



Onventis Buyer – Katalogsuche

Grundsätze der Katalogsuche, Gewichtungstabelle und Analyse eines Anwendungsfalles

Version	1.0
Datum der Version	21.02.2020
Erstellt durch	Onventis GmbH
Vertraulichkeitsstufe	Copyright Onventis 2020
Integritätsstufe	Informationell unkritisch
Verfügbarkeitsstufen	Wiederherstellen nach Bedarf

Copyright:

Alle in diesem Handbuch genutzten Warenzeichen und Produktnamen sind unter Umständen eingetragene Warenzeichen der Herstellerfirmen und dürfen nicht gewerblich oder in sonstiger Weise verwendet werden. Die in diesem Dokument enthaltenen Beschreibungen zeigen die Interpretation der Software-Nutzung der Onventis GmbH zum Zeitpunkt der Erstellung dieses Leitfadens. Da die Onventis GmbH auf sich ändernde Marktanforderungen reagiert, stellt dies keine Verpflichtung seitens der Onventis GmbH dar. Bei der Zusammenstellung dieses Leitfadens und der Abbildungen wurde mit größter Sorgfalt gearbeitet. Dennoch können Fehler nicht vollständig ausgeschlossen werden. Die angegebenen Daten dienen lediglich der Produktbeschreibung und sind nicht als zugesicherte Eigenschaft im Rechtssinne zu verstehen. Die Onventis GmbH und dessen Autoren können für fehlerhafte Angaben und deren Folgen weder juristisch noch in irgendeiner anderen Form haftbar gemacht werden. Alle Rechte vorbehalten; kein Teil dieses Handbuchs darf in irgendeiner Form (Druck, Fotokopie oder die Speicherung und/oder Verbreitung in elektronischer Form) ohne schriftliche Genehmigung der Onventis GmbH verändert, reproduziert oder vervielfältigt werden. Eine öffentliche Verwendung des Dokumentes darf nur mit Zustimmung der Onventis GmbH erfolgen.

Inhalt

1.	Einführung	3
2.	Technologie der eingesetzten Katalogsuche	3
3.	Grundsätze der Katalogsuche	4
4.	Gewichtungstabelle	10

1. Einführung

Diese Dokumentation beschreibt das Verhalten der Onventis Katalogsuche, insbesondere aus Anwendersicht und zeigt die grundsätzlichen Such Paradigmen auf, welche von Benutzern der Onventis Katalogsuche berücksichtigt werden können. Zudem wird aufgezeigt, wie die "Apache Solr Search" bei der Indexierung von Artikeldaten vorgeht, insbesondere für die Bildung von Textfregmenten, sogenannten "ngrams".

Zudem wird auf die Gewichtungstabelle der "Apache Solr Search" eingegangen, welche aufzeigt, welche Felder indexiert werden, welche Gewichtung diese haben und für welche Felder Textfragmente für die Suche gebildet werden.

Abschließend wird auf die Recherche von beispielhaften Katalogsuchen im Onventis bzw. E/D/E-Katalog eingegangen und die entsprechenden Suchergebnisse analysiert.

Nicht enthalten sind die technische Beschreibung und die Dokumentation der Implementierung der eingesetzten Suche-Technologie.

2. Technologie der eingesetzten Katalogsuche

Im Katalogmanagement bzw. für die Katalogsuche in Onventis Buyer wird die Apache Solr in Version 4.5 eingesetzt.

Diese Version wird auch für die 'Volltextsuche' in Onventis Buyer eingesetzt.

3. Grundsätze der Katalogsuche

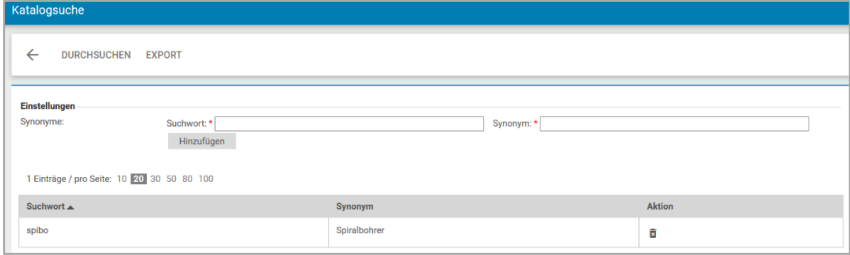
#	Such Paradigma	Beispiele
1	Textfragmentierung	<p>Ein Wort wird vollständig indexiert, unabhängig von der Länge des Wortes (z. B. "W", "mm", "IT", "Kit").</p> <p>Ab 4 Zeichen wird das Wort zusätzlich von vorne und von hinten in Textfragmente aufgeteilt, sogenannte "n-gram".</p> <p>Beispiele der Textfragmentierung durch die 'Apache Solr Search Engine' anhand einer Artikelnummer. Es werden sowohl die vollständige Artikelnummer, als auch die zerlegten Textfragmente in den Suchindex aufgenommen.</p> <ul style="list-style-type: none"> • 4540849 <ul style="list-style-type: none"> ○ 4 ○ 45 ○ 454 ○ 4540 ○ 45408 ○ 454084 ○ 540849 ○ 40849 ○ 0849 ○ 849 ○ 49 ○ 9 <p>Beispiel 1: Die Suche nach der Zeichenfolge 4207 ergibt Suchtreffer für Artikel mit der Lieferanten Artikel-Nr.:</p> <ul style="list-style-type: none"> • 5864207 (vgl. in der Gewichtungstabelle den Wert "article_supplier_article_nr_ngram_front" im Feld "Lieferanten Artikel-Nr.; Textfragment hinten") • 4207095 (vgl. in der Gewichtungstabelle den Wert "article_supplier_article_nr_ngram_back" im Feld "Lieferanten Artikel-Nr.; Textfragment vorne") <p>Beispiel 2: Suche nach der Zeichenfolge 40849 ergibt Suchtreffer für Artikel mit der Lieferanten Artikel-Nr.:</p> <ul style="list-style-type: none"> • 40849 (vgl. in der Gewichtungstabelle den Wert "article_supplier_article_nr" im Feld "Lieferanten Artikel-Nr.") • 3240849 (vgl. in der Gewichtungstabelle den Wert "article_supplier_article_nr_ngram_back" im Feld "Lieferanten Artikel-Nr.; Textfragment hinten")

		<ul style="list-style-type: none"> 4540849 (vgl. in der Gewichtungstabelle den Wert "article_supplier_article_nr_ngram_back" im Feld "Lieferanten Artikel-Nr.; Textfragment hinten") <p>Weiteres Beispiel der Textfragmentierung durch die Apache Solr Search Engine:</p> <ul style="list-style-type: none"> Soap <ul style="list-style-type: none"> S So Soa oap ap p
2	Textfragmentierung - Limitierungen	<ul style="list-style-type: none"> Schreibweise im Artikel lautet "DIN333-A"; Der Bindestrich fungiert als Trennzeichen, somit werden 2 Worte indexiert mit folgenden Fragmenten: DIN333 <ul style="list-style-type: none"> D DI DIN DIN3 DIN33 IN333 N333 333 33 3 A <p>Die Textfragmentierung funktioniert dabei nicht, wenn der indexierten Zeichenfolge bei der Suche zusätzliche Zeichen hinzugefügt werden, oder Zeichen innerhalb der Zeichenfolge ausgetauscht werden, bspw. DIN333A oder DIN323</p> <ul style="list-style-type: none"> Suchergebnisse nach div. Suchen: <ul style="list-style-type: none"> DIN3 = 6.454 Treffer DIN33 = 3.166 Treffer DIN333 = 91 Treffer DIN333- = 91 Treffer DIN333* = 91 Treffer DIN333-a = 46 Treffer DIN333A = 0 Treffer DIN323 = 0 Treffer Hierfür müsste die Schreibweise im Artikel DIN333A oder DIN323A lauten; Dabei werden indexiert: DIN333A <ul style="list-style-type: none"> D DI DIN

		<ul style="list-style-type: none"> ○ DIN3 ○ DIN333 ○ IN333A ○ N333A ○ 333A ○ 33A ○ 3A ○ A <ul style="list-style-type: none"> • DIN323A <ul style="list-style-type: none"> ○ D ○ DI ○ DIN ○ DIN3 ○ DIN32 ○ DIN323 ○ IN323A ○ N323A ○ 323A ○ 23A ○ 23A ○ 3A ○ A <ul style="list-style-type: none"> • 3,15mm <ul style="list-style-type: none"> ○ 3 ○ 3, (= > Angehängte Zeichen wie Kommas (,) werden allerdings bei der Suchanfrage ignoriert, vgl. Hinweise in Zeile 7) ○ 3,1 ○ 3,15 ○ 3,15m ○ ,15mm (= > Angehängte Zeichen wie Kommas (,) werden allerdings bei der Suchanfrage ignoriert, vgl. Hinweise in Zeile 7) ○ ,15m ○ ,15 ○ ,1 ○ ,
3	"Stemming"	<p>Dieses Verfahren, bei dem verschiedene morphologische Varianten eines Wortes auf ihren gemeinsamen Wortstamm zurückführen, wird <u>nicht</u> unterstützt.</p> <ul style="list-style-type: none"> • Suche nach "Haus" findet auch "Häuser" • Suche nach "Wörter" findet auch "Wort"
4	Initiale Sortierung der Suchergebnisse	<ul style="list-style-type: none"> • Die Suchergebnisse werden initial nach Relevanz sortiert. • Die Relevanz orientiert sich an der Gewichtung des Suchergebnisses (siehe Tabelle in Kapitel 4). • Die Gewichtung errechnet sich aus den summierten Werten der Gewichtungstabelle. • Je höher die Summe der Gewichtung ist, desto höher ist die Relevanz. • Je höher die Relevanz, desto höher ist die Position des Artikels in der Suchergebnisliste. • Wählt der Benutzer eine andere Sortierung, bspw. "Preis aufsteigend", so wird diese eine gewisse Zeit im Browser vorgehalten und bei

		nachfolgenden Suchen automatisch verwendet.
5	Sortierung bei gleicher Gewichtung von Suchergebnissen.	Wenn das Ergebnis der die Gewichtung von Suchergebnisse gleich ist, so werden die Suchergebnisse alphabetisch nach "Lieferanten Artikel-Nr." sortiert.
6	Verundung der Suchbegriffe	<p>Es gibt eine sogenannte Verundung bei der Trennung von einzelnen Suchbegriffen durch Leerzeichen.</p> <p>Bei der Verwendung mehrere Suchbegriffe und Trennung dieser durch Leerzeichen werden nur Artikel gefunden, die alle Suchbegriffe beinhalten. Die Reihenfolge, in der die Begriffe innerhalb eines Artikels auftreten, spielt dabei keine Rolle.</p> <p>Suche nach: FORMAT D333A Findet z. B. den Artikel "Zentrierb. D333A HSS 1,00mm FORMAT"</p> <p>Je mehr Suchbegriffe gleichzeitig eingegeben werden, desto geringer wird die Trefferzahl, da eine Schnittmenge aller Begriffe gebildet wird:</p> <ul style="list-style-type: none"> • Beispiele: <ul style="list-style-type: none"> ○ Suche nach "Bleistift" = 77 Treffer ○ Suche nach "gelb" = 1.964 Treffer ○ Suche nach "Bleistift gelb" = 16 Treffer <p>Dabei werden nur Suchergebnisse ausgegeben, die mindestens in einem Feld "Bleistift" (oder ein Textfragment davon) enthalten UND in mindestens einem Feld das Wort "gelb" (oder ein Textfragment) davon enthalten.</p>
7	Wenn Zeichen wie Punkt (.) Bindestrich (-) oder Komma (,) direkt vor oder hinter einen Suchbegriff angehängt werden, so werden diese Zeichen bei der Suche ignoriert.	<ul style="list-style-type: none"> • Suche nach "Spiralb." = 35.580 Treffer • Suche nach "Spiralb" = 35.580 Treffer • Suche nach "Spiralb-" = 35.580 Treffer • Suche nach "Spiralb ." = 0 Treffer => Es wird kein Suchergebnis ausgespielt aufgrund des Leerzeichens. • Suche nach "Spiral -" = 0 Treffer => Es wird kein Suchergebnis ausgespielt aufgrund des Leerzeichens. • Suche nach ",15m" => Suchergebnis mit ",15m" als auch "15m" im bspw. Artikelnamen werden ausgespielt. • Suche nach "," => 0 Treffer
8	Sternchen (*)	Das Sternchen (*) ersetzt beliebig viele Zeichen bzw. fungiert als Platzhalter für beliebig viele Zeichen.
9	Fragezeichen (?)	Das Fragezeichen (?) ersetzt genau ein Zeichen.
10	Trennzeichen (.) oder (,)	Die Trennzeichen Punkt (.) und Komma (,) werden bei der Suche berücksichtigt.
11	Minus (-)	Suchtreffer mit dem Begriff hinter dem (-) Symbol werden aus dem Suchergebnis ausgeschlossen.

		Beispiel-Suche: "40849 -Alu" => Es werden alle Suchtreffer ausgeblendet welche in einem Feld die Zeichenfolge Alu enthalten.																																							
12	Phrasensuche (" ")	(" ") Phrasensuche durch Setzen von Anführungszeichen. Suche nach der Zeichenkette innerhalb der Anführungszeichen.																																							
13	Schlagwörter (Feld: Suchbegriffe)	<ul style="list-style-type: none"> Wer: Katalog Provider Schlagwörter können auf Artikel-Ebene eingesehen werden: Stammdatenverwaltung > Kataloge > Katalog auswählen > Artikel auswählen > Reiter: Artikel > Bereich "Artikel" > Feld "Suchbegriffe" Hier können durch Schlagwörter eingegeben werden, welche bei der Suche ebenfalls Artikel in der Suchergebnisliste ausgespielt werden. Beispiel 1: Suche nach "Bleistift" bildet einen Artikel in der Suchergebnisliste ab, welcher als Schlagwort "Bleistift" hinterlegt hat, und diese Information sonst in keinem Artikelattribut enthält (Lieferanten Artikel-Nr.: 97640006) <p>Suchbegriffe: Lyra Beschriftungselement Schulkreide QR4029282 Bleistift Beschriftungselement, Bleistift,</p> <ul style="list-style-type: none"> Beispiel 2: Suche nach "wapuzange" findet Artikel mit der Lieferanten Artikel-Nr.: 3055990240 <div style="border: 1px solid black; padding: 5px; margin: 10px 0;"> <p>Artikel</p> <p>Titel: Wasserpumpenzange-CV-rot 240mm DIN/ISO 8976 FORMAT</p> <p>Beschreibung: Wasserpumpenzange Format, Kopf geschliffen</p> <p>Ausführung: DIN ISO 8976. Das durchgesteckte Gleitgelenk ermöglicht durch die doppelte Führung eine hohe Stabilität. Der Klemmschutz verhindert Quetschverletzungen.</p> <p>Aus Chrom-Vanadium-Stahl, geschmiedet und hochwertig vergütet, Kopf geschliffen.</p> <p>Herstellernamen: -</p> <p>Suchbegriffe: DINISO8976 Standard Zangen Chrom-Vanadium-Stahl Format Format, 8976 ISO Zange DIN Wasserpumpenzangen FORMAT Wasserpumpenzange Wapuzange</p> </div> <ul style="list-style-type: none"> Feldbezeichnung in der Apache Solr: "ArticleKeywords" Anlage der Schlagwörter beim Katalog-Import im Import_Article_Template.xlsx <div style="border: 1px solid black; padding: 5px; margin: 10px 0;"> <table border="1"> <thead> <tr> <th>E</th> <th>F</th> <th>G</th> </tr> </thead> <tbody> <tr> <td>Keywords_de-DE</td> <td>[Optionales Feld]</td> <td></td> </tr> <tr> <td>Bildschirm, Monitor</td> <td>S</td> <td>Alphanumerisch, max. 1000 Zeichen</td> </tr> <tr> <td>Bildschirm, Monitor</td> <td>S</td> <td>Schlüsselwörter zum Artikel. Schlüsselwörter sind für den Kunden nicht sichtbar, werden bei der Suche aber berücksichtigt.</td> </tr> <tr> <td>OS-System, Server</td> <td>L</td> <td></td> </tr> <tr> <td>Drucker, Printer</td> <td>S</td> <td>Sprachcode siehe Spalte "Titel".</td> </tr> <tr> <td>Laptop</td> <td>K</td> <td></td> </tr> <tr> <td></td> <td>L</td> <td>[Optional Field]</td> </tr> <tr> <td></td> <td></td> <td>Alpha numeric, max. 1000 characters</td> </tr> <tr> <td></td> <td></td> <td>Article keywords. Keywords are not visible for customer, but are included in search.</td> </tr> <tr> <td></td> <td></td> <td>Language code see column "Title".</td> </tr> </tbody> </table> </div> <ul style="list-style-type: none"> BMEcat Katalog im Feld ARTICLE_DETAILS / KEYWORD <table border="1" style="margin: 10px 0;"> <thead> <tr> <th>KEYWORD</th> <th>String</th> <th>Schlüsselwörter die bei der Suche berücksichtigt werden</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> <td></td> </tr> </tbody> </table> <ul style="list-style-type: none"> Für diese Schlagwörter werden keine Teilfragmente gebildet: <ul style="list-style-type: none"> Artikelnummer: 97550005 (1,69€) hat folgende Schlagwörter hinterlegt: "Beschriftungselement Löt-Schweißen-Werkstattbedarf-Arbeitsschutz Marke Turnus Jeton Bleistift Werkzeugmarke Kreide Beschriftungselement, Bleistift," 	E	F	G	Keywords_de-DE	[Optionales Feld]		Bildschirm, Monitor	S	Alphanumerisch, max. 1000 Zeichen	Bildschirm, Monitor	S	Schlüsselwörter zum Artikel. Schlüsselwörter sind für den Kunden nicht sichtbar, werden bei der Suche aber berücksichtigt.	OS-System, Server	L		Drucker, Printer	S	Sprachcode siehe Spalte "Titel".	Laptop	K			L	[Optional Field]			Alpha numeric, max. 1000 characters			Article keywords. Keywords are not visible for customer, but are included in search.			Language code see column "Title".	KEYWORD	String	Schlüsselwörter die bei der Suche berücksichtigt werden			
E	F	G																																							
Keywords_de-DE	[Optionales Feld]																																								
Bildschirm, Monitor	S	Alphanumerisch, max. 1000 Zeichen																																							
Bildschirm, Monitor	S	Schlüsselwörter zum Artikel. Schlüsselwörter sind für den Kunden nicht sichtbar, werden bei der Suche aber berücksichtigt.																																							
OS-System, Server	L																																								
Drucker, Printer	S	Sprachcode siehe Spalte "Titel".																																							
Laptop	K																																								
	L	[Optional Field]																																							
		Alpha numeric, max. 1000 characters																																							
		Article keywords. Keywords are not visible for customer, but are included in search.																																							
		Language code see column "Title".																																							
KEYWORD	String	Schlüsselwörter die bei der Suche berücksichtigt werden																																							

		Ergebnis: Die Suche nach "Werkstattbed" oder "Jeto" ergibt keine Suchergebnisse.
14	Synonyme	<ul style="list-style-type: none"> • Wer: Katalog Nutzer <p>Synonyme können vom Katalog Nutzer selbst gepflegt werden: Stammdatenverwaltung > Katalogverwaltung > Katalogsuche > Einstellungen Bspw. Spibo => Spiralbohrer.</p> 
15	Artikelmerkmale	<ul style="list-style-type: none"> • Durchsuchung möglich: nein • Berücksichtigung bei der After Search Filterung: nein
16	After Search/ Facetierung	<ul style="list-style-type: none"> • Filterung nach Facetten/Facetten-Paradigma <ul style="list-style-type: none"> ○ Einschränkung des Preises/Preisspanne ○ Einschränkung auf Kernsortiment ○ Einschränkung auf Kataloge ○ Einschränkung auf Artikelkategorie ○ Einschränkung auf Hersteller • Speicherung der Facetten ist möglich: <ul style="list-style-type: none"> ○ Organisationseinstellungen / Reiter „Katalog“: Facettenauswahl bei einer neuen Suche merken.
17	After Search/ Sortierung	<ul style="list-style-type: none"> • Relevanz (Voreingestellt/Standard) <ul style="list-style-type: none"> ○ Gewichtung nach Feld, genauer Schreibweise und Teilbegriff. • Preis <ul style="list-style-type: none"> ○ aufsteigend ○ absteigend • Beliebtheit <ul style="list-style-type: none"> ○ Wie häufig war dieser Artikel in einer Bestellung vorhanden. Die Bestellmenge innerhalb einer Bestellung wird dabei berücksichtigt.

4. Gewichtungstabelle

Die Gewichtungstabelle zeigt auf welche Gewichtung für Suchtreffer in einem Feld in der Onventis Katalogsuche definiert ist. Je höher der Wert desto höher ist die Relevanz des Suchtreffers und somit dessen Rang im Suchergebnis. Für eine Suchanfrage werden die Gewichtungen aus allen relevanten Feldern summiert.

Wert	Gewichtung	Feld im Onventis Katalog aus Anwendersicht
article_title	3	Artikelname
article_title_ngram_front	1	Artikelname, Textfragment vorne
article_title_ngram_back	1	Artikelname, Textfragment hinten
article_description	1	Artikelbeschreibung
article_description_ngram_front	1	Artikelbeschreibung, Textfragment vorne
article_description_ngram_back	1	Artikelbeschreibung, Textfragment hinten
article_manufacturer_nr	10	Hersteller Artikel Artikel-Nr.
article_manufacturer_name	2	Herstellername
article_manufacturer_name_idx_ngram_front	2	Herstellername ID, Textfragment vorne
article_manufacturer_name_idx_ngram_back	2	Herstellername ID, Textfragment hinten
article_ean_nr	10	EAN Nummer
article_supplier_article_nr	10	Lieferanten Artikel-Nr.:
article_supplier_article_nr_ngram_front	10	Lieferanten Artikel-Nr., Textfragment vorne
article_supplier_article_nr_ngram_back	10	Lieferanten Artikel-Nr., Textfragment hinten
article_category_title	1	Kategorie Titel
article_keywords	1	Keywords
catalog_facet_de-de	2	Facetten
article_nr_extern	10	Externe Artikelnummer
article_internal_article_nr	10	Interne Artikelnummer
article_internal_article_nr_ngram_front	10	Interne Artikelnummer, Textfragment vorne
article_internal_article_nr_ngram_back	10	Interne Artikelnummer, Textfragment hinten