEN		
ElasticSearch aktivieren	<input type="checkbox"/>	
ElasticSearch Server / Port	<input type="text" value="localhost"/>	<input type="text" value="9200"/>
Node für Artikeldaten	<input type="text" value="node-1"/>	
Erfolgreiche Übertragungen	<input type="text" value="elastic"/>	
ES Password	<input type="password" value="....."/>	

Elasticsearch aktivieren: Aktiviert die Suchengine im Shop und tauscht die Standard-Suche in allen Frontend-Bereichen aus.

ACHTUNG: Aktivieren Sie die App nur nach erfolgreichen Tests. Sollten Sie falsche Zugangsdaten eingegeben haben oder sollte der Server nicht erreichbar sein, steht Ihnen keine Shop-Suche in diesem Zeitraum zur Verfügung.

Elasticsearch Server / Port: Adresse und Port des vom Shopbetreiber zur Verfügung gestellten Suchservers.

Node für Artikeldaten: Der auf dem Suchserver eingerichtete Node, in dem die verschiedenen Suchindizes angelegt werden sollen.

ES Benutzer / ES Password: Falls das X-Pack installiert oder Zugangsdaten für den Suchserver eingerichtet worden sind, hinterlegen Sie an dieser Stelle die entsprechenden Daten.

Suchfelder

Neben den Zugangsdaten zur Einrichtung des Servers müssen auch die Einstellungen für die Suche selbst hinterlegt werden. Sie können innerhalb der untenstehenden Tabelle die Felder definieren, die für eine erfolgreiche Suche herangezogen werden sollen. Dazu setzen Sie bitte innerhalb der Spalte „Zu durchsuchende Felder“ ein Häkchen.

Die Berechnung und Erstellung der Scoring-Werte übernimmt weitestgehend der Suchserver selbst. Auf diese Berechnung kann direkt kein Einfluss genommen werden. Sie haben allerdings die Möglichkeit für bestimmte Felder eine höhere Gewichtung vorzunehmen. Dies können Sie über die Spalte „Scoring“ vornehmen. Hierbei handelt es sich um einen Faktorwert, der den durch die Suchengine errechneten Score für dieses Feld um den eingetragenen Wert multipliziert. Keine Eintragung entspricht dem Wert „1“.

Suchindex	Zu durchsuchende Felder	Scoring*
	<input checked="" type="checkbox"/> Artikelnummer	3
	<input checked="" type="checkbox"/> Produktbezeichnung	2
	<input type="checkbox"/> Produktbeschreibung	
	<input type="checkbox"/> Produkteigenschaften	
	<input type="checkbox"/> Keywords	
	<input checked="" type="checkbox"/> EAN	
	<input checked="" type="checkbox"/> Lieferanten Artnr.	
	<input type="checkbox"/> Katalog Nummer / Kundenartikel Nummer	

* Faktor mit dem das Scoring dieses Feldes im Suchergebnis erhöht werden soll

Mindest Score
(% vom Max Score) %

Zeige nur Ergebnisse, die der prozentuellen Angabe vom Max Score entsprechen

Zum Beispiel sind die Felder Produktbezeichnung, Produktbeschreibung und Keywords mit den

zugewiesenen Scoring Faktoren 2 für Produktbezeichnung und 4 für die Keywords. Eine Suche nach dem Wort „Bohrer“ liefert beispielsweise folgendes Ergebnis:

	Artikel 1 ohne Faktor	Artikel 1 mit Faktor	Artikel 2 ohne Faktor	Artikel 2 mit Faktor
Produktbezeichnung	2,5	5	2	4
Produktbeschreibung	4	4	8	8
Keyword	2	8	1	4
Scorewert	8,5	17	11	16

Bei der Suche ohne Faktor würde Artikel 2 vor Artikel 1 angezeigt werden, durch mehrfache Trefferleistung innerhalb des Feldes „Produktbeschreibung“. Gibt man der Bezeichnung und den Keywords für die eigene Suche eine höhere Priorisierung, ändert sich das Suchergebnis und Artikel 1 würde vor Artikel 2 dargestellt werden.

Mindest-Score

Aufgrund verschiedener Filter und Analyser ist bei Elasticsearch das jeweilige Suchergebnis deutlich größer als bei der Standard-Shopsuche. In diesem Ergebnis können sich Resultate befinden, die gegebenenfalls nicht erwünscht sind oder mit der Sucheingabe wenig zu tun haben.

Damit solche Resultate im Suchergebnis nicht erscheinen, kann die Qualität des Suchergebnisses durch die Angabe eines Mindest-Scores erhöht werden. Durch die internen Berechnungsroutinen von Elasticsearch lässt sich dieser Wert nicht fest definieren, daher wird in prozentualer Abhängigkeit des besten Ergebnisses dieser Wert errechnet.

Zum Beispiel ist der höchste Score bei einer Suche 100 und Ihre Angabe für den Mindest Score beträgt 75%, dann müssen alle nachfolgenden Suchresultate mindestens einen Score von 75 aufweisen.

Index Einstellungen

Für die optimale Suche im Shop ist der Aufbau des Index und die Verwendung der verschiedenen Analyser und Tokenizer weitestgehend festgelegt. Änderungen an diesen Einstellungen setzen größtenteils einen Neu-Import oder eine Re-Indexierung voraus. Dieser Prozess kann je nach Datenbankgröße, mehrere Tage beanspruchen. Ausgenommen von der Re-Indexierung ist die Anpassung der Tokenizerlängen.

Für die Suche über die Hauptfelder Produktbezeichnung und Keywords wird ein eigener Analyser verwendet, der sich aus dem folgenden Elementen zusammensetzt:

Folgende Filter werden für die Suche verwendet: Standard, Lowercast und einem Stemmer, entsprechend der Sprache des Indexes. Um Teil-Elemente von Suchbegriffen z.B. Schuh in Arbeitsschuh zu finden, wird der „Ngram - Tokenizer“ eingesetzt.

Produktbeschreibungen und die Suche über Artikelmerkmale werden anhand der „Language Analyzers“ der jeweiligen Index-Sprache durchsucht.

Tokenizer „Artikelnummern“

Hier werden die Nummern betreffenden Tokenzier-Einstellungen vorgenommen. Verwendet wird der „Edge NGram“ - Tokenizer. Dies bedeutet die einzelnen Suchtokens werden nur von Beginn der Nummer vorgenommen und anhand der eingestellten Mindest- und Maximalgröße erstellt.

Beispiel: '2 Quick Foxes.' / Tokengröße: Min/Max 3 Tokens: Qui, Fox

Minimale Tokengröße	<input type="text" value="3"/>
Maximale Tokengröße	<input type="text" value="13"/>
Einzuschliessende Zeichen	<input checked="" type="checkbox"/> Zahlen <input type="checkbox"/> Buchstaben <input type="checkbox"/> Leerzeichen <input type="checkbox"/> Interpunktion <input type="checkbox"/> Symbole

Minimale Tokengröße: Anzahl der Zeichen aus denen die kleinsten Suchtokens gebildet werden sollen.

Maximale Tokengröße: Anzahl der Zeichen aus denen die größten Suchtokens gebildet werden sollen.

Einzuschließende Zeichen: Zeichen die zur Bildung von Zeichen herangezogen werden sollen. Nicht berücksichtigte Zeichen werden als Trenner angesehen.

Tokenizer „Texte“

Hier werden Tokenzier-Einstellungen für den Kurztext und die Keywords vorgenommen. Verwendet wird der „NGram“ - Tokenizer. Hierbei werden die verschiedenen Suchbegriffe und die im Index übertragenen Begriffe in Suchtokens verschiedener Größen anhand der eingestellten Mindest- und Maximalgröße erstellt.

Beispiel: „2 Quick Foxes.“ / Tokengröße: Min/Max 3 Tokens: Qui, uic, ick, Fox, oxe, xes

Minimale Tokengröße	<input type="text" value="4"/>
Maximale Tokengröße	<input type="text" value="6"/>
Einzuschliessende Zeichen	<input checked="" type="checkbox"/> Zahlen <input checked="" type="checkbox"/> Buchstaben <input type="checkbox"/> Interpunktion <input type="checkbox"/> Symbole

Minimale Tokengröße: Anzahl der Zeichen aus denen die kleinsten Suchtokens gebildet werden sollen.

Maximale Tokengröße: Anzahl der Zeichen aus denen die größten Suchtokens gebildet werden sollen.

Einzuschließende Zeichen: Zeichen die zur Bildung von Zeichen herangezogen werden sollen. Nicht berücksichtigte Zeichen werden als Trenner angesehen.

Informationen zu den Analysis-Modulen

Genauere Informationen zu den Analysis-Modulen und deren Arbeitsweise finden Sie in der offiziellen Dokumentation von Elasticsearch unter:

<https://www.elastic.co/guide/en/elasticsearch/reference/current/analysis.html>

Vorschau

In allen Einstellungsbereichen haben Sie auf der rechten Seite die Möglichkeit eine kurze Testsuche auszuführen, um die Auswirkungen Ihrer Einstellungen zu überprüfen.

Die Suche verhält sich entsprechend der definierten Standardeinstellungen für Varianten und Sortierung wie die Hauptsuche im Shop Frontend. Angezeigt wird hierbei die erste Ergebnisseite. Neben der Anzeige der ersten Resultate erhalten Sie Informationen über die Ausführungszeit, Anzahl der ungruppierten Treffer und dem Scorewert des besten Treffers.

Die Ausführungszeit bei neuen Eingaben kann aufgrund fehlenden Caches etwas erhöht sein.

VORSCHAU **SUCHVORSCHAU**

Ausführungszeit (in ms) 715

Anzahl Treffer (ungruppiert) 2

Höchste Relevanz 46.643871307373

Geräuschpegelmessgerät Sound Test-Master 	Öltester RTK Retrofit-Test-Kit 
---	--

Sortimente übertragen

Über diesen Bereich haben Sie die Möglichkeit bestehende Sortimente manuell an den Elasticsearch-Server zu übertragen.

In dieser Liste werden alle Sortimente des Shops nach der Spracheinstellung sortiert angezeigt. Zur Übertragung wählen Sie bitte das entsprechende Sortiment aus der Liste aus. Für eine Mehrfachauswahl halten Sie die STRG-Taste gedrückt.

- Global - 24.07.2017 - Standard Sortiment
- DE - 24.07.2017 - Beispielkatalog
- DE - Neu - Maschinen_1
- IT - 14.07.2017 - Test

Nur neue Daten

DATEN ÜBERTRAGEN

Pro Auswahlzeile können Sie drei Informationen entnehmen.

1. Der Index, in den das Sortiment übertragen wird. Global steht für die Sortimente, welche für alle Sprachen freigegeben sind.
2. Entweder das Stichwort „Neu“ oder das Datum der letzten Übertragung des Sortiments.
3. Name des Sortimentes

Die Option „Nur neue Daten“ vergleicht das letzte Übertragungsdatum mit der letzten Aktualisierung des Sortiments und überträgt anschließend alle Datensätze, welche neuer als die letzte Übertragung sind.

Index Statistiken

Auf der rechten Seite finden Sie die Statistiken zu den verschiedenen Indizes. Es werden dabei nur die Indexe angezeigt, welche auf dem Suchserver auch Daten enthalten.

Globaler Index (Statistiken einklappen)

Artikel Daten: 94
Speichergröße: 237.65KB

Enthaltene Sortimente	Artikel (ohne Vars)	Letztes Update
Maschinen	1	24.07.2017 16:53
Test	1	24.07.2017 16:53

⊗ Index Löschen

Jedem Block können Sie die folgenden Daten entnehmen.

1. Artikel Daten ohne VAR - Artikel
2. Speichergröße in der Node
3. Die zehn größten übertragenen Sortimente mit der Anzahl Artikel und dem letzten Übertragungsdatum

Sie haben an dieser Stelle auch die Möglichkeit einen kompletten Index zu löschen.

Kibana

Um Statistiken zu der Suche zu verarbeiten und darzustellen, bietet Elasticsearch ein weiteres Server Tool auf ihrer Homepage an. Mit dem Tool „Kibana“ lassen sich ihre an Elasticsearch übertragenen Daten visualisieren.

Für die Visualisierung muss jeder Index separat in Kibana eingerichtet werden. Hierzu müssen Sie ihre Kibana App starten und den Reiter „Management“ aufrufen. Unter dem Tab „Index Patterns“ fügen

Sie nun einen neuen Index hinzu, in dem Sie auf das + - Symbol drücken.

In der nachfolgenden Maske tragen Sie unter „KIBANA-WILDCARD_DYNAMIC_INDEX_PATTERNS“ den Namen ihres Indexes ein und entfernen falls notwendig den Haken bei „KIBANA-CONTAINS_TIME_BASED_EVENTS “. Der Button zur Anlage „KIBANA-CREATE“ sollte nun erscheinen und anklickbar sein.

Die Namen der Indexe besteht entweder nur aus dem Länderkürzel der jeweiligen Sprache (de, en, it, etc.) oder aus dem Keyword „multi_lang“, welches für die Sortimentssprachzuweisung „Für alle Sprachen“ steht.

Nach der erfolgreichen Einrichtung erhalten Sie nun unter „Index Patterns“ eine Auflistung sämtlicher Felder, die übertragen worden sind.

Des Weiteren steht ihnen der Index nun in sämtlichen Bereichen von Kibana zur Auswahl zur Verfügung.

Informationen zu Kibana

Alle Informationen zu Kibana sowie den Download und die Dokumentation finden Sie auf der offiziellen Produktseite: <https://www.elastic.co/de/products/kibana>