Berufsakademie Sachsen

Staatliche Studienakademie Leipzig

**Analyse des Einsatzes einer Volltextsuche für die Sage HR Suite am Beispiel von Apache Solr**

Bachelorarbeit

zur Erlangung des akademischen Grades eines

„Bachelor of Science“

in der Studienrichtung Informatik

Eingereicht von: Angela Stöckert

Herrmann-Meyer-Straße 11, 04207 Leipzig

Seminargruppe: CS14-1 /CS15-1

Matrikelnr.: 5000559

Betreuer: M. o. Science Andre Kierzkowski

Arvato Systems Perdata GmbH

Martin-Luther-Ring 7-9

04109 Leipzig

Leipzig, 29.06.2018

**Inhaltsverzeichnis**

[1 Einleitung 3](#_Toc506303989)

[1.1 Motivation 3](#_Toc506303990)

[1.2 Unternehmen 3](#_Toc506303991)

[1.3 Anforderungen an einen Prototyp 3](#_Toc506303992)

[1.4 Thesen 3](#_Toc506303993)

[2 Theoretische Vorbetrachtung 4](#_Toc506303994)

[3 Vorüberlegungen zur Umsetzung in 5](#_Toc506303995)

[4 Implementierung 6](#_Toc506303996)

[4.1 Technische Infrastruktur 6](#_Toc506303997)

[4.2 Auswahl und Darstellung der Suchergebnisse 6](#_Toc506303998)

[4.3 Navigation vom Suchergebnis zur entsprechenden Ansicht 6](#_Toc506303999)

[5 Diskussion der Thesen 7](#_Toc506304000)

[6 Zusammenfassung, Fazit und Ausblick 8](#_Toc506304001)

[6.1 Zusammenfassung 8](#_Toc506304002)

[6.2 Fazit 8](#_Toc506304003)

[6.3 Ausblick 8](#_Toc506304004)

[7 Abkürzungsverzeichnis 9](#_Toc506304005)

[8 Abbildungsverzeichnis 10](#_Toc506304006)

[9 Tabellenverzeichnis 11](#_Toc506304007)

[10 Literaturverzeichnis 12](#_Toc506304008)

[11 Anhang 16](#_Toc506304009)

[11.1 Auszug aus Solr-Schema-Datei für Prototyp 16](#_Toc506304010)

[11.1.1 Beispiel für Kalenderübersichtmodul 16](#_Toc506304011)

[Selbstständigkeitserklärung 17](#_Toc506304012)

# Einleitung

## Motivation

## Unternehmen

## Anforderungen an einen Prototyp

## Thesen

Folgende Thesen gilt es in der Analyse näher zu untersuchen.

# Theoretische Vorbetrachtung

# Vorüberlegungen zur Umsetzung in

# Implementierung

## Technische Infrastruktur

## Auswahl und Darstellung der Suchergebnisse

## Navigation vom Suchergebnis zur entsprechenden Ansicht

# Diskussion der Thesen

# Zusammenfassung, Fazit und Ausblick

## Zusammenfassung

## Fazit

## Ausblick

# Abkürzungsverzeichnis

**DSD** *Application Programming Interface*

**CSS** *Cascading Style Sheets*

**DIH** *Data Import Handler*

**DLL** *Dynamic Link Library*

**DTO** *Data Transfer Object*

**GUID** *Globally Unique Identifier*

**HTTP** *Hypertext Transfer Protocol*

**HTML** *Hypertext Markup Language*

**IIS** *Internet Information Services*

**IR** *Information Retrieval*

**JDBC** *Java Database Connectivity*

**JSON** *JavaScript Object Notation*

**MVC** *Model View Controller*

**MVVM** *Model View ViewModel*

**NoSQL** *Not only SQL*

**PDF** *Portable Document Format*

**PW** *Sage HR Suite*

**REST** *Representational State Transfer*

**SQL** *Standard Query Language*

**URL** *Uniform Resource Locator*

**XML** *Extensible Markup Language*

# Abbildungsverzeichnis

[Abbildung 1: Indexierungsprozess von Lucene [4 S. 34] 12](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755558)

[Abbildung 2: Indexierungsvorgang mit Zusammenhang zwischen Dokument, Feldern, Analysierer und Indexdatenbank 14](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755559)

[Abbildung 3: Beziehung zwischen den Hauptklassen von Lucene zum Durchsuchen des Indexes 15](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755560)

[Abbildung 4: Schematische Darstellung des Architekturmusters MVC (durchgezogene Linie: direkte Beziehung; gestrichelte Linie: Benachrichtigung über Ereignis [roter Blitz]) 23](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755561)

[Abbildung 5: Bildschirmausschnitt zum Aufruf der Resturlaubansicht in Kalenderübersicht des HR Portals. Die gesamte Ansicht zum Menüpunkt Übersicht ist eine SPA. 24](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755562)

[Abbildung 6: Schematische Darstellung des Architekturmusters MVVM (durchgezogene Linie: direkte Beziehung; gestrichelte Linie: Benachrichtigung über Ereignis [roter Blitz]) 25](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755563)

[Abbildung 7: Menü zu den Hauptmodulen des HR Portals 27](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755564)

[Abbildung 8: Code-Beispiel zum Umschalten zwischen zwei Views mit Hilfe des Kendo UI Frameworks 28](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755565)

[Abbildung 9: Ausschnitt aus der Detailansicht innerhalb des HR Portal-Moduls Zeiterfassung 32](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755566)

[Abbildung 10: C#-Quelltext zur Generierung eines möglichst einzigartigen Dokumentenschlüssels anhand einer gegebenen Zeichenkette 33](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755567)

[Abbildung 11: Startseite der Weboberfläche zum Solr-Server 37](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755568)

[Abbildung 12:UML-Verteilungsdiagramm zur Client-Server-Architektur des Produktivsystems HR Portal mit Volltextsuche 37](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755569)

[Abbildung 13: Eintrag des DIH als RequestHandler in die Hauptkonfigurationsdatei von Solr 41](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755570)

[Abbildung 14: Definition der Datenquellen innerhalb der DIH-Konfigurationsdatei 42](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755571)

[Abbildung 15: Definition eines Dokuments in DIH-Konfigurationsdatei für Suche der Resturlaubsansicht 43](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755572)

[Abbildung 16: DTO-Klasse für Datenobjekt zur Indexierung 45](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755573)

[Abbildung 17: WebAPI-Controller-Methode zum Abfragen von Daten zur Indexierung 45](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755574)

[Abbildung 18: Methode aus der *FullTextSearchService*-Klasse zur Beschaffung der angeforderten Daten für die Indexierung 46](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755575)

[Abbildung 19: Benutzeroberfläche der prototypsartigen Sucheingabe im HR Portal 47](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755576)

[Abbildung 20: JavaScript-Funktion zum Behandeln von Mausklickereignissen nach Betätigung der Suchschaltfläche 47](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755577)

[Abbildung 21: URL für Suchanfrage an den Solr-Server 48](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755578)

[Abbildung 22: Ausschnitt aus Ansicht der Ergebnisse der Volltextsuche 49](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755579)

[Abbildung 23: Beispiel für Zuordnung von allen anzuzeigenden Views pro View-Schlüssel 51](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755580)

[Abbildung 24: Funktion eines Hilfsobjekts zum Navigieren zur Resturlaubsübersichts-View 52](file:///H:\BA%20Leipzig\Bachelorarbeit\Bachelorarbeit.docx#_Toc491755581)

# Tabellenverzeichnis

[Tabelle 1: Zu definierende Eigenschaften einer Solr-Schema-Datei 13](#_Toc491755582)

# Literaturverzeichnis

[1]: **Chris, A.** Top 10 Search Engines In The World. [Online] 2017. https://www.reliablesoft.net/top-10-search-engines-in-the-world/.

[2]: **Kuc, R.** Top 15 Solr vs. Elasticsearch Differences. [Online] Juni 19, 2017. [Cited: August 5, 2017.] https://sematext.com/blog/2017/06/19/solr-vs-elasticsearch-differences/.

[3]: **Yigal, A.** Solr vs. Elasticsearch: Who’s The Leading Open Source Search Engine? [Online] Sept. 2016. https://logz.io/blog/solr-vs-elasticsearch/.

[4]: **McCandless, M., Hatcher, E. und Gospodnetic, O.** *Lucene in Action.* Greenwich : Manning Publications Co, 2010.

[5]: **Lucidworks.** Full Text Search Engines vs. DBMS. [Online] Sept. 2009. https://lucidworks.com/2009/09/02/full-text-search-engines-vs-dbms/.

[6]: **GmbH, Sage.** Das Unternehmen. [Online] 2017. http://www.sage.de/ueber-uns/ueber-sage/unternehmensprofil.

[7]: **Grainger, T. und Potter, T.** *Solr in Action.* New York : Manning Publications Co., 2014.

[8]: **Edlich, S. et al.** *NoSQL - Einstieg in die Welt nichtrelationaler Web 2.0 Datenbanken.* Münschen : Carl Hanser Verlag, 2011.

[9]: **Solr, Apache.** Apache Solr Reference Guide. [Online] Juni 2017. https://lucene.apache.org/solr/guide/6\_6/.

[10]: **Goll, J. and Dausmann, M.** *Architektur- und Entwurfsmuster der Softwaretechnik: Mit lauffähigen Beispielen in Java.* Wiesbaden : Springer Vieweg, 2013.

[11]: **Fettke, P.** Client-Server-Architektur. [Online] Sept. 2016. http://www.enzyklopaedie-der-wirtschaftsinformatik.de/lexikon/is-management/Systementwicklung/Softwarearchitektur/Architekturparadigmen/Client-Server-Architektur.

[12]: **Microsoft.** The MVVM Pattern. [Online] Feb. 2012. https://msdn.microsoft.com/en-us/library/hh848246.aspx.

[13]: **Fowler, M.** GUI Architectures. [Online] Juli 2006. https://martinfowler.com/eaaDev/uiArchs.html.

[14]: **Wikipedia.** Apache Solr - Integrating Solr. [Online] Aug. 2017. https://en.wikipedia.org/wiki/Apache\_Solr#Integrating\_Solr.

[15]: **Wikipedia**. Hasfunktion. [Online] Juni 2017. https://de.wikipedia.org/wiki/Hashfunktion.

[16]: **Sieben, J.** Informatik Aktuell. [Online] Oktober 20, 2015. [Cited: Juli 4, 2017.] https://www.informatik-aktuell.de/entwicklung/programmiersprachen/to-trigger-or-not-to-trigger.html.

[17]: **Miller, B.** How to Download and Compile Solr 6 in Eclipse. [Online] Mai 2016. http://coding-art.blogspot.de/2016/05/how-to-download-and-compile-solr-6-in.html.

[18]: **Fowler, M.** Data Transfer Object. [Online] https://martinfowler.com/eaaCatalog/dataTransferObject.html.

[19]: **Osmani, A.** Learning Javascript Design Patterns. [Online] 2017. https://addyosmani.com/resources/essentialjsdesignpatterns/book/.

[20]: **Zala, A.** An Introduction To The HTML5 History API. [Online] Okt. 2014. https://code.tutsplus.com/tutorials/an-introduction-to-the-html5-history-api--cms-22160.

[21]: **Schiering, M.** SEO und SEA im Kreuzfeuer von AJAX und Single-Page-Applications. [Online] Okt. 2015. https://www.mso-digital.de/seo-und-sea-im-kreuzfeuer-von-ajax-und-single-page-applications/.

[22]: **Microsoft.** Übersicht über ASP.NET MVC. [Online] 2017. https://msdn.microsoft.com/de-de/library/dd381412(v=vs.108).aspx.

[23]: **Solr, Apache.** Solr. [Online] 2017. http://lucene.apache.org/solr/.

[24]: **Wikipedia.** Unstrukturierte Daten. [Online] https://de.wikipedia.org/wiki/Unstrukturierte\_Daten.

[25]: **W3C.** Lokalisierung vs. Internationalisierung. [Online] Sept. 2010. https://www.w3.org/International/questions/qa-i18n.de.php.

[26]: **Wikipedia.** Webanwendung. [Online] Juli 2017. https://de.wikipedia.org/wiki/Webanwendung.

[27]: **Wikipedia**. Subskription - Software. [Online] Juni 2017. https://de.wikipedia.org/wiki/Subskription#Software.

[28]: **Wikipedia.** Softwarearchitektur. [Online] Mai 2017. https://de.wikipedia.org/wiki/Softwarearchitektur.

[29]: **Wikipedia**. Programmierschnittstelle. [Online] Mai 2017. https://de.wikipedia.org/wiki/Programmierschnittstelle.

[30]: **Wikipedia**. Anwendungsserver. [Online] Aug 2017. https://de.wikipedia.org/wiki/Anwendungsserver.

[31]: **Wikipedia**. Hypertext Transfer Protocol. [Online] Juli 2017. https://de.wikipedia.org/wiki/Hypertext\_Transfer\_Protocol.

[32]: **Wirtschaftslexikon, Gabler.** URL. [Online] 2017. http://wirtschaftslexikon.gabler.de/Archiv/55200/url-v8.html.

[33]: **Wikipedia.** Framework. [Online] März 2017. https://de.wikipedia.org/wiki/Framework.

[34]: **selfhtml.** JavaScript/ Anwendung und Praxis/ Webanwendung. [Online] Feb. 2017. https://wiki.selfhtml.org/wiki/JavaScript/Anwendung\_und\_Praxis/Webanwendung.

[35]: **selfhtml**. JavaScript/ Ajax. [Online] Juli 2017. https://wiki.selfhtml.org/wiki/JavaScript/Ajax.

[36]: **Wikipedia.** Message-Digest Algorithm 5. [Online] Aug. 2017. https://de.wikipedia.org/wiki/Message-Digest\_Algorithm\_5.

[37]: **Wirtschaftslexikon, Gabler.** Primärschlüssel. [Online] 2017. http://m.wirtschaftslexikon.gabler.de/Definition/schluessel.html?referenceKeywordName=Prim%C3%A4rschl%C3%BCssel.

[38]: **Microsoft.** Exemplarische Vorgehensweise: Erstellen einer Website mit Razor-Syntax in Visual Studio. [Online] 2017. https://msdn.microsoft.com/de-de/library/gg606533(v=vs.100).aspx.

[39]: **Wirtschaftslexikon, Gabler.** Datenbankmanagementsystem (DBMS). [Online] http://wirtschaftslexikon.gabler.de/Archiv/74907/datenbankmanagementsystem-dbms-v9.html.

[40]: **Microsoft.** Was ist eine DLL? [Online] Mai 2017. https://support.microsoft.com/de-de/help/815065/what-is-a-dll.

[41]: **Kecher, C. and Salvanos, A.** *UML 2.5. Das umfassende Handbuch.* Bonn : Rheinwerk Verlag GmbH, 2015.

[42]: **Oettinger, R.** Was ist eine Softwarelizenz. [Online] Juni 2014. https://www.computerwoche.de/a/was-ist-eine-softwarelizenz,1913465.

[43]: **Wirtschaftslexikon, Gabler.** Open Source. [Online] http://wirtschaftslexikon.gabler.de/Archiv/77360/open-source-v8.html.

[44]: **Wikipedia.** Microsoft Access. [Online] März 2017. https://de.wikipedia.org/wiki/Microsoft\_Access.

[45]: **Wikipedia**. Globally Unique Identifier. [Online] Juni 2017. https://de.wikipedia.org/wiki/Globally\_Unique\_Identifier.

[46]: **Schwichtenberg, Dr. H.** Erklärung des Begriffs: Internet Information Server (IIS). [Online] Dez. 2015. http://www.it-visions.de/glossar/alle/109/Der\_Internet\_Information\_Server\_IIS\_und\_ASPNET.aspx.

[47]: **Tilkov, S. et al.** *REST und HTTP - Entwicklung und Integration nach dem Architekturstil des Web.* Heidelberg : dpunkt.verlag, 2015.

[48]: **Ullenboom, C.** *Java ist auch eine Insel - Das umfassende Handbuch.* Bonn : Rheinwerk Verlag, 2010.

[49]: **Wikipedia.** Anwendungssoftware. [Online] März 2017. https://de.wikipedia.org/wiki/Anwendungssoftware.

[50]: **Wikipedia**. PL/SQL. [Online] Mai 2017. https://de.wikipedia.org/wiki/PL/SQL.

# Anhang

## Auszug aus Solr-Schema-Datei für Prototyp

### Beispiel für Kalenderübersichtmodul

# Selbstständigkeitserklärung

Ich versichere, dass ich die vorliegende Arbeit ohne fremde Hilfe selbstständig verfasst und nur die angegebenen Quellen und Hilfsmittel benutzt habe. Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen sind unter Angabe der Quellen kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form weder veröffentlicht, noch einer anderen Prüfungsbehörde vorgelegt.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Ort, Datum Unterschrift