

# State-of-the-Art und vielversprechende Studienideen zur (datengetriebenen) Ableitung von Future Skills

Arbeitspaket 02

09. Dezember 2022

Prof Dr. Mathias Klier | Julia Brasse | Dr. Maximilian Förster | Lars Moestue

Arbeitspaket 02

# Überblick – Bestehende Studien und (methodische) Ansätze

# Forschungsrichtungen im Bereich Future Skills

Ansätze zur Identifizierung zukünftiger Fähigkeiten



**Literature  
Review**



**Experten  
Interviews**



**Datengestützte  
Ansätze**



**Mixed  
Methods**

# Forschungsrichtungen im Bereich Future Skills

## Vertiefung | Ansätze zur Identifizierung zukünftiger Fähigkeiten

### Literature Review

Eine Möglichkeit, Future Skills zu ermitteln, besteht in der **Durchsicht der einschlägigen Fachliteratur**. Dies erfolgt auf der Grundlage etablierter Ansätze für Literature Reviews. Auf diese Weise kann eine **Vielzahl unterschiedlicher Perspektiven** auf Future Skills **erfasst, kategorisiert und evaluiert** werden. Dies führt zu eindeutigen Future Skills, da die Forscher sie aus der bisherigen Forschung ableiten.

Wir unterscheiden **zwei Arten** von Literature Reviews:



Literature Reviews, die Studien und frühere Forschungen zu bestimmten Trends (bspw. Industrie 4.0 oder Digitalisierung) und deren **Auswirkungen auf die Beschäftigung und die Arbeitskräfte in allen Branchen** zusammenfassen, um daraus abzuleiten, welche Skills zur Bewältigung dieser Änderungen erforderlich sind.



Literature Reviews, welche einen **spezifischen und vertieften Einblick** zu Future Skills **in einer einzelnen Branche** (bspw. Ingenieurwesen oder maritime Industrie) geben, indem sie sowohl Studien zu Future Skills für alle Branchen als auch Trends für die jeweilige Branche untersuchen.

**Aber:** Literature Reviews sind immer **vergangenheitsorientiert**, da sie nur bereits veröffentlichte Arbeiten überprüfen. Dadurch könnten sie **künftige Fähigkeiten übersehen**, die noch nicht etabliert sind, insbesondere wenn sie aus neuen Trends entstehen, die in der bisherigen Literatur noch nicht behandelt wurden.

(Cicek et al. 2019; Kotsiou et al. 2022; Kurtzo et al. 2016; Lieu et al. 2018; Ra et al. 2019; Rahmat et al. 2020; vom Brocke et al. 2015)

# Forschungsrichtungen im Bereich Future Skills

## Vertiefung | Ansätze zur Identifizierung zukünftiger Fähigkeiten

### Experten Interviews

Mit dem Ziel, Future Skills zu identifizieren, die sich aus aufkommenden Trends ergeben, kann die Methode der **Expertenbefragung** als **qualitative Forschungsmethode** verwendet werden. **Experten** sind sich in der Regel sowohl **etablierter Trends** als auch **aufkommender neuer Entwicklungen** und damit verbundener **zukünftiger Kompetenzen** bewusst. Um ein umfassendes Verständnis der Trends und ihrer Folgen für künftige Kompetenzen zu erlangen, sollten sowohl der **Werdegang** (bspw. Forscher, Dozenten, Schulungsleiter, private oder öffentliche Manager) als auch das **Dienstalter** der befragten Experten variieren.

Wir unterscheiden **drei** separate **Methoden**, wobei eine Kombination ebenfalls möglich ist:



**Einzelinterviews mit mehreren Experten;** ein Vorteil dieser Einzelinterviews ist, dass sie **vertiefende Befragungen** und **Nachfragen** ermöglichen. Sie sind jedoch **zeitaufwändig**, entsprechend ist die **Zahl der Experten** in der Regel eher **klein**.



**Online-Umfragen;** bei dieser Methode ist die Befragung einer **größeren Anzahl von Personen** möglich, jedoch sind **Rückfragen nicht** möglich. Zudem werden die Ergebnisse **aggregiert**, was zu einer **Verzerrung** hin zu extremen Antworten führen kann.



**Fokusgruppen;** hier **diskutieren** Experten über künftige Kompetenzen, um zu einem entsprechenden Konsens zu gelangen. Um eine Voreingenommenheit gegenüber rhetorisch geschickten Experten oder Experten mit höherem Status zu vermeiden, betonen die Forscher die Bedeutung einer **sorgfältigen Gestaltung und Moderation der Diskussion**, um Situationen des Gruppenzwangs in Fokusgruppen zu vermeiden.

(Aichholzer 2009; Bogner et al. 2009; Evans and Mathur 2005; Kirchherr et al. 2018; Kurtzo et al. 2016; Leopold et al. 2016; Tremblay et al. 2010; Wentling and Palma-Rivas 1998; Zahidi et al. 2020)

# Forschungsrichtungen im Bereich Future Skills

## Vertiefung | Ansätze zur Identifizierung zukünftiger Fähigkeiten

### Daten- gestützte Ansätze

Angetrieben durch die **große Verfügbarkeit von Daten** und durch enorme **technologische Fortschritte** wie maschinelles Lernen, ermöglichen datengetriebene Ansätze die Identifizierung von Future Skills.



Als **zugrundeliegende Datenbasis** werden meist **Stellenanzeigen** verwendet. Diese bestehen in der Regel aus Informationen über das Unternehmen, über Vorteile der Stelle, über den Bewerbungsprozess und aus **detaillierten Informationen über die geforderten Fähigkeiten**, die von den Arbeitgebern als notwendig für die Zukunft angesehen werden. Da heutzutage fast alle Stellenanzeigen online veröffentlicht werden, können **Online-Stellenanzeigen** eine **repräsentative Stichprobe der vom Arbeitsmarkt in der Zukunft geforderten Qualifikationen** liefern.

Datengetriebene Ansätze zur Ermittlung von Future Skills bestehen in der Regel aus **zwei Schritten**:



In einem ersten Schritt werden alle Fähigkeiten aus den Stellenanzeigen mit verschiedenen Methoden **extrahiert**.



In einem zweiten Schritt werden die extrahierten Fähigkeiten mit verschiedenen **quantitativen Methoden** analysiert.

- **Empirische Analyse**; Ermittlung von künftigen Fähigkeiten für bestimmte Stellen. Die Methodik bietet einen guten Überblick über häufig nachgefragte Future Skills, könnte aber neu entstehende Zukunftskompetenzen übersehen.
- **Explorative Methoden**; Ermittlung von Fähigkeiten, die häufig zusammen in Stellenanzeigen erscheinen. Dieses Vorgehen hat Vorteile im Hinblick auf die Identifizierung von neu entstehenden Fähigkeiten. Beispiele: Topic Modeling oder Clustering-Algorithmen.

Datengestützte Ansätze können für ganze Regionen und Branchen repräsentative Future Skills ermitteln.

(Ang et al. 2013; Buchmann et al. 2022; Carnevale et al. 2016; Ganesan et al. 2018; Litecky et al. 2010; Maer-Matei et al. 2019; Rios et al. 2020; Todd et al. 1995)

# Forschungsrichtungen im Bereich Future Skills

## Vertiefung | Ansätze zur Identifizierung zukünftiger Fähigkeiten

A circular icon with a green border containing the text 'Mixed Methods' and a small graphic of a person at a computer.

### Mixed Methods

In der Literatur werden hauptsächlich **drei Ansätze** zur Ermittlung künftiger Kompetenzen verwendet:

- **Literature Reviews**
- **Experteninterviews**
- **Datengestützte Ansätze.**

Um die Schwächen der jeweiligen Ansätze zu überwinden und ihre Stärken zu vereinen, werden auch Kombinationen dieser Ansätze in der Forschung genutzt, sog. **Mixed Methods**.

Dabei wird z.B. ein **datengestützter Ansatz** mit **Experteninterviews und / oder vergleichbaren Methoden kombiniert**.

Diese Kombination von meist qualitativen und quantitativen Ansätzen innerhalb einer Untersuchung bietet Vorteile. So lassen diese Untersuchungen **aussagekräftige Schlüsse und vielseitige Einblicke in ein interessantes Phänomen** zu.

Eine Kombination von Methoden ermöglicht es zudem, **Schlussfolgerungen aus unterschiedlichen, manchmal gegensätzlichen Ergebnissen** der qualitativen und quantitativen Untersuchungen zu ziehen. Diese Kombination macht die abgeleiteten theoretischen Ergebnisse zum untersuchten Phänomen oft **robuster**.

(Creswell and Clark 2017; Reis et al. 2022)

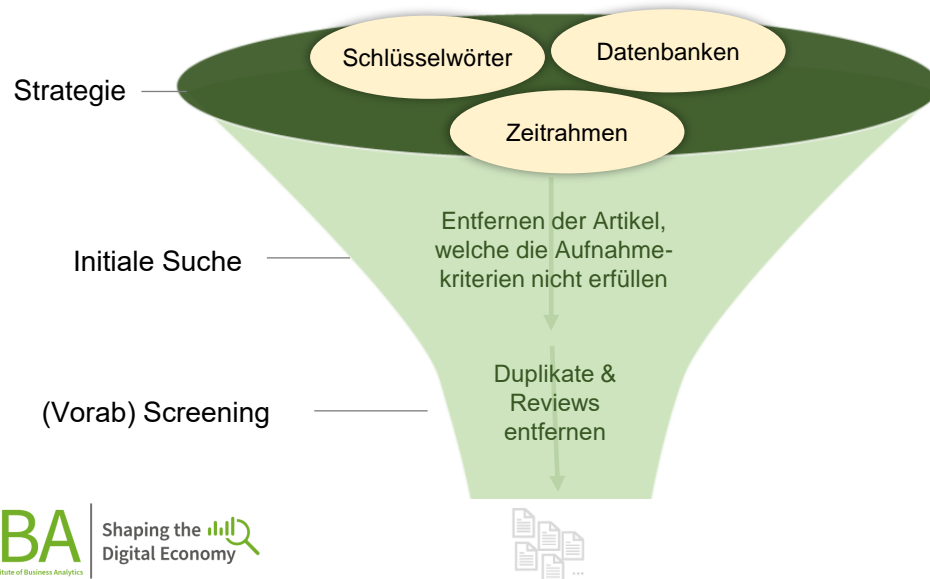
# Strukturierte Literaturrecherche

## Überblick

Ausgehend vom Überblick über bestehende Ansätze zur Identifizierung zukünftiger Fähigkeiten, werden in der nachfolgenden Untersuchung **datengestützte Ansätze und Mixed Methods auf Basis von Online-Stellenanzeigen** näher **analysiert**. Dafür wird eine **systematische und strukturierte Literaturrecherche in drei Schritten** durchgeführt:

### 1 Suchstrategie

Zur Vorbereitung ist es notwendig, eine umfassende und replizierbare Literatursuchstrategie anzuwenden, die relevante Datenbanken, geeignete Schlüsselwörter und einen angemessenen Zeitrahmen umfasst.



### 2 Labeling



Mehrere Forscher kodieren die Artikel nach einem festgelegten Schema.



Dimensionen u.a.: data-driven, mixed-methods.



Bei Unstimmigkeiten zwischen den Forschern wird ein Konsens auf der Grundlage von Diskussionen im Team erzielt.



### 3 Analyse



Wir analysieren Muster innerhalb der verschiedenen Artikel.



Wir leiten Schlussfolgerungen aus den für die Zukunft identifizierten Forschungslücken in den wichtigsten Papern ab.

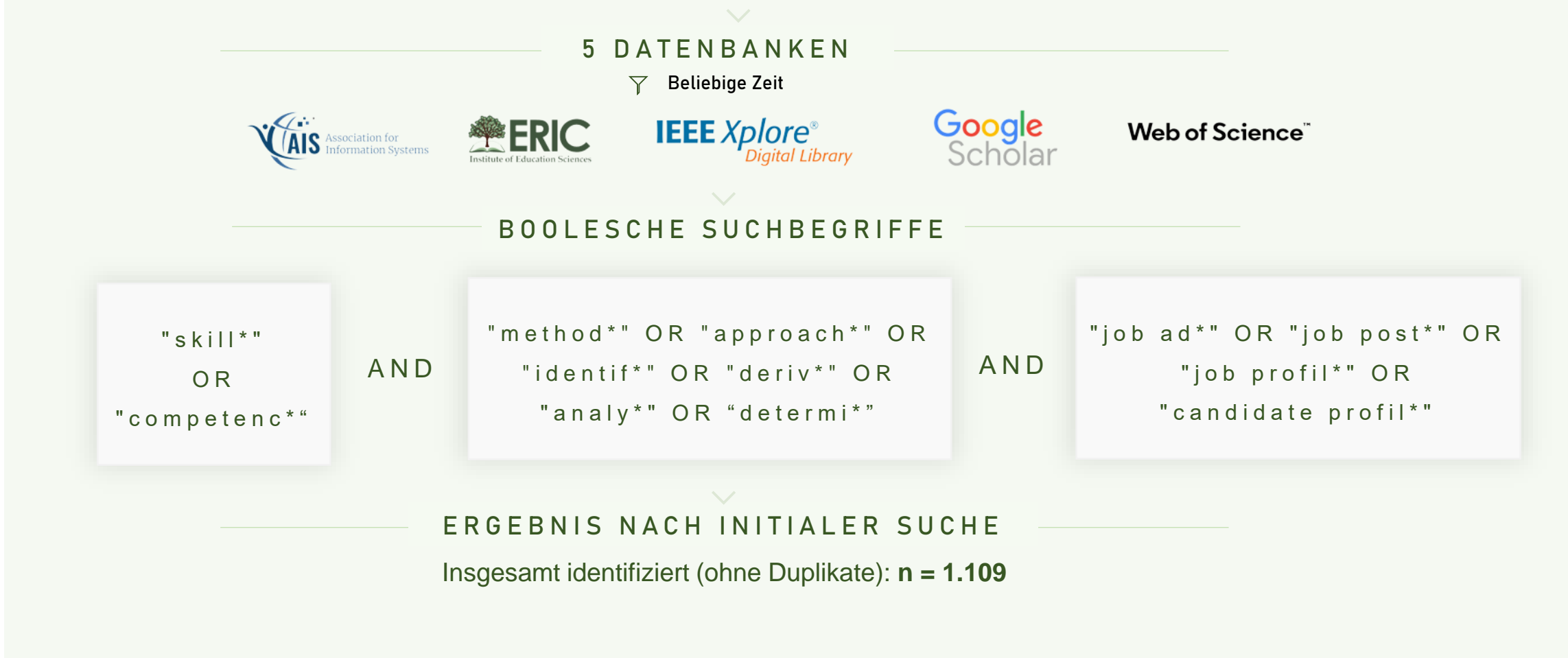


Ergebnisse fließen in die neuen Studienideen in AP 03 ein.



# Strukturierte Literaturrecherche

## 1 Suchstrategie | Rechercheergebnisse



# Strukturierte Literaturrecherche

## 1 Suchstrategie | Rechercheergebnisse

### SCREENING

Mehrere Forscher screenen die **1.109 Artikel** anhand **Titel** und **Abstract** nach **Relevanz** der Artikel.

#### 1. Schritt

Forscher bewerten Artikel hinsichtlich Relevanz (data-driven + job ads).

Alle Artikel werden mindestens 1x geratet.

50% der Artikel werden von zwei Forschern bewertet und dann verglichen.

Bei Unstimmigkeiten wird ein Konsens auf der Grundlage von Diskussionen im Team erzielt.

**Ergebnis: 301 Artikel verbleiben**

Mehrere Forscher analysieren **301 Artikel**, ob **automatisierte Methoden** verwendet werden.

#### 2. Schritt

Forscher analysieren, ob automatisierte Methoden verwendet werden.

Alle Artikel werden mindestens 1x analysiert.

50% der Artikel werden von zwei Forschern analysiert und dann verglichen.

Bei Unstimmigkeiten wird ein Konsens auf der Grundlage von Diskussionen im Team erzielt.

**Ergebnis: 136 Artikel verbleiben**

### ERGEBNIS NACH SCREENING

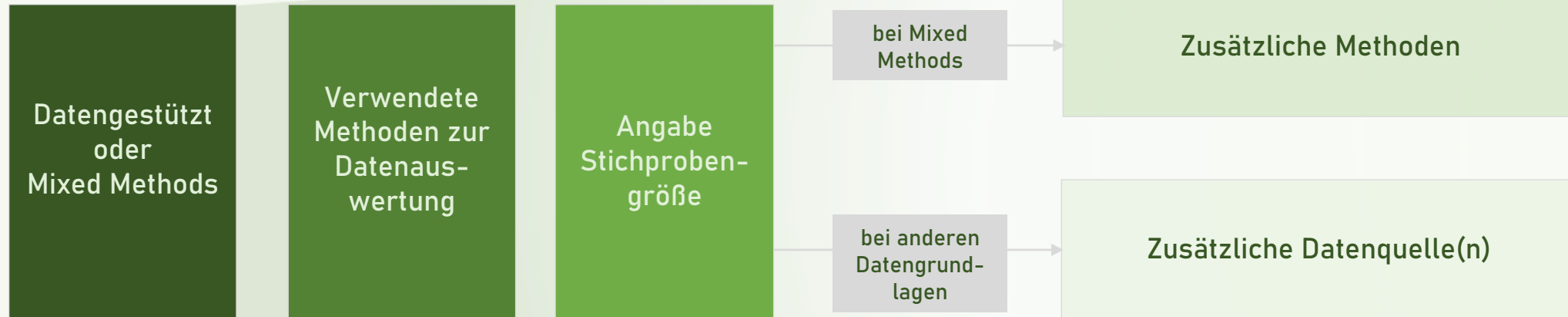
Insgesamt identifiziert: **n = 136<sup>1</sup>**

# Strukturierte Literaturrecherche

## 2 Labeling | Überblick Dimensionen

### RATING NACH METHODIK

Mehrere Forscher raten die Artikel anhand **5 fachlicher Kategorien**.  
Bei Unstimmigkeiten zwischen den Forschern wird ein Konsens auf der Grundlage von Diskussionen im Team erzielt.

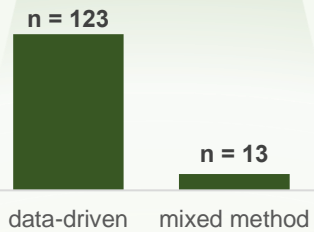


# Strukturierte Literaturrecherche

## 3 Analyse | Rechercheergebnisse

### ERGEBNISSE

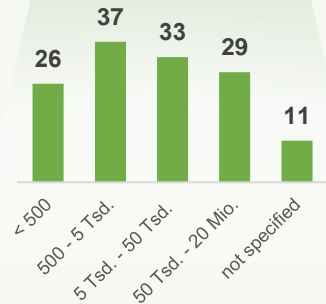
Datengestützt  
oder  
Mixed Methods



Verwendete  
Methoden zur  
Datenauswertung

Vertiefung  
auf Folie 15

Angabe  
Stichproben-  
größe



bei Mixed  
Methods  
n = 13

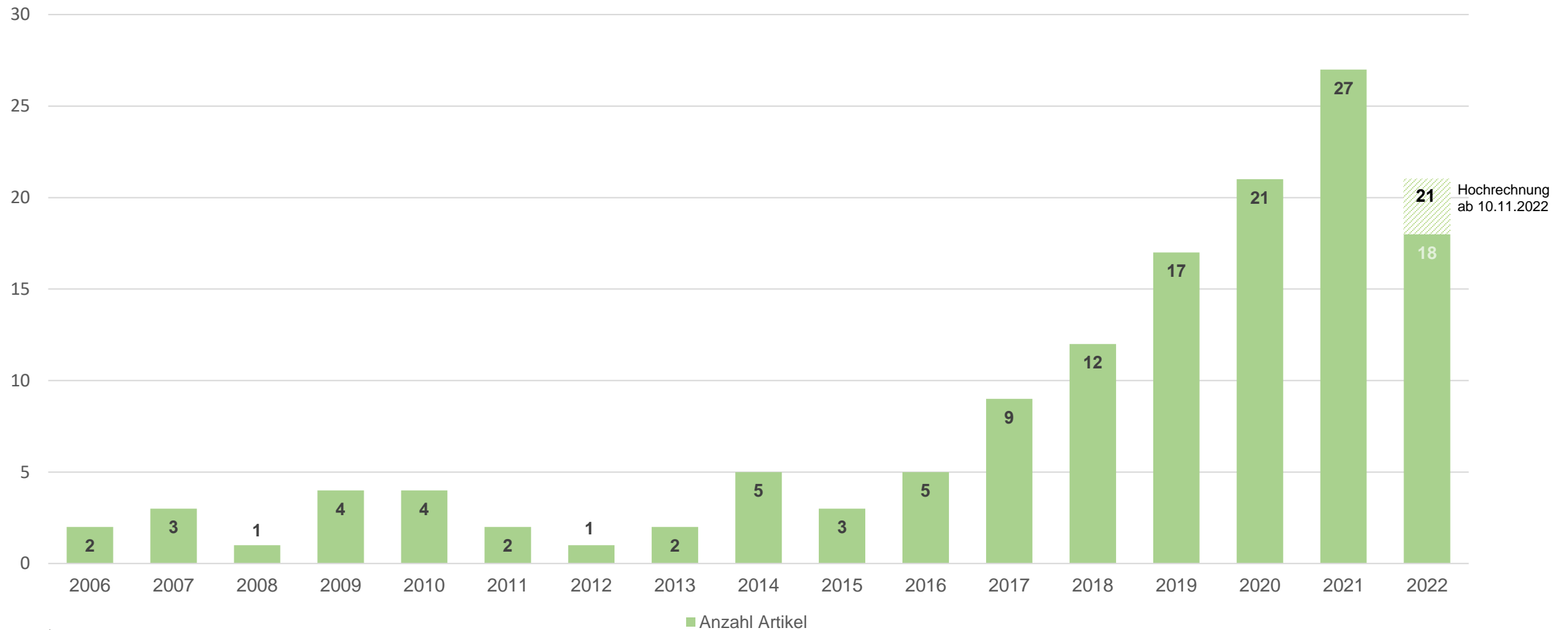
Zusätzliche Methoden:  
Experteninterview,  
Literature Review, Survey, ...

bei anderen  
Datengrund-  
lagen  
n = 16

Zusätzliche Datenquelle:  
Praktikumsanzeigen, Lebensläufe,  
Ausbildungsprogramme, LinkedIn-Profile,...

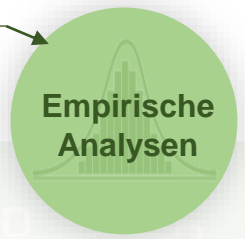
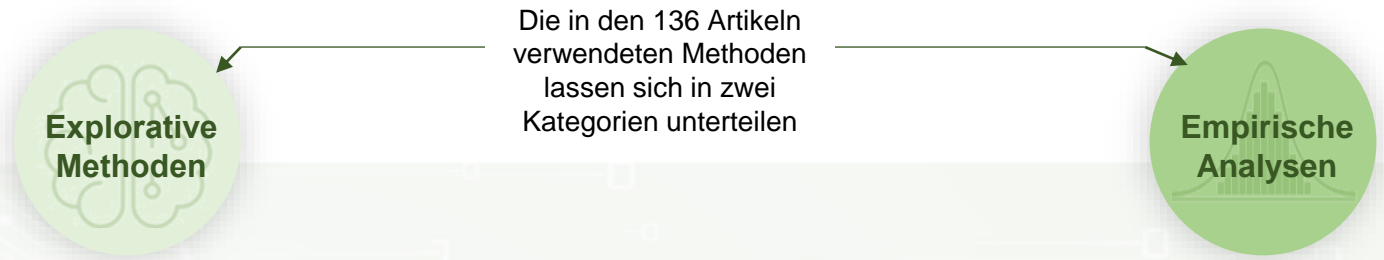
# Strukturierte Literaturrecherche

## 3 Analyse | Entwicklung im Zeitverlauf



# Strukturierte Literaturrecherche

## 3 Analyse | Kategorisierung verwendeter Methoden



**Ziel:** Erkennen von ähnlichen Skills, die häufig zusammen vorkommen

**Häufig verwendete Methoden:**

**Topic Modeling:** Methodik, um abstrakte Themen (Oberskills) in Texten (Stellenanzeigen) sichtbar zu machen. Jedes Thema ist dabei eine Verteilung einzelner Wörter, welche auch mehrfach in unterschiedlichen Themen vorkommen können. Häufig verwendete Algorithmen:

- Latent Semantic Analysis (LSA)
- Latent Dirichlet Allocation (LDA).

**Word Embeddings:** Methodik, um einzelne Worte (z.B. Skills) als niedrig-dimensionale reelle Vektoren darzustellen, sodass semantisch ähnliche Worte eine geringe Distanz besitzen. Aufwändig zu erstellen, da Millionen von Dokumenten benötigt werden.

**Clustering:** Methodik, um eine Menge von Objekten (Skills) so zu gruppieren, dass Objekte innerhalb einer Gruppe in einem gewissen Sinne ähnlicher zueinander sind, als Objekte in unterschiedlichen Gruppen (z.B. häufiges gemeinsames Vorkommen in Stellenanzeigen). Dabei kann jedes Objekt nur zu einer einzigen Gruppe gehören. Häufig verwendete Algorithmen:

- Hierarchisches Clustering
- k-Means.

**Ziel:** Analyse, welche Skills besonders häufig vorkommen

**Häufig verwendete Methoden:**

**Content Analysis:** Analyse, welche Wörter (z.B. Skills) am häufigsten vorkommen und somit am relevantesten sind, häufig auch in Kombination mit *Co-Word-Occurrence-Analysis* um zu analysieren, welche Wortpaare am häufigsten zusammen vorkommen.

**Lineare Regression:** Methodik, um zu modellieren, wie sich die Skillnachfrage über die nächsten Jahre entwickelt. Daraus wird abgeleitet, welche Skills besonders relevant werden. Kann auch bspw. dazu verwendet werden, um zu analysieren, inwiefern gewisse Skills Einfluss auf das Gehalt haben.

Methodiken aus diesen Kategorien lassen sich sowohl innerhalb der Kategorie als auch mit der jeweils anderen Kategorie kombinieren. Daneben gibt es noch ein paar wenige andere Methodiken, die sich nicht in diese Kategorien einordnen lassen (z.B. Social Network Analysis, Klassifikation & Recommender Systems).

(Almeida and Xexeo 2019; Case 2007; Omrana et al. 2007; Sykes 1993; Vayansky and Kumar 2020)

# Ausgewählte Artikel

## Steckbrief 01 / 11

### Demystifying Job Roles in Data Science: A Text Mining Approach (2021)

Sven Michalczyk; Mario Nadj; Alexander Maedche; and Christoph Gröger

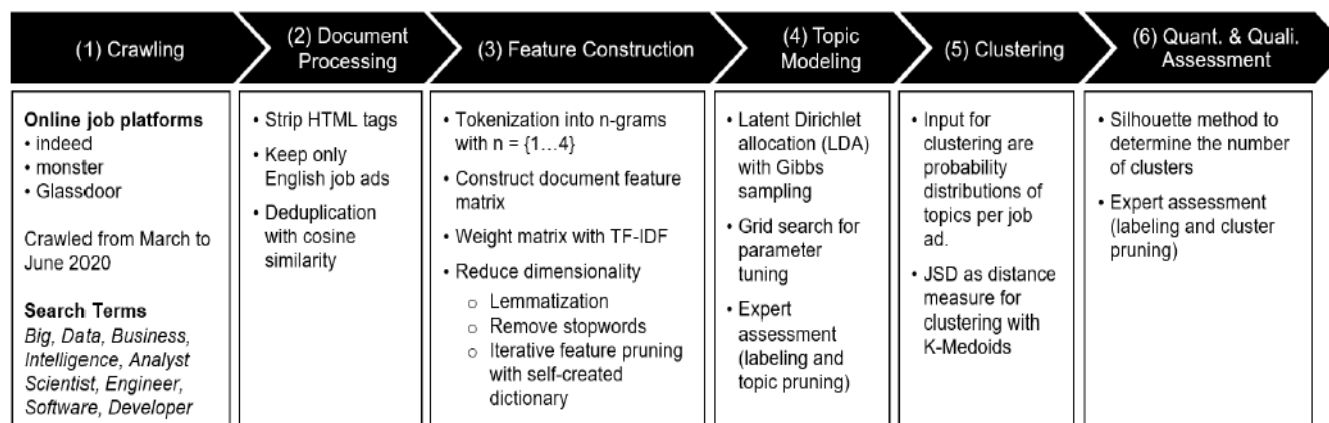
European Conference on Information Systems 2021

 Download: Paper (pdf)  
Size: 614 KB



#### Method | Data-driven

- text mining approach combining **topic modeling**, clustering, and expert assessment
- methodological approach follows six steps:



- to reduce bias in the algorithms' parameter settings, the proposed approach is based on established methodological recommendations



#### Future Research

- text mining approach can be reproduced to study new topics such as job roles in other disciplines like Industry 4.0
- application of a seeded topic model using the results of a series of interviews
- investigation how identified job roles change over time, for instance, by applying an LDA-based topic model over time



#### Scope

job roles in the field of data science



#### Data

25,104 job advertisements published at online job platforms Indeed, Monster, and Glassdoor



#### Key findings

- identify six job roles in data science in request by organizations
- describe 24 topics classified in three major knowledge domains (business, analytical, und technical)

# Ausgewählte Artikel

## Steckbrief 02 / 11

### Understanding Demand for Project Manager Competences in the Construction Industry: Data Mining Approach (2020)

Junping Zheng; Qi Wen; and Maoshan Qiang

Journal of Construction Engineering and Management  
Vol. 146 Issue 8

Download: Paper (pdf)  
Size: 2.5 MB



#### Scope

construction project manager job roles



#### Data

243,521 job advertisements published at top five Chinese job search websites



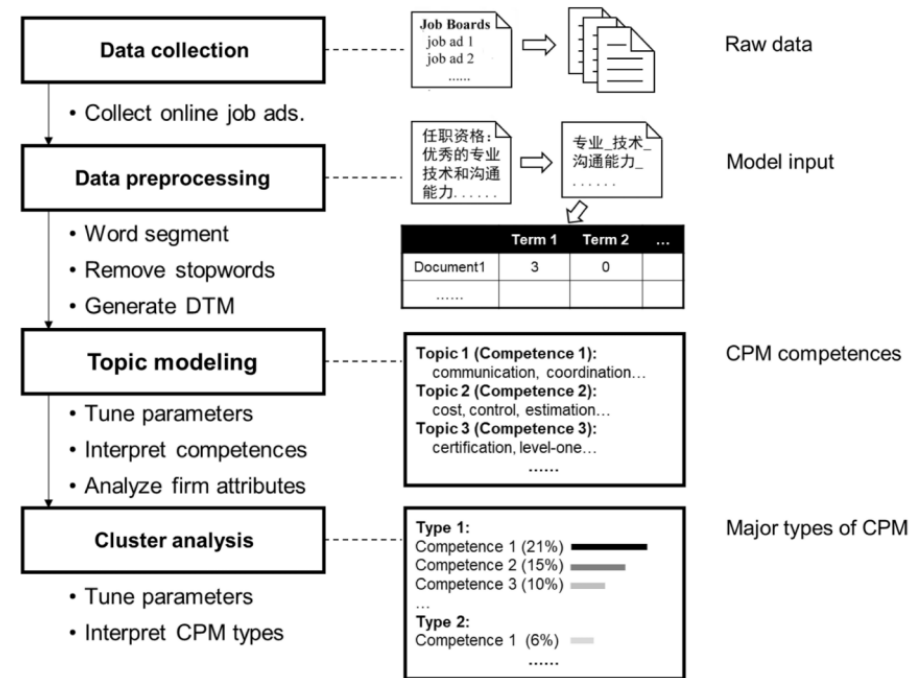
#### Key findings

identify eight major competence dimensions expected by recruiters and five major types of construction project managers



#### Method | Data-driven

- text mining approach combining structure **topic modeling** and clustering (job advertisements are clustered based on their topic distribution)
- methodological approach follows four steps:



#### Future Research

- analyzing data over time to reveal evolution of demands
- operationalizing high-level competences into specific indicators and testing with questionnaire survey data
- extending findings by collecting more data from other countries outside China



# Ausgewählte Artikel

## Steckbrief 03 / 11

### A Longitudinal Analysis of Job Skills for Entry-Level Data Analysts (2020)

Tianxi Dong and Jason Triche

Journal of Information Systems Education Vol. 31 Issue 4

 Download: Paper (pdf)  
Size: 1.3 MB



#### Scope

entry-level data analytics jobs



#### Data

9,163 job advertisements published at Indeed



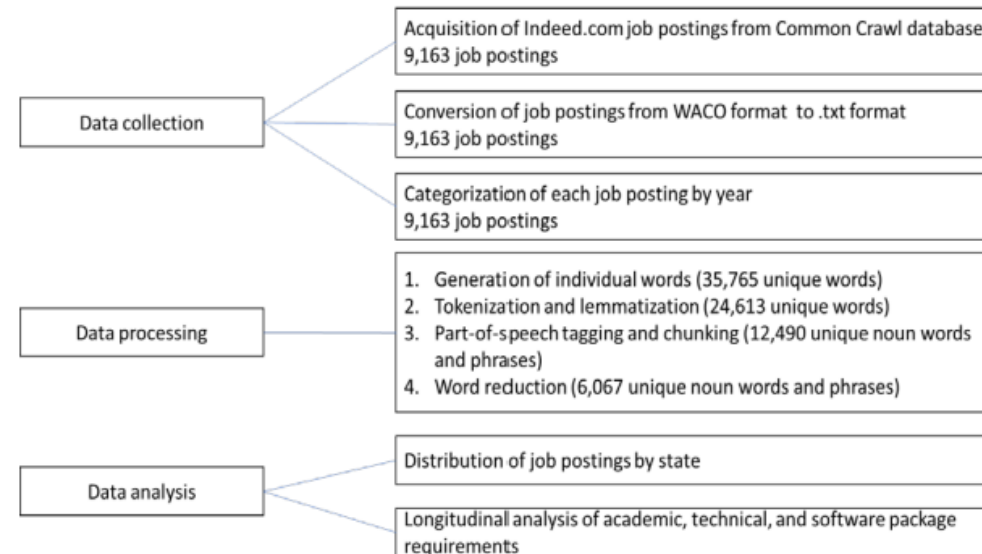
#### Key findings

- identify a preliminary set of key skills sought in practice (programming and data visualization)
- analyze evolution of key skills over time



#### Method | Data-driven

- text mining approach with a custom text mining dictionary
- methodical approach follows three steps:



- empirical analysis of differences, using **linear regression**, in the skillsets between 2014 and 2018 by difference of proportions



#### Future Research

- custom text mining dictionary can assist other researchers in the area of data analytics
- results can help universities and instructors to make better-informed curriculum and teaching decisions

# Ausgewählte Artikel

## Steckbrief 04 / 11

### Bridge the Terminology Gap Between Recruiters and Candidates: A Multilingual Skills Base built from Social Media and Linked Data (2016)

Emmanuel Malherbe and Marie-Aude Aufaure

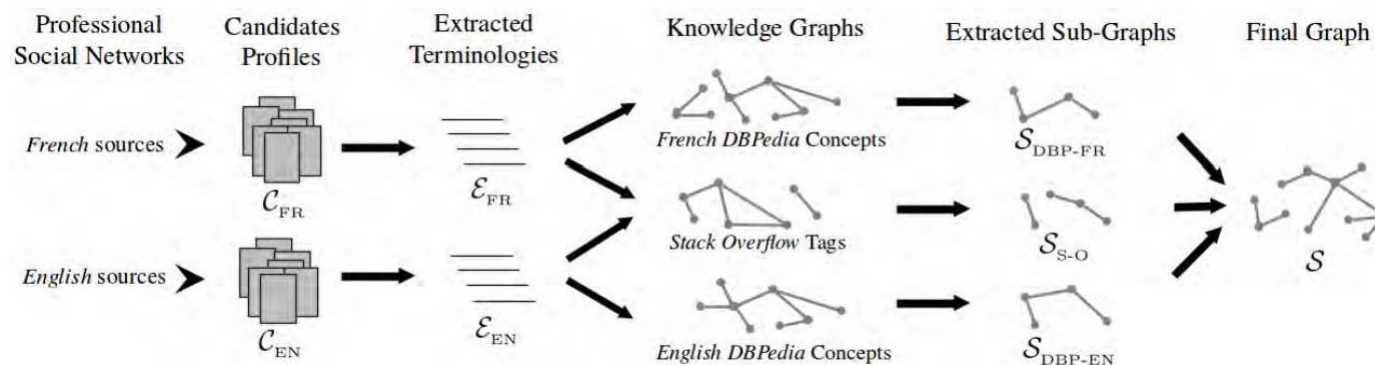
IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2016

Download: Paper (pdf)  
Size: 907 KB



#### Method | Data-driven

- **social network analysis** approach using knowledge graphs from DBpedia and tags from StackOverflow



- constructing skills as concept graphs leads to an automatically assigned category



#### Future Research

- skills base can be updated automatically to follow the new trends in social media and professional social networks
- future work will focus on exploiting the skills hierarchy extracted in the construction, and on extracting and handling the experience for a given skill



#### Scope

multilingual skills base, mainly from France and the U.S.



#### Data

- 100,000 job offers each in French and English from CareerBuilder, Indeed, JobisJob and APEC
- 4,352,948 candidate profiles from Viadeo, XING, SkillsPage and Naymz



#### Key findings

method to generate a knowledge base of multilingual skills by terminologies extracted on candidates' profiles

# Ausgewählte Artikel

## Steckbrief 05 / 11

### SCODIS: Job advert-derived time series for high-demand skillset discovery and prediction (2020)

Elisa Margareth Sibarani and Simon Scerri

International Conference on Database and Expert Systems Applications 2020

 Download: Paper (pdf)  
Size: 967 KB



#### Method | Data-driven

- **clustering approach** with critical indices and cluster categories
  - extract skills with the Skills and Recruitment Ontology (SARO)
  - construct a temporally varying skill graph with the skills as nodes and the number of co-occurrences as weighted edges
  - identify clusters in a temporal snapshot of the skill graph using the Louvain algorithm
  - compute the critical indices (centrality, density) for the clusters
  - classify the clusters into categories (isolated, secondary, principal, crossroads)
  - generate a time series of skill clusters by computation of the cluster similarities between consecutive snapshots
  - use the cluster series to identify trends
- observation of skills clusters in a time series



#### Future Research

- enables the discovery of in-demand skillsets
- extend the framework using other types of learning-based forecasting models
- improve the performance of the framework using semantic similarity



#### Scope

job adverts mostly related to data science and data analysis from 17 European countries



#### Data

620,760 job adverts from the online job portals Adzuna, Indeed and Trovit



#### Key findings

SCODIS: a quadrant-based framework that can observe the evolution of dynamic skills graphs and forecast which skills will be requested

# Ausgewählte Artikel

## Steckbrief 06 / 11

### Skills in the MIS Job Market (2009)

Chuck Litecky; Andrew Aken; Bipin Prabhakar  
and Kirk Arnett

Americas Conference on Information Systems 2009

 Download: Paper (pdf)  
Size: 708 KB



#### Method | Data-driven

- **content analysis** approach of extracted skills
  - data mining of job ads with a set of 1,717 skill search terms built by a combination of previous research
  - manual skill classification based on job definitions of previous research: business skills, soft skills and general technical skills
  - with technical skills broken down into application development skills, database management skills, programming skills, system administration skills and networking skills
  - analysis of skill frequencies in the MIS job ads



#### Future Research

- developing curricula better attuned to the MIS labor market
- making better choices of courses to fit labor market demands
- compare companies' inventory of job skills to those demanded and popular in the labor market



#### Scope

job skills in the Management Information Systems (MIS) job market in the U.S.



#### Data

approximately 218,000 job advertisements from the job web sites Monster, SimplyHired and HotJobs



#### Key findings

the most highly demanded job skills in the current MIS job market

# Ausgewählte Artikel

## Steckbrief 07 / 11

### Learning Representations for Soft Skill Matching (2018)

Luiza Sayfullina; Eric Malmi and Juho Kannala

International Conference on Analysis of Images, Social Networks and Texts 2018

 Download: Paper (pdf)  
Size: 257 KB



#### Method | Data-driven

- **classification approach** with neural networks and different input representations to match CVs and job advertisements
  - creating a supervised dataset for training through a crowdsourcing experiment for the job ads and the CVs corpora
  - training classifiers on the job ads corpus:
    - Long short-term memory (LSTM)
    - Convolutional Neural Network (CNN)
    - Hierarchical Attention Model (HAN)
  - each with three ways of input representation:
    - soft skill masking
    - combination of soft skill masking with a soft skill embedding
    - soft skill tagging
- differentiate between soft skill phrases referring to a candidate vs. something else



#### Future Research

soft skill disambiguation methods open new possibilities for refined analyses of soft skill demands and supply on the job market



#### Scope

soft skill phrases that are ambiguous



#### Data

245,000 job advertisements as a publicly available dataset from Kaggle and 525 CVs from Indeed



#### Key findings

LSTM yields the best disambiguation performance on sentences with tagged soft skills on the job corpus and with a raw sentence representation on the CV corpus

# Ausgewählte Artikel

## Steckbrief 08 / 11

### In Search of the Older Worker: Framing Job Requirements in Recruitment Advertisements (2021)

Martine van Selm and Linda van den Heijkant

Work, Aging and Retirement Vol. 7 Issue 4

 Download: Paper (pdf)  
Size: 281 KB



#### Method | Mixed Method with expert interviews

- automated **content analysis** approach using co-word occurrence analysis
- comparison between job advertisements targeted at older job seekers and general job seekers
- definition of a dictionary based on requested abilities that occurred more than 50 or 100 times, respectively
- **Clustering** using principal component analysis (based on the document-term-matrix) to identify skill clusters
- design of suitable labels for the resulting requirements
- additionally, 11 interviews with HR professionals and recruiters
- interviews used a semi-structured interview protocol structured along five main themes regarding experience with older applicants or capabilities of older employees



#### Future Research

- analyzing the need for meaningful roles and social interaction among older workers
- examining how demographic and conjuncture changes may challenge the older workers' stereotype for the better
- focusing on job seekers in higher job segments



#### Scope

job advertisements targeting older job seekers (45 years old and above)



#### Data

1,158 job advertisements, mostly lower-wage and part-time, published at nine Dutch agencies and job boards



#### Key findings

reveal that stereotypes about older workers in regard to soft and hard skills are reflected in job advertisements and thinking of recruiters

# Ausgewählte Artikel

## Steckbrief 09 / 11

### The Humans Behind Artificial Intelligence – An Operationalisation of AI Competencies (2020)

Eduard Anton, Alina Behne and Frank Teuteberg

European Conference on Information Systems 2020

Download: Paper (pdf)  
Size: 1,6 MB



#### Scope

job roles in the field of artificial intelligence



#### Data

9,247 job advertisements from 60 countries published at the online job platform Indeed



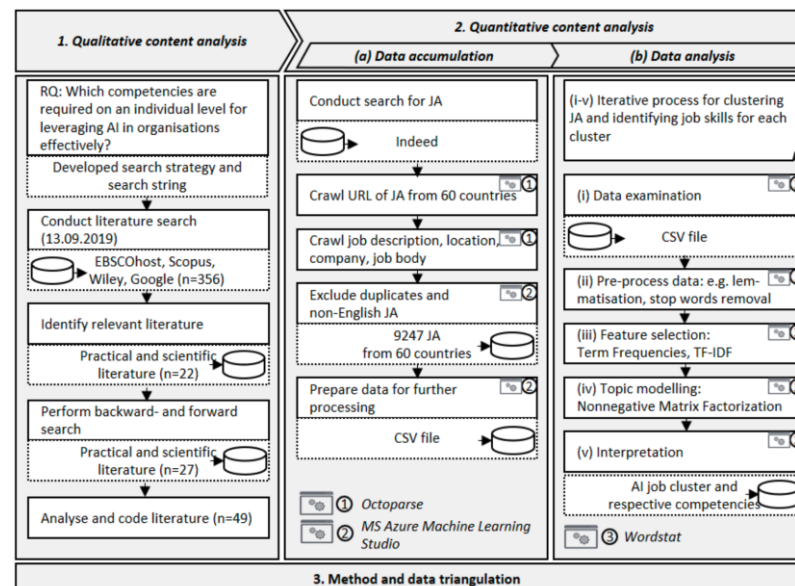
#### Key findings

- comprehensive overview of key technical and managerial competencies essential for implementing and utilizing AI on an individual level
- highlighting the demand for AI-related competencies in three occupational fields



#### Method | Mixed Method with qualitative content analysis

- qualitative content analysis through systematic literature analysis
- text mining approach combining **topic modeling** (nonnegative matrix factorization) and manual clustering of the skills of each topic



#### Future Research

- starting points for developing and promoting the necessary competencies
- examine AI skills on the job to validate the results and identify the skill supply to assess the skill gaps between supply and demand
- study differences in required competencies between sectors

# Ausgewählte Artikel

## Steckbrief 10 / 11

### Human resources for Big Data professions: A systematic classification of job roles and required skill sets (2018)

Andrea De Mauro; Marco Greco; Michele Grimaldi and Paavo Ritala

Elsevier / Information Processing & Management

 Download: Paper (pdf)  
Size: 745KB



#### Method | Mixed Method with expert judgement

- analytical methodology based on a combination of machine learning algorithms and expert judgement
- 4-step methodology:
  - web scraping
  - identification of job families (text mining techniques, expert judgement)
  - identification of skill sets (**topic modeling**)
  - mapping of skill sets by job families (empirical analysis)



#### Future Research

- useful insights for organizations and managers
- useful guidance to educational institutions that aim to focus their efforts in developing skills and competences that are needed in the future
- original method can be reused in similar future studies focusing on any other professional field



#### Scope

job roles in the field of Big Data Analytics



#### Data

2,786 job posts published at the online platform Dice



#### Key findings

- identify four Big Data 'job families'
- recognize nine homogeneous groups of Big Data skills (skill sets)



# Ausgewählte Artikel

## Steckbrief 11 / 11

### Skills2Job: A recommender system that encodes job offer embeddings on graph databases (2021)

Anna Giabelli; Lorenzo Malandri; Fabio Mercorio; Mario Mezzanica and Andrea Seveso

Elsevier / Applied Soft Computing

Download: Paper (pdf)  
Size: 1.1 MB



#### Scope

General job vacancies posted in United Kingdom, France and Germany



#### Data

2,500,000 online job vacancies



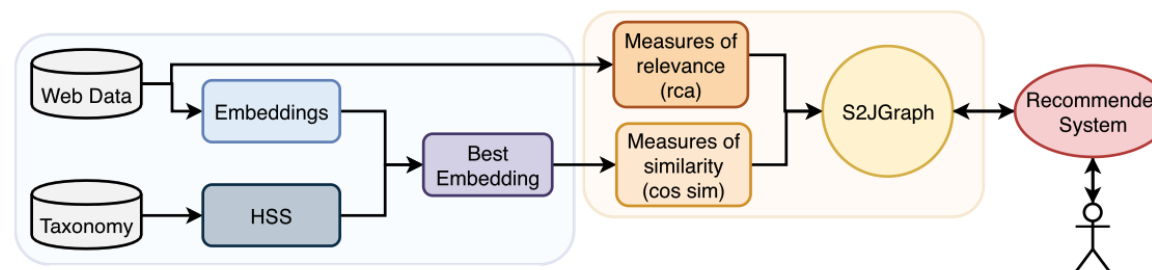
#### Key findings

skills2job, a recommendation system that recommends the most suitable jobs based on the skills specified by the user



#### Method | Mixed Method with expert judgement

- training of skill **word embeddings** based on the co-occurrence of words in a job ad and the ESCO taxonomy using a CBOW approach
- creation of a directed labeled multi-graph derived from both the descriptions of the OJAs, and skills based on the co-occurrences of occupations and skills
- skill based recommendations through revealed comparative advantage-based method and cosine similarity-based method



- Results evaluated through interviews with 10 labor market experts



#### Future Research

- skills2job can be adapted to different countries/industries and easily updated over time
- develop a method to evaluate skill–job fit by employing the information derived from the word embeddings models

# Quellen Arbeitspaket 02

- **Aichholzer, G. (2009):** The Delphi Method: Eliciting Experts' Knowledge in Technology Foresight, in Interviewing Experts, A. Bogner, B. Littig and W. Menz (eds.), Houndmills, UK, pp. 252-274.
- **Almeida, F. and Xexeo, G. (2019):** Word Embeddings: A Survey. arXiv:cs.CL/1901.09069.
- **Ang, S., Joseph, D., and Slaughter, S. A. (2013):** IT Professionals and the IT Profession, *Encycl. Manag. Inf. Syst.*, pp. 1-13.
- **Anton, E. et al (2020):** The Humans Behind Artificial Intelligence—An Operationalisation of AI Competencies.
- **Buchmann, M., Buchs, H., Busch, F., Clematide, S., Gnehm, A.-S., and Müller, J. (2022):** Swiss Job Market Monitor: A Rich Source of Demand-Side Micro Data of the Labour Market, *European Sociological Review*, pp. 1-14.
- **Bogner, A., Littig, B., and Menz, W. (2009):** Introduction: Expert Interviews – An Introduction to a New Methodological Debate,” in Interviewing Experts, A. Bogner, B. Littig and W. Menz (eds.), Houndmills, UK, pp. 1-16.
- **Carnevale, A. P., Jayasundera, T., and Gulish, A. (2016):** America's Divided Recovery: College Haves and Have-Nots, Georgetown University Center on Education and the Workforce, pp. 1-36.
- **Case, D. (2007):** Power and Sample Size, *Methods in molecular biology* (Clifton, N.J.).
- **Cicek, K., Akyuz, E., and Celik, M. (2019):** Future Skills Requirements Analysis in Maritime Industry, *Procedia Computer Science* (158), pp. 270-274.
- **Creswell, J. W., and Clark, V. L. P. (2017):** Designing and conducting mixed methods research, Thousand Oaks, CA: SAGE Publications.

- **De Mauro, A. et al. (2018):** Human resources for Big Data professions: A systematic classification of job roles and required skill sets.
- **Dong, T. and Triche, J. (2020):** A Longitudinal Analysis of Job Skills for Entry-Level Data Analysts (2020) | Journal of Information Systems Education Vol. 31 Issue.
- **Evans, J. R., and Mathur, A. (2005):** The value of online surveys, *Internet Research* (15:2), pp. 195-219.
- **Ganesan, M., Antony, S. P., and George, E. P. (2018):** Dimensions of job advertisement as signals for achieving job seeker's application intention, *Journal of Management Development* (37:5), pp. 425-438.
- **Giabelli, A. et al (2021):** Skills2Job: A recommender system that encodes job offer embeddings on graph databases.
- **Kirchherr, J., Klier, J., Lehmann-Brauns, C., and Winde, M. (2018):** Future Skills: Welche Kompetenzen in Deutschland fehlen, Stifterverband (ed.), Essen, Germany.
- **Kotsiou, A., Fajardo-Tovar, D. D., Cowhitt, T., Major, L., and Wegerif, R. (2022):** A scoping review of Future Skills frameworks, *Irish Educational Studies* (41:1), pp. 171-186.
- **Kurtzo, F., Hansen, M. J., Rucker, K. J., and Edgar, L. D. (2016):** Agricultural Communications: Perspectives from the Experts, *Journal of Applied Communications* (100:1), pp. 33-45.
- **Leopold, T. A., Ratcheva, V. S., and Zahidi, S. (2016):** The Future of Jobs: Employment, Skills, and Workforce Strategy for the Fourth Industrial Revolution," *World Economic Forum* (ed.), Geneva, Switzerland.
- **Lieu, T. T. B., Duc, N. H., Gleason, N. W., Hai, D. T., and Tam, N. D. (2018):** Approaches in Developing Undergraduate IT Engineering Curriculum for the Fourth Industrial Revolution in Malaysia and Vietnam," *Creative Education* (09:16), pp. 2752-2772.

- **Litecky, C., Aken, A., Prabhakar, B., and Arnett, K. (2009):** Skills in the MIS job market. AMCIS 2009 Proceedings, 255.
- **Litecky, C., Aken, A., Ahmad, A., and Nelson, H. J. (2010):** Mining for Computing Jobs, IEEE Software (27:1), pp. 78-85.
- **Maer-Matei, M. M., Mocanu, C., Zamfir, A.-M., and Georgescu, T. M. (2019):** Skill Needs for Early Career Researchers - A Text Mining Approach, Sustainability (11:10), pp. 1-17.
- **Malherbe, E., and Aufaure, M. A. (2016):** Bridge the terminology gap between recruiters and candidates: A multilingual skills base built from social media and linked data. In 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), IEEE, pp. 583-590.
- **Michalczyk, Sven; Nadj, Mario; Maedche, Alexander; and Gröger, Christoph (2021):** Demystifying Job Roles in Data Science: A Text Mining Approach, ECIS 2021 Research Papers. 115.
- **Omana, M.G.H., Engelbrechtb, A., Salmanc, A. (2207):** An overview of clustering methods. Intelligent Data Analysis 11, pp. 583-605.
- **Ra, S., Shrestha, U., Khatiwada, S., Yoon, S. W., and Kwon, K. (2019):** The rise of technology and impact on skills, International Journal of Training Research (17:51), pp. 26-40.
- **Rahmat, A. M., Adnan, A. H. M., and Mohtar, N. M. (2020):** Becoming Industry 4.0 workers: Technical, methodological, social, personal and digital processing capabilities, The management of work performance, safety and entrepreneurship trend in Malaysia, pp. 103-111.
- **Reis, L., Maier, C., and Weitzel, T. 2022.** “Mixed-Methods in Information Systems Research: Status Quo, Core Concepts, and Future Research Implications,” Communications of the Association for Information Systems (51).

- **Rios, J. A., Ling, G., Pugh, R., Becker, D., and Bacall, A. (2020):** Identifying Critical 21st Century Skills for Workplace Success: A Content Analysis of Job Advertisements, *Educational Researcher* (49:2), pp. 80-89.
- **Sayfullina, L., Malmi, E., and Kannala, J. (2018):** Learning representations for soft skill matching. In *International conference on analysis of images, social networks and texts*, Springer, Cham, pp. 141-152.
- **Sibarani, E. M., and Scerri, S. (2020):** SCODIS: Job advert-derived time series for high-demand skillset discovery and prediction. In *International Conference on Database and Expert Systems Applications*, Springer, Cham, pp. 366-381.
- **Sykes (1993):** An Introduction to Regression Analysis, Coase-Sandor Institute for Law & Economics Working Paper No. 20.
- **Todd, P. A., McKeen, J. D., and Gallupe, R. B. (1995):** The evolution of IS job skills: A content analysis of IS job advertisements from 1970 to 1990, *MIS Quarterly*, pp. 1-27.
- **Tremblay, M. C., Hevner, A. R., and Berndt, D. J. (2010):** Focus Groups for Artifact Refinement and Evaluation in Design Research, *Communications of the association for Information Systems* (26:1), Article 27.
- **Van Selm, M. and van den Heijkant, L. (2021):** In Search of the Older Worker: Framing Job Requirements in Recruitment Advertisements (2021), *Aging and Retirement* Vol. 7 Issue 4.
- **Vayansky, I. and Kumar, S.A.P. (2020):** A review of topic modeling methods, *Information Systems* 94:101582.
- **vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., and Cleven, A. (2015):** Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research,” *Communications of the Association for Information Systems* (37:9), pp. 205-224.
- **Wentling, R. M., and Palma-Rivas, N. (1998):** Current status and future trends of diversity initiatives in the workplace: Diversity experts’ perspective, *Human Resource Development Quarterly* (9:3), pp. 235-253.

- **Zahidi, S., Ratcheva, V. S., Hingel, G., and Brown, S. (2020):** The Future of Jobs Report 2020, World Economic Forum (ed.), Geneva, Switzerland.
- **Zheng, J., Wen, Q. and Qiang, M. (2020):** Understanding Demand for Project Manager Competences in the Construction Industry: Data Mining Approach. Journal of Construction Engineering and Management Vol. 146 Issue 8.