

## Analyse | Beratung | Schulung

*Andreas Wygrabek - Partner für Data Science, Big-Data,  
Datenanalyse, Python und R.*



Ich biete ich Full-Stack Projektmitarbeit im Bereich Data Science an: Codeentwicklung, Support, Training und Projektmanagement. Die zentralen Werkzeuge in meinen Projekten und Trainings sind die 2 populärsten freien Programmiersprachen zur statistischen Datenanalyse – R und Python.

Technisches Profil, Stand: 09. Juni 2019

Verfügbar unter: <https://www.data-science-architect.de/ueber-mich/>

### Inhalt

Übersicht.....	2
Technisches Profil .....	2
Onlinereferenzen .....	3
Projektauszug.....	3
Angebotsübersicht: Trainings und Workshops.....	5
Karriere .....	5
Talks .....	6

## Übersicht

- + > 9 Jahre Erfahrung in der Programmierung mit R
- + > 6 Jahre Erfahrung in der Programmierung mit Python
- + Full-Stack Projektarbeit: Datenmanagement, ETL, Analyse, Prognose, Interpretation und Integration von Data Science in Business-Prozesse über Web-Applikationen.
- + Expertise in Machine Learning, Multivariater Statistik, Zeitreihenanalyse und den State-of-the-Art Frameworks.
- + Erfahrener Trainer und Berater im Bereich Programmierung und Datenanalyse.
- + Projektleitung in zahlreichen Data-Science Projekten.
- + Branchenübergreifendes Angebot: Finance, Energie, akademische Einrichtungen, Forschung, Handel und Andere.
- + Sprachen: Deutsch und Englisch

## Technisches Profil

### Programmieren

R +++  
Python +++  
JavaScript ++  
Web-Entwicklung: html/CSS ++  
SQL ++

### Methoden

Machine Learning / Data Mining +++  
Deep Learning +++  
Multivariate Statistik +++  
Zeitreihenanalyse +++  
Inferenzstatistik +++

### Big-Data Technologien

Spark +++  
Hadoop ++  
Hive ++  
MongoDB ++  
Cassandra ++

### Frameworks/Libraries

mlr / caret +++  
tensorflow/ Keras +++  
Scikit-Learn +++  
MLlib +++  
shiny / htmlwidgets +++

### Weitere Technologien

Version Control (Git, SVN) +++  
Atlassian Jira +++  
SPSS +++  
Grafana ++  
SAS +

## Onlinereferenzen

Website	<a href="https://www.data-science-architect.de">https://www.data-science-architect.de</a>
Blog	<a href="https://www.data-science-architect.de/blog">https://www.data-science-architect.de/blog</a>
Xing	<a href="https://www.xing.com/profile/Andreas_Wygrabek/">https://www.xing.com/profile/Andreas_Wygrabek/</a>
Linkedin	<a href="https://www.linkedin.com/in/andreas-wygrabek-7b805a154/">https://www.linkedin.com/in/andreas-wygrabek-7b805a154/</a>
Github	<a href="https://github.com/dascar">https://github.com/dascar</a>

## Projektauszug

**Skizze:** Prognose- und Risikomodellierung für mögliche Kreditausfälle bei einer Bankgesellschaft.

Position: Algorithmus-Entwickler  
Technologien: R, caret, Textmining, Ensemble Modelling, AWS

**Skizze:** Hochrechnung und Prognose der Energieeinspeisung in das Elektrizitätsnetz durch erneuerbare Energieerzeugeranlagen.

Position: Algorithmus-Entwickler  
Technologien: Python, Scikit-Learn, Keras, Deep Learning, Spark, Hadoop, MongoDB

**Skizze:** Erstellen eines R-Pakets mit übersetzten Befehlen der SPSS-Syntax.

Position: Projektleitung, Entwickler  
Technologien: R, SPSS, Roxygen2, Devtools, Git, Jira

**Skizze:** Erstellen einer Shiny Applikation – Die Applikation dient dem Qualitätsmanagement eines Medizinprodukteherstellers als führendes BI-Tool.

Position: Shiny-Entwickler  
Technologien: R, Shiny, htmlwidgets, Git, Jira

**Skizze:** Proof of Concept für Preisprognosen im Energiesektor durch die automatisierte NLP-Verarbeitung und DataMining-Analyse von Meldungen einer Nachrichtenagentur.

Position: Algorithmus-Entwickler  
Technologien: R, tm, NLP, Machine Learning

**Skizze:** Integration von R in Business-Umgebungen.

Position: Projektleitung, System-Architekt, System-Administrator  
Technologien: RStudio-Server, Git, Jira

**Skizze:** Implementierung von ETL-Prozessen bei einem Marktforschungsunternehmen - Automatisierung von Workflows, Datencleaning, automatisierte Plausibilisierung der Daten während der Aufbereitung, Anlegen der Datenbankschemata auf einem Microsoft SQL-Server, importieren der Daten in die Datenbank, implementieren von Datenintegritätschecks.

Position: Entwickler  
Technologien: RStudio-Server, Git, Jira

**Skizze:** Erstellen einer Web-Appliation mit Leaflet. Die Applikation visualisiert die globale Sonneneinstrahlung sowie die Energieerzeugung von erneuerbaren Energieanlagen.

Position: Projektleitung, Web-Entwickler  
Technologien: Leaflet, html, Javascript, Python

**Skizze:** Werbewirkungsanalyse unter Berücksichtigung zeitverzögerter Effekte anhand von GRP-Werten und Kunden- bzw. deren Kosumverhalten im Internet.

Position: Projektleitung, Datenanalyst  
Technologien: R, mlr, Distributed Lag Models

## Angebotsübersicht: Trainings und Workshops

Die hier aufgeführten Veranstaltungen bilden mein derzeitiges Trainingsangebot. Weitere Informationen finden Sie auf meiner Website unter: <https://www.data-science-architect.de/leistungen/trainings-schulungen/>.

Titel	Sprache	Reguläre Dauer
Einführung in die Programmiersprache	R / Python	2 Tage R / 3 Tage Python
Fortgeschrittene Programmierung	R / Python	2 Tage
Machine Learning	R / Python	2 Tage
Deep Learning	R / Python	2 Tage
Reinforcement Learning (mit Partnern)	Python	2 Tage
Multivariate Statistik	R / Python	2 Tage
Zeitreihenanalyse	R / Python	2 Tage
Python in Operativumgebungen	Python	2 Tage
Hadoop und Spark für Python-Entwickler	Python	1 Tag
Webapps mit Shiny und htmlwidgets	R	1 Tag
Versuchsplanung und statistische Testverfahren	R	1 Tag

## Karriere

**1/2017 – 12/2018**

<b>Role:</b>	Data Scientist
<b>Unternehmen:</b>	Fraunhofer Gesellschaft
<b>Tätigkeiten:</b>	Software Development, Product Development, Statistical Programming
<b>Technologische Schwerpunkte:</b>	Python, JavaScript, HTML5, CSS3, Leaflet, jQuery, Hadoop, Spark, NoSQL, MongoDB, JSON, SVN, Scrum, Grafana, InfluxDB, Oracle, Pycharm

1/2011 – 12/2016

**Role:** Data Scientist  
**Unternehmen:** eoda GmbH  
**Tätigkeiten:** Statistical Programming, Software Development, Project Management and Lecturer / Trainer  
**Technologische Schwerpunkte:** R, JavaScript, HTML5, CSS3, MySQL, XML, JSON, Git, Jira, Scrum, Microsoft-SQL-Server, Hadoop, Spark, Machine Learning, Statistics

## Talks

Datum	Titel	Konferenz
9. November 2017	Hadoop und Spark in Data-Science-Anwendungen	[R]-Kenntnistage 2017
12. Oktober 2016	Application Field of R in classical industrial analytics	rRum – European R users meeting

Kassel, 2. Juni 2019



Andreas Wygrabek