



IW-Report 35/19

People Analytics

Evidenzbasierte Entscheidungsfindung im Personalmanagement
Andrea Hammermann; Christopher Thiele

Köln, 02.10.2019

Inhaltsverzeichnis

Zusammenfassung	2
1 Einleitung	3
2 Stand der Forschung und begriffliche Verortung	4
2.1 Zum Begriff People Analytics	4
2.2 Bedeutung von People Analytics in der wissenschaftlichen Literatur	7
3 Vorgehen, Daten und Methodik	10
3.1 Aufbau eines People Analytics Projekts	10
3.2 Daten und Datenquellen	11
3.3 Analyse und Interpretation	14
4 Anwendungsfelder im HR	17
4.1 Schwerpunkt Personalbeschaffung	18
4.2 Schwerpunkt: Personalentwicklung und -führung	19
5 Rechtliche und ethische Betrachtung	21
5.1 Entscheidungsfindung	21
5.2 Gleichbehandlung	22
5.3 Datenschutz	24
5.4 Betriebliche Mitbestimmung	25
6 Handlungsfelder und Ausblick	27
Abstract	28
Tabellenverzeichnis	29
Abbildungsverzeichnis	29
7 Literaturverzeichnis	30

JEL-Klassifikation:

M12 Personnel Management; Executives; Executive Compensation

M50 – Personnel Economics

Zusammenfassung

Die Menge an digitalen Daten steigt rasant. Im Jahr 2018 wurden 33 Zettabytes (33 Milliarden Terabytes) an Daten generiert und für 2020 werden 47 Zettabytes prognostiziert. In den meisten Unternehmen wird das Potenzial verfügbarer Daten zur Unterstützung von (Investitions-) Entscheidungen noch nicht voll ausgeschöpft. Dies gilt insbesondere für das Personalwesen, in dem Entscheidungen über die Einstellung neuer Mitarbeiter bis hin zu Entwicklungsschritten und Beförderungsentscheidungen häufig allein auf Intuition und Erfahrung basieren.

Unter den Begriffen People Analytics oder HR Analytics werden vor allem in der Managementliteratur, aber zunehmend auch in der wissenschaftlichen Forschung, datenbasierte Ansätze der Entscheidungsfindung diskutiert und ihre Potenziale nicht selten überzeichnet. Vermeintliche Vorteile wie Objektivität und Diskriminierungsfreiheit einer datenbasierten Entscheidungsfindung sind aus wissenschaftlicher Sicht jedoch kritisch zu hinterfragen.

Die vorliegende Studie setzt sich mit den Ansätzen, Datenquellen und Methodiken von People Analytics auseinander und diskutiert ihre Chancen und Risiken aus einer ökonomischen, rechtlichen und ethischen Betrachtung heraus. Eine klare Empfehlung für mehr People Analytics in Unternehmen gibt die Studie nicht. Die Autoren möchten vielmehr für Hemmnisse in der Datennutzung sensibilisieren, deren Überwindung für eine wertschöpfende Anwendung von People Analytics notwendig ist. Die Hemmnisse werden vor allem in der verbreiteten Rechtsunsicherheit im Umgang mit Daten, einer mit Ängsten und Vorbehalten behafteten Diskussion über die Datennutzung im Arbeitskontext und fehlenden Erfahrungen und Kompetenzen in der Datenaufbereitung und -analyse gesehen.

1 Einleitung

Der Film „Die Kunst zu gewinnen – Moneyball“ von Bennet Miller aus dem Jahr 2011 zeigt eindrucksvoll die Bedeutung von statistischen Analyseverfahren im Spitzensport. Erzählt wird, basierend auf der Buchvorlage von Michael Lewis (Lewis, 2013), der Aufstieg der Oakland Athletics – einer Baseballmannschaft, die von ihrem Manager Billy Beane in den 1990er Jahren anhand von wissenschaftlich fundierten Auswahlkriterien zusammengestellt wurde. Damit setzte sich der Manager über das Urteil erfahrener Scouts hinweg, warb kostengünstigere Spieler an und erzielte dadurch, trotz finanzieller Unterlegenheit des Vereins gegenüber der Konkurrenz, viel beachtete Erfolge.

Statistische Auswertungen sind aus dem Spitzensport nicht mehr wegzudenken. In kaum einem anderen Berufsfeld liegen so viele Beobachtungen über die Anstrengung und Leistungen des Einzelnen und die Interaktion im Team vor, die zur Auswahl von Sportlern sowie zur Optimierung des Trainingsplans und des taktischen Vorgehens genutzt werden. Doch auch abseits des Sports gewinnen quantitative Analysen bei der Personalauswahl und der Förderung von Talenten an Bedeutung. Getrieben wird der Trend nicht zuletzt aufgrund der exponentiell wachsenden Datenmenge, die Aufschluss über eine der wichtigsten betrieblichen Ressourcen bieten kann: die Mitarbeiter. Dabei hat das Zeitalter von Big Data gerade erst begonnen. Im Jahr 2018 wurden rund 33 Milliarden Terabyte erzeugt. Das entspricht rund 660 Milliarden Bluray CDs und es wird ein exponentieller Anstieg erwartet (Statista, 2019).

Während andere Unternehmensbereiche wie Vertrieb, Controlling oder Marketing eine lange Tradition haben, Daten für sich nutzbar zu machen, werden Kennzahlen zum effizienten Personaleinsatz, zum effektiven Trainingsangebot oder dem Wertschöpfungsbeitrag von HR-Aktivitäten nur selten erhoben und ausgewertet. Hinzu kommt, dass die HR-Abteilung von der Geschäftsführung häufig nicht als strategischer Partner auf Augenhöhe wahrgenommen wird (Boudreau/Lawler, 2015). Der Fülle an Beratungsangeboten und den Predigten von Management-Gurus zufolge könnte sich dies nun ändern. Das Schlagwort lautet People Analytics.

Die vorliegende Analyse setzt sich kritisch mit dem Trendthema People Analytics auseinander. In Kapitel zwei wird der Begriff eingeordnet und der wissenschaftliche Diskurs analysiert. Das Vorgehen von People Analytics Projekten, die möglichen Datentypen, -quellen und Methodiken werden in Kapitel drei beleuchtet. In Kapitel vier werden Schwerpunkte zur Anwendung von People Analytics untersucht und im Personalprozess verortet, gefolgt von einer Betrachtung der rechtlichen und ethischen Rahmenbedingungen. Abschließend werden Handlungsfelder aufgezeigt und ein Ausblick skizziert.

2 Stand der Forschung und begriffliche Verortung

„It is clear that HRA [Human Resource Analytics] in a relatively short period of time has risen to prominence in the HR community. [...] Concepts that fit well with the current zeitgeist will stand a greater chance of becoming fashionable. In this regard, HRA benefits from the strong position of the evidence-based movement in the social sciences [...] and the historical and ongoing quest within the field of HR to play a more important and strategic role within the organization [...].“

(Madsen/Slåtten, 2017, 153)

2.1 Zum Begriff People Analytics

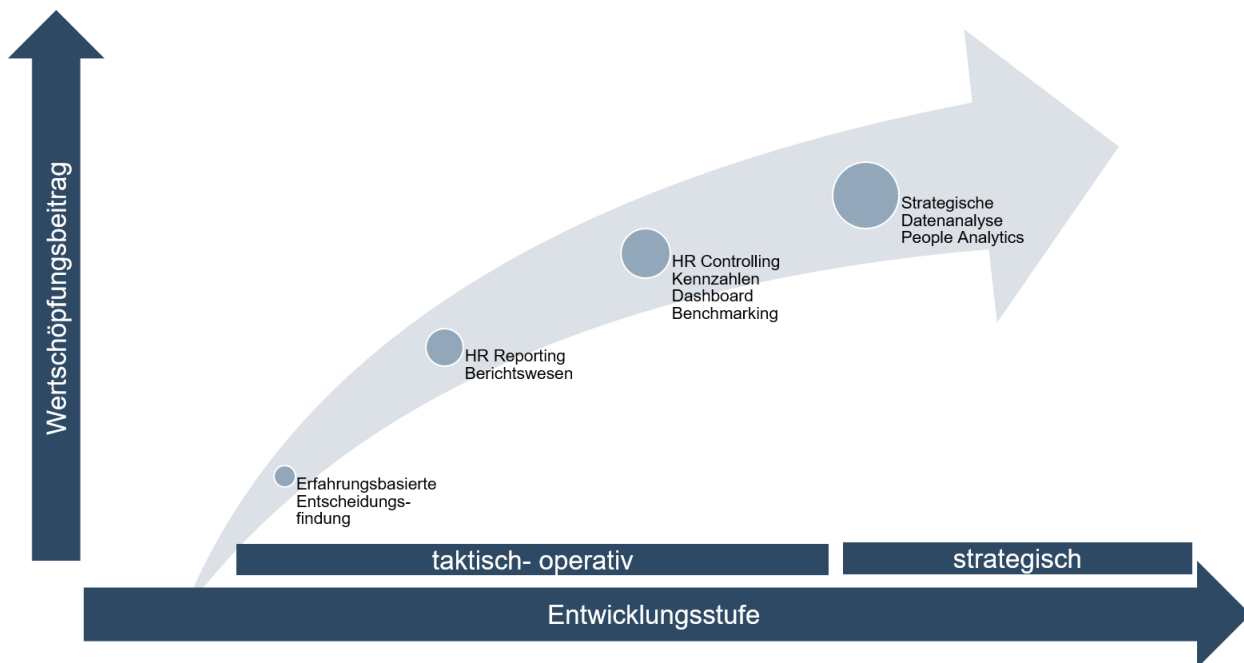
Der HR-Bereich ist häufig einer der am wenigsten datengetriebenen Unternehmensbereiche und unterliegt im Umgang mit Daten besonderen Vertraulichkeits- und Datenschutzbestimmungen. Mit zunehmender Datenmenge, generiert durch IT-gestützte Prozesse sowie digitale Kommunikations- und Informationstechnik in Kombination mit der hohen Bedeutung des Faktors Arbeit für die unternehmerische Wertschöpfung ist jedoch anzunehmen, dass die Datenanalyse im Personalwesen an strategischer Bedeutung gewinnen wird. In der Literatur haben sich hierfür die Begriffe People Analytics oder auch HR Analytics durchgesetzt. Es finden sich aber auch zahlreiche Begriffe, die synonym verwendet werden wie Workforce Analytics, Workforce Intelligence oder Talent Analytics. People Analytics sind eine Teilmenge von Business Analytics, in denen geschäftsrelevante, meist vergangenheitsbezogene Daten ausgewertet werden, um Ableitungen für die zukünftige Entwicklung vornehmen und Entscheidungen evidenzbasiert treffen zu können. Entscheidungen sollen so unabhängiger von den Vorerfahrungen und der subjektiven Haltung des Einzelnen werden.

Abbildung 2-1 skizziert die Entwicklung der Datenorientierung im HR anhand von vier Entwicklungsstufen. Während die Güte von erfahrungsbasierten Entscheidungen primär von dem Erfahrungswissen des Entscheiders abhängt und in der Regel über die berufliche Laufbahn hinweg zunimmt, schafft das Berichtswesen einen Überblick über wichtige Kennzahlen im HR wie beispielsweise Krankenquoten, Fluktuationsraten oder Vakanzenzeiten offener Stellen, die über den zeitlichen Verlauf hinweg Rückschlüsse auf (Fehl-) Entwicklungen zulassen. Durch die rückblickende Perspektive fallen Handlungsbedarfe je nach Berichtsintervall (in der Regel Quartals- oder Jahresberichte) allerdings nur mit einiger Verzögerung auf.

Während sich das Unternehmenscontrolling Ende des 19. Jahrhunderts in den USA zeitgleich mit dem Entstehen von Großunternehmen entwickelte, fand das Controlling erst in den 1980er Jahren Einzug in die HR-Abteilungen. Die ersten HR-Kennzahlen und HR-Benchmarks veröffentlichte Dr. Jac Fitz-enz in den USA 1978 bzw. 1984 (Fitz-enz, 2001; Reindl/Krügl, 2017, 32). Im Gegensatz zum Berichtswesen bietet ein modernes HR-Controlling mit einem umfangreichen Kennzahlensystem verdichtete Informationen über den aktuellen Stand und die zeitliche Entwicklung von zentralen HR-Aktivitäten. Mit der von Kaplan und Norton (1996) entwickelten Balance Scorecard stehen neben den Finanzkennzahlen erstmals auch prozess-, mitarbeiter- und kundenbezogene Kennzahlen im Fokus, die als mögliche Frühindikatoren für die finanzielle Entwicklung des Unternehmens dienen können (Bhattacharyya, 2017, 103; Kaplan/Norton, 2009). Aufbauend auf der Arbeit von Kaplan und Norton stellt die

Veröffentlichung von Becker et al. im Jahr 2001 zur HR-Scorecard einen weiteren Meilenstein für die strategische Bedeutung des kennzahlenbasierten Personalwesens dar (Becker et al., 2007; Bhattacharyya, 2017, 105). Die Autoren plädieren dafür, HR als strategischen Business Partner zu entwickeln. Gerade Mitarbeiterkennzahlen wie hohe Fluktuationsraten oder Krankenquoten können darauf hindeuten, dass interne Prozesse nicht optimal laufen. Dies dürfte sich in den meisten Fällen mittel- bis langfristig auch auf Kennzahlen wie den Umsatz und die Kundenzufriedenheit auswirken, wenn sich die fehlenden internen Kapazitäten in verzögerten Lieferungen oder Qualitätseinbußen niederschlagen. Je stärker die Ursache-Wirkungsbeziehungen von Kennzahlen im Fokus stehen, umso besser lassen sich auf Basis der Kennzahlen strategische Ableitungen ziehen, die Fehlentwicklungen verhindern, Investitionen und Ressourcen steuern und damit Wertschöpfung für das Unternehmen erbringen können.

Abbildung 2-1: Entwicklung der Datenorientierung im HR



Quelle: Reindl, 2016, 194

People Analytics werden als nächster Evolutionsschritt gesehen, indem Daten aus verschiedenen Datenquellen kombiniert werden (z. B. aus dem Qualitätsmanagement, dem Vertrieb oder den sozialen Medien), um Zusammenhänge zu identifizieren und damit Vorhersagen über zukünftige Entwicklungen bzw. Szenarien treffen zu können. Tabelle 2-1 fasst die wichtigsten Differenzierungsmerkmale zwischen People Analytics und dem Personalcontrolling zusammen. Die Übergänge sind in der Realität fließend und in einem Betrieb lassen sich mehrere koexistierende Entwicklungsstufen feststellen. Der gesteigerte Wertschöpfungsbeitrag durch die Verwendung von Daten im HR-Wesen unterliegt allerdings dem Vorbehalt, dass die Forschungsfragen, Analysemethoden, Datengrundlagen sinnvoll gewählt und die daraus gezogenen Schlüsse in wertschöpfende Aktivitäten umgesetzt werden.

Tabelle 2-1: People Analytics versus Personalcontrolling

Personalcontrolling	People Analytics
Sekundäranalytisch, d. h. die Daten werden nicht primär für das Personalcontrolling erfasst	Daten aus unterschiedlichen betriebsinternen wie -externen Quellen werden kombiniert
Im Fokus stehen Kennzahlen (Key Performance Indikatoren)	Im Fokus stehen (im besten Fall kausale) Zusammenhänge
Grundlage ist ein festgelegtes Kennzahlensystem	Grundlage ist eine Hypothese bzw. Forschungsfrage
Rückwärtsgewandter Ansatz	Prädiktiver Ansatz

Quelle: in Anlehnung an Reindl/Krügl, 2017, 35 ff.

People Analytics können Wertschöpfung dadurch generieren, indem Entscheidungsprozesse verkürzt und Entscheidungen anhand der Datengrundlage leichter und „besser“ im Sinne von evidenzbasiert gefällt werden. Zudem kann die quantitative Analyse auch das Verständnis über Wirkungszusammenhänge der Organisation – beispielsweise zwischen ungewollter Fluktuation und dem Führungsverhalten – stärken. Anhand der empirischen Befunde lassen sich Abläufe, Strukturen oder betriebliche Angebote möglicherweise effizienter gestalten oder auch die Einhaltung von gesetzlichen Auflagen z. B. mit Blick auf den Arbeitsschutz kontrollieren. Für das Personalwesen bieten People Analytics darüber hinaus die Chance, die eigenen Aktivitäten noch besser an der Unternehmensstrategie und den Bedürfnissen der Belegschaft auszurichten. Zudem kann der Nutzen von HR-Aktivitäten messbarer und für die Geschäftsführung sichtbarer gemacht werden und damit Investitionen in HR-Projekte begünstigen. Mitarbeiterdaten werden darüber hinaus auch als Trainingsmasse für maschinelles Lernen verwendet, um Abläufe mithilfe von Algorithmen zu automatisieren (Borowczak/Yang, 2018).

Den Potenzialen stehen aber auch eine Reihe struktureller Probleme gegenüber, die der Verbreitung von People Analytics im Weg stehen. Dazu gehören insbesondere inkompatible Dateninfrastrukturen im Betrieb sowie fehlende Analysekompetenzen. Für die Analyse komplexer und unstrukturierter Daten, wie beispielsweise Texte, sind spezielle Informatik- und Statistikkenntnisse notwendig, die über gängige Controllingkenntnisse hinausgehen. Die sich ändernden Anforderungen schlagen sich auch in einer neuen Berufsbezeichnung nieder (Heupel/Lange, 2019). Der Data Scientist analysiert und verarbeitet „(..) mit Methoden der Mathematik, Informatik und Statistik große Datenmengen in Echtzeit aus verschiedenen Quellen ("Big Data"), um z. B. über potenzielle Kunden oder Markttrends nutzbare Informationen zu erlangen“ (Berufenet, 2019).“

Beachtung finden People Analytics bislang vor allem in der Managementliteratur. Darin sind Beiträge von Beraterfirmen und Technologieunternehmen wie IBM, SAP oder Oracle vorherrschend, die ein sehr geschöntes Bild des Themas zeichnen. Neben anbieterseitigen Artikeln, die durch Personalmagazine große Verbreitung in Personalerkreisen finden, werden People

Analytics auch durch zahlreiche Fachtagungen und Ratgeber zum Trendthema (Madsen/Slåtten, 2017). Die wissenschaftliche Fachliteratur ist dagegen deutlich vorsichtiger, inwieweit der Trend die Arbeit im Personalwesen wirklich verändern wird und wie die Chancen und Risiken zu bewerten sind (Angrave et al., 2016; Holthaus et al., 2015; Jain/Maitri, 2018). Insbesondere empirische Evidenz zum Effekt von People Analytics auf Erfolgskennzahlen ist bislang kaum vorhanden (Marler/Boudreau, 2017, 20). Eine der wenigen Ausnahmen ist die Studie von Aral et al. (2012), die einen produktivitätssteigernden Effekt von Human Capital Management Software in Verbindung mit leistungsabhängiger Vergütung und Leistungsbewertung nachweist.

2.2 Bedeutung von People Analytics in der wissenschaftlichen Literatur

Wie bedeutend das Thema People Analytics in der wissenschaftlichen Forschung ist, soll mithilfe der Anzahl an veröffentlichten Artikeln zu diesem Thema in den letzten Jahren näherungsweise bestimmt werden (Marler/Boudreau, 2017; Marler et al., 2017, 63 ff.). Für die Analyse wird die Literaturdatenbank Microsoft Academic (Sinha et al., 2015) genutzt. Microsoft Academic ist eine Suchmaschine für akademische Veröffentlichungen, die seit 2016 als Nachfolge der Microsoft Academic Search aktiv und mit Suchmaschinen wie Web of Science, Scopus und Google Scholar vergleichbar ist (zum Vergleich der Suchmaschinen siehe Paszcza, 2016). Der hier verwendete Datensatz¹ enthält 124.923.119 Einträge (Stand: 14.6.2019) mit Abstracts.

In einem ersten Schritt werden die Titel und Abstracts der Veröffentlichungen nach einer ersten Vorbereitung, wie der Umwandlung in Kleinbuchstaben, nach den Begriffen „human resource“, „hr“, „workforce“, „people“ oder „talent“ durchsucht. Diese Texte stellen die Datengrundlage für explorative Analysen dar. Diese Analysen sowie die Literaturrecherche ergeben, dass die Begriffsfolgen „human resource analytic“, „hr analytic“, „people analytic“, „talent analytic“, „workforce analytic“, und „workforce intelligence“ das wissenschaftliche Gebiet zu einem großen Teil abdecken. Die Begriffe „talent intelligence“ und „people intelligence“ sind nicht trennscharf und werden daher nicht explizit mit aufgenommen. Darüber hinaus werden nur Texte mit dem Dokumententyp „Book“, „BookChapter“ und „Journal“ berücksichtigt². Mit dieser Begriffsdefinition können 228 Veröffentlichungen zwischen 1986 und 2018 identifiziert werden. Von diesen können nach Durchsicht der Abstracts 14 als unpassend eingestuft und aussortiert werden. Von den verbleibenden 214 Fachartikeln und Buchbeiträgen sind 167 seit dem Jahr 2010 und 120 seit dem Jahr 2014 veröffentlicht worden (siehe Abbildung 2.2).

Über die Zeit betrachtet werden die Begriffsfolgen mit „human resource“ und „hr“ am häufigsten in dem hier beschriebenen People Analytics Kontext benutzt.

Zur weiteren Untersuchung der Veröffentlichungen wird eine Clusteranalyse mit drei Gruppen vorgenommen. Dies geschieht mithilfe eines strukturierten Topicmodels (Roberts et al., 2013). Die Vorverarbeitung umfasste neben dem Entfernen von Füllwörtern nach ihrer absoluten Häufigkeit anhand der empfohlenen Einstellungen des R Pakets stm (Roberts et al., 2018) auch

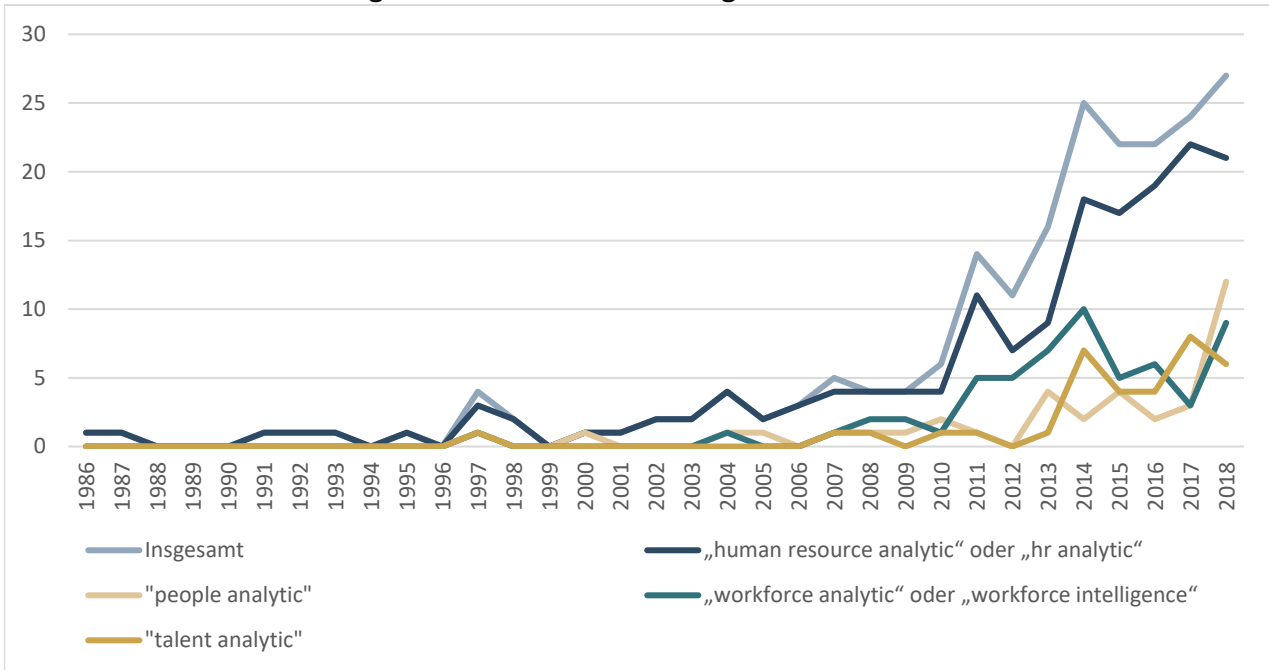
¹ Die Analyse nutzt Informationen von Microsoft Academic. Die Datennutzung unterliegt der ODC Attribution Licence.

² Das Datenschema kann unter <https://docs.microsoft.com/en-us/academic-services/graph/reference-data-schema> eingesehen werden; zuletzt abgefragt am 10.7.2019

eine Reduktion auf den Wortstamm. Die Anzahl an Themen wurde, durch Inspektion der Exklusivität und Kohärenz der einzelnen Themen in Modellen von bis zu 20 Themen, auf drei festgelegt. Auch nach dem Lesen aller Abstracts und Titel erscheinen diese Gruppierungen sinnvoll.

Abbildung 2-2: Auswertung der wissenschaftlichen Fachliteratur

Anzahl der Veröffentlichungen nach verwendeten Begriffen



Je nach den im Abstract verwendeten Begriffen, können die Papiere mehrfach in unterschiedlichen Kategorien gezählt werden.

Quelle: Microsoft Academic Graph 2019

Ein Cluster umfasst 12 Prozent der Dokumente, die sich explizit mit dem Gesundheitssektor befassen. Diese sind größtenteils Veröffentlichungen des in England im Jahr 2010 gegründeten „Centre for Workforce Intelligence“, welches mit der Beratung über die Personalplanung im Gesundheitswesen beauftragt ist. Allerdings handelt nicht jede Veröffentlichung von (daten-) analytischen Themen.

Die restlichen 187 Veröffentlichungen betreffen verschiedene Aspekte des Kernthemas People Analytics. Das Clusterverfahren teilt diese in zwei weitere Cluster, die allerdings bereits deutlich schwieriger zu differenzieren sind. Ein Unterschied stellt beispielsweise der Anteil an technischen oder methodischen Begriffen dar. Dies führt dazu, dass ein Cluster mit 62 Dokumenten häufig selbst empirische Analysen beinhaltet oder ein empirisches Modell und Metriken beschreibt. Das letzte Cluster umfasst 125 Dokumente und behandelt Entwicklungen und Anwendungen von HR-Themen aus unterschiedlichsten Blickwinkeln. Tabelle 2-2 zeigt die Titel von drei Veröffentlichungen je Cluster.

Auf Basis des Microsoft Academic Graph, lässt sich festhalten, dass die absolute Anzahl an wissenschaftlichen Publikationen zu dem Thema People Analytics steigt (wie auch die gesamte Anzahl an Publikationen), dass diese sich auf die Begrifflichkeiten des „Human Resource

Management“ fokussieren und die Mehrzahl der Papiere keinen eigenen datenanalytischen Schwerpunkt haben, sondern sich mit (theoretischen) Anwendungsfällen beschäftigen. Die Tatsache, dass Institutionen, wie das „Centre for Workforce Intelligence“ entstehen, bestärkt die Vermutung, dass eine datengetriebene und empirisch fundierte Sichtweise auf Personalthemen kein kurzlebiger Managementtrend ist.

Tabelle 2-2: Cluster

Drei exemplarische Titel pro Cluster

Cluster 1: Centre for Workforce Intelligence	Workforce planning and flexible working in the NHS of the future
	Training should focus on shift to community care report urges
	Boost GP trainee posts at expense of hospital specialties recommends centre for workforce intelligence
Cluster 2: Empirie	Some Insights into Procedures and Practices to Acquire the Right Talent
	Predicting knowledge workers' participation in voluntary learning with employee characteristics and online learning tools
	Big Data and Hris Used by Hr Practitioners: Empirical Evidence from a Longitudinal Study
Cluster 3: Anwendungen und Trends	HRM practices and organizational performance in the UAE banking sector
	Significant developments and emerging issues in human resource management
	Human resources management in multinational companies in Africa: a systematic literature review

Quelle: Microsoft Academic Graph 2019, eigene Darstellung

3 Vorgehen, Daten und Methodik

„Data by themselves are informative but do not give insights that are usable in decision making or planning. Only through context [..], and supplied with meaning and by understanding relationships [..], do data become information. Information transforms into knowledge when combined with experience, cognition and competence [..].“

(Scholz, 2017, 13)

3.1 Aufbau eines People Analytics Projekts

People Analytics Projekte lassen sich in fünf Schritte untergliedern, von denen die eigentliche statistische Auswertung nur einen Teil ausmacht (vgl. Bhattacharyya, 2017, 77 f.; Biemann et al., 2017, 10; Scholz, 2017, 20; Strohmeier/Piazza, 2015; Biemann et al., 2017, 10):

- 1. Planung/Festlegung des Forschungsdesign:** Ausgehend von der Hypothese, die durch die Datenanalyse untersucht werden soll, wird das Forschungsdesign festgelegt. Dafür werden der Informationsbedarf, die Datenquellen und Analysemethoden ausgewählt und die Durchführbarkeit, basierend insbesondere auf den rechtlichen Vorgaben zur Mitbestimmung und des Datenschutzes und der technischen Umsetzbarkeit, geprüft.
- 2. Aufbereitung der Daten:** Die Aufbereitung und Zusammenführung der Daten aus unterschiedlichen Quellen hängt davon ab, wie kompatibel die einzelnen Datenquellen sind und in welcher Form die Daten vorliegen. Softwareanbieter stellen eine Reihe an Analysewerkzeugen zur Verfügung, die jeweils unterschiedliche statistische Verfahren und grafische Darstellungen ermöglichen.
- 3. Data-Mining:** Data-Mining bedeutet, anders als der Begriff nahelegt, nicht die Sammlung von Daten, sondern die Gewinnung von Erkenntnissen über Zusammenhänge und Muster aus großen Datenmengen mithilfe von statistischen Verfahren, Algorithmen und maschinellem Lernen.
- 4. Interpretation:** Basierend auf Erfahrungen und praktischem wie statistischem Wissen müssen die empirischen Befunde eingeordnet und analysiert werden, um Entscheidungen ableiten zu können. Damit ist jede evidenzbasierte Analyse auch ein Stück weit subjektiv, kontextabhängig und beruht auf getroffenen Annahmen. Zu der Interpretation gehört auch die Bewertung zur Güte der Schätzung. Stichprobenselektionen oder andere (strukturelle) Verzerrungen wie Scheinkorrelationen müssen bei der Interpretation berücksichtigt werden.
- 5. Ableitung von Maßnahmen:** Bereits in der Planungsphase ist es wichtig, ein interdisziplinäres Team unter Beteiligung der betroffenen Geschäftsbereiche zusammenzustellen und auch den Betriebsrat und die Leitungsebene einzubeziehen. Damit das People Analytics Projekt mehr ist als eine empirische Übung, müssen die Entscheidungsebenen von dem praktischen Nutzen überzeugt werden, abgeleitete Maßnahmen und Veränderungen auch umzusetzen.

Der vorgestellte Ablauf in fünf Schritten zeigt, dass ökonomische Expertise und praktische Erfahrungen nur gemeinsam zu einem erfolgreichen People Analytics Projekt führen können. Simón und Ferreiro (2018) beschreiben anhand eines Projekts im Handel, wie Forschung und Praxis erfolgreich zusammenarbeiten können, um Mitarbeiterdaten in Unternehmen besser zu verstehen. Die Autoren weisen jedoch auch auf den sogenannten *scholar-practitioner gap* hin, der sich aufgrund des verschiedenen Sprachgebrauchs und der unterschiedlichen Zielsetzungen ergeben kann. Während für Praktiker klare Ableitungen und ein möglicher Effizienzgewinn im Vordergrund stehen, sind für Forscher die Genauigkeit der Schätzung und die Verwendung methodischer Exzellenz wichtig. Lassen sich Forscher und Praktiker jedoch auf die jeweils andere Perspektive ein, können People Analytics Projekte das generelle Verständnis von Zusammenhängen in der Forschung verbessern und gleichzeitig betriebspezifische Beratung für das Unternehmen bieten (vgl. Rasmussen/Ulrich, 2015, 238 ff.).

3.2 Daten und Datenquellen

Daten werden auch als das „neue Öl“ bezeichnet, welches noch in vielen Unternehmen ungenutzt unter der Oberfläche schlummert. So bewerten die meisten Unternehmen der Industrie und industrienahen Dienstleistungen in Deutschland ihre Daten nicht, beispielsweise anhand der Aktualität und Qualität der Daten, und haben keinen Überblick über die ökonomischen Nutzungsmöglichkeiten oder eine Grundlage für eine effiziente Bewirtschaftung der Daten (Engels, 2018). Denn um Daten für das Geschäftsmodell wertschöpfend einsetzen zu können, braucht es ein Verständnis über die Datenstruktur, um mittels statistischer Methoden aus den verfügbaren Daten Informationen beispielsweise über Kundenwünsche, Produktionsprozesse und den effizienten Einsatz des Personals gewinnen zu können. Durch die große Verbreitung von digitalen Informations- und Kommunikationsmitteln, große Speichervolumen in Cloud-Systemen und die zunehmende Vernetzung von Menschen, Maschinen und Objekten im Wertschöpfungsprozess hat die Menge an verfügbaren Daten in den letzten Jahren enorm zugenommen. Diese liegen in Unternehmen zum größten Teil unstrukturiert in Form von Bild, Ton oder Textdateien vor (vgl. Strohmeier, 2015, 9).

In diesem Zusammenhang hat sich in der Literatur und auch der öffentlichen Wahrnehmung der Begriff Big Data etabliert (für eine Bestandsaufnahme siehe Engels/Goecke (2019)). Big Data beschreibt große komplexe Datensätze, die sich durch verschiedene Kriterien von herkömmlichen Datensätzen unterscheiden. Zu den wichtigsten Dimensionen von Big Data gehören (Finlay, 2014, 13 f.; Scholz, 2017, 14 f.):

- **Großer Datenumfang (volume):** Big Data beschreibt Datensätze, die nicht mehr auf einem einzigen Rechner verarbeitet werden können. Finlay gibt in seiner Veröffentlichung 2014 ein Datenvolumen ab einer Größe von einem Terabyte (1000 Gigabyte) an.
- **Verschiedene Datentypen (variety):** Big Data besteht häufig aus strukturierten und unstrukturierten Daten in Form von Texten, maschinell generierten Daten (z. B. über GPS-Tracker), Audio- oder Videodateien. Strukturierte Daten sind beispielsweise Angaben zur Person in der Personalakte. E-Mail-, Chatverläufe, transkribierte Telefoninterviews oder Social Media Beiträge sind Beispiele für unstrukturierte Daten.

- **Verschiedene Datenquellen (multi-sourced):** Big Data verbindet häufig mehrere unterschiedliche betriebsinterne wie externe Datenquellen miteinander.
- **Variabilität der Ausprägungen (volatility/velocity):** Neben eher zeitinvarianten Daten beinhaltet Big Data auch Daten, deren Ausprägungen sich sehr schnell ändern können, wie beispielsweise der Gemütszustand einer Person.

Um zu entscheiden, ob die Aufbereitung und Analyse von Big Data lohnen, stellt sich die Frage, inwieweit die Zusammenführung der Datensätze auch einen zusätzlichen Informationsgehalt für die Beantwortung der Fragestellung liefert. Für People Analytics sind Verhaltensdaten von besonderem Interesse, mit deren Hilfe Prognosen über den Arbeitgeberwechsel eines Mitarbeiters (Likely-Leaver-Profile) oder zukünftige Leistungen getroffen werden können. Möchte man beispielsweise untersuchen, wann und warum Mitarbeiter kündigen, können möglicherweise Daten über das vergangene Verhalten (Anzahl der Arbeitgeberwechsel), aber auch Daten aus dem Umfeld des Mitarbeiters hilfreich sein (siehe Tabelle 3-1).

Tabelle 3-1: Datentypen zur Messung des Verhaltens

Darstellung inklusive eines Beispiels für die Vorhersage der Wechselbereitschaft von Mitarbeitern

Datentyp	Beschreibung
Primäre Verhaltensdaten	Informationen über vergangenes Verhalten der gleichen Art, wie dasjenige, welches vorhergesagt werden soll, z. B. Anzahl und Frequenz früherer Arbeitgeberwechsel
Sekundäre Verhaltensdaten	Informationen über vergangenes Verhalten ähnlich zu dem, welches vorhergesagt werden soll, z. B. Anzahl und Frequenz von Wohnortwechsel
Tertiäre Verhaltensdaten	Informationen über vergangenes Verhalten, welches keinen direkten Bezug zu demjenigen hat, welches vorhergesagt werden soll, z. B. Kündigung des Fitnessclubs
Geo-demographische Daten	Informationen über die gegenwärtige Situation der Person, z. B. Alter der Person, Familienstand
Daten von Bezugspersonen	Informationen über Personen, zu denen eine Beziehung besteht, z. B. Erwerbsstatus des Partners
Einstellungsdaten (Gefühle und Meinungen)	Informationen über die Einstellung der Person zu bestimmten Sachverhalten oder die Gefühlslage, z. B. Angaben zur Arbeitszufriedenheit
Netzwerkdaten	Informationen über das Netzwerk an Beziehungen, in dem sich die Person befindet, z. B. Verortung des sozialen Umfelds

Quelle: in Anlehnung an Finlay, 2014, 68

Um (Früh-) Indikatoren für die Wechselbereitschaft von Mitarbeitern zu sammeln, ist die Frage zu beantworten, welche Anzeichen im Vorfeld oder zusammen mit dem Arbeitgeberwechsel auftreten. Dafür ist ein kausaler Zusammenhang zwischen dem Indikator und der vorherzusagenden Handlung nicht zwingend notwendig.

Die Gefühlslage und Stimmung ihrer Beschäftigten erheben Unternehmen häufig über Mitarbeiterbefragungen, die in der Regel alle zwei oder drei Jahre durchgeführt werden. Mit kurzen Onlinefragebögen zur wöchentlichen Messung der Arbeitszufriedenheit (z. B. über kununu), lässt sich das Intervall deutlich verkürzen, um auch aktuellere Stimmungsbilder aus der Belegschaft zu erhalten. Auch Texte und die gesprochene Sprache können mit Sentiment Analysen, einer Methode des Text Mining, analysiert werden, um durch positive oder negative Konnotationen der Äußerungen auf die zugrundeliegende Haltung des Senders Rückschlüsse ziehen zu können (Kornwachs, 2018, 123; Marr, 2018, 76). Informationsströme innerhalb des Unternehmens können über Netzwerkdaten nachvollzogen werden, die auch Rückschlüsse auf die Zusammenarbeit in Teams und die Identifikation von Kommunikationsbarrieren ermöglichen.

Entwicklungen in der Text-, Sprach-, Bild- und Videoanalyse haben den Kreis der analysierbaren Daten in den letzten Jahrzehnten stark erweitert. Rein aus technischer Sicht bieten die E-Mailkommunikation und Äußerungen der Mitarbeiter in sozialen Medien viele Anhaltspunkte über ihre Zufriedenheit mit der Arbeit. Auch Bilder und Videos können Aufschluss über die Gemütslage und das Wesen einer Person geben, aber auch mit Blick auf die Sicherheit und den Arbeitsschutz verwendet werden. So könnte man beispielsweise per Video automatisch erfassen, wenn ein Mitarbeiter die Helmpflicht auf der Baustelle verletzt und in Echtzeit ein Warnsignal abgeben, um schweren Unfällen vorzubeugen (Marr, 2018, 75 f.). Mit Ausweiskarten, die Sensoren oder auch GPS enthalten, sogenannten Smart-ID-Batches, lassen sich zudem Bewegungs- und Kommunikationsprofile erstellen. Mithilfe der vom MIT entwickelten Sociometric Badges werden beispielsweise die Anzahl an Face-to-Face Interaktionen, Gesprächszeiten und sprachliche Merkmale wie Intonation und Sprechrhythmus erfasst, um das Sozialverhalten von einzelnen Personen in Gruppen messbar zu machen (Waber, 2013, 14 ff.). Die Anwendung solcher „intelligenten Ausweiskarten“ ist für das Beschäftigungsverhältnis jedoch aus verschiedenen rechtlichen und ethischen Erwägungen kritisch zu sehen. So verbietet das deutsche Gesetz die Totalüberwachung von Beschäftigten. Eine Opt-in Regelung, die die Anwendung durch Freiwilligkeit legitimiert, wäre denkbar, allerdings müsste die Anonymität der Entscheidung für oder gegen das Tragen der Smart-ID-Batches durch Attrappen gewahrt bleiben. Die rechtlichen Rahmenbedingungen und auch der rechtliche Diskurs über die Freiwilligkeit im Arbeitsverhältnis finden sich in Kapitel 5. Neben rechtlichen und ethischen Bedenken ist auch der Informationsgehalt einiger Datentypen kritisch zu sehen.

Wie informativ Daten sind, ist im Vorfeld der Analyse nicht immer leicht abzuschätzen. Auch scheinbar irrelevante Daten können Rückschlüsse auf Personen zulassen. Dies zeigt eine Studie von Kosinski et al. (2013). Anhand von wenigen Informationen darüber, welche Inhalte eine Person auf Facebook gelikt hat, konnten die Autoren auf das Geschlecht, die sexuelle Orientierung, das Alter und die politische Gesinnung der Person schließen. Ein weiteres Beispiel für problematische Rückschlüsse von verfügbaren auf sensible Personendaten ist der Fall von

Target. Als einer der größten Einzelhändler in den USA untersuchte Target, welche Produkte werdende Mütter im frühen Stadium ihrer Schwangerschaft kauften. Ziel des Unternehmens war es, in einem möglichst frühen Stadium der Schwangerschaft entsprechende Kaufangebote für Babybedarf machen zu können. Dies funktionierte so gut, dass die Strategie bekannt wurde, als sich ein Vater über die unangemessene Werbung für Schwangere an seine Teenager-Tochter beschwerte, nur um kurz darauf festzustellen, dass das Unternehmen mehr über seine Tochter wusste als er selbst (Hill, 2012; Waber, 2013, 3). Diese Schlagzeile zeigt die Macht, aber auch die Gefahr von Daten, Privates offenzulegen und zu kommerzialisieren. Für die Verwendung von personenbezogenen Daten im Arbeitskontext ergibt sich aus dieser Erkenntnis eine besondere Verpflichtung zum sensiblen Umgang bei der Verknüpfung von Datensätzen.

3.3 Analyse und Interpretation

Die Vorgehensweise in der Analyse und die Auswahl der Daten unterscheiden sich je nach gesetztem Ziel des Vorhabens. Grob können drei Analyseformen unterschieden werden (siehe Tabelle 3-2). Für die Beschreibung von Zuständen und Entwicklungen in der Vergangenheit eignen sich deskriptive Analysen, die aggregierte Daten zu einem bestimmten Stichtag abbilden. Für Aussagen über zukünftige Entwicklungen braucht es dagegen Informationen über die Zusammenhänge zwischen Indikatoren und der zu prognostizierenden Entwicklung (siehe Beispiel aus Tabelle 3-1).

Tabelle 3-2: Analyseformen

	Deskriptive Analyse	Prädiktive Analyse	Präskriptive Analyse
Ziele /Fragestellungen	Analyse, die Fakten und Daten zu einem bestimmten Zeitpunkt oder im Zeitablauf beschreiben.	Vorhersage was, wann und vor allem auch warum passieren wird.	Empfehlungen, was zur Erreichung eines vorgegebenen Ziels unter vorliegenden Bedingungen zu tun ist.
Methoden	Reports, Dashboards, Scorecards	Data Mining, Text Mining, Web Media Mining, Forecasting	Optimierung, Simulation, Decision Modeling, Expertensysteme

Quelle: in Anlehnung an Kornwachs, 2018, 124

Die höchste Anforderung an die Datenstruktur und die Analyseverfahren stellt die präskriptive Analyse, in der es um die Aufdeckung von kausalen Zusammenhängen geht, um daraus Handlungsempfehlungen ableiten zu können. Möchte man beispielsweise herausfinden, ob eine Maßnahme des Gesundheitsmanagements den Krankenstand reduziert, reicht es nicht zu wissen, ob nach Einführung der Maßnahme der Krankenstand gesunken ist. So wäre nicht auszuschließen, dass Veränderungen im Krankenstand der Belegschaft auf andere Gründe wie

beispielsweise saisonale Effekte zurückzuführen sind, die nichts mit der gesundheitsförderlichen Maßnahme zu tun haben. Für einen kausalen Einfluss einer Variable X (Gesundheitsmaßnahme) auf eine Variable Y (Krankenstand in der Belegschaft) müssen drei Bedingungen gelten (Cheng, 2017, 5):

- i) X muss zeitlich vor Y liegen, also die Maßnahme muss der Periode gemessener Krankheitstage vorangehen.
- ii) X und Y müssen korreliert sein, das heißt der Krankenstand und die Gesundheitsmaßnahme müssen miteinander in einem strukturellen Zusammenhang stehen, der nicht rein zufällig auf die Stichprobenauswahl zurückzuführen ist.
- iii) Alternative Erklärungen für den Zusammenhang von X und Y müssen ausgeschlossen werden, inklusive umgekehrte Kausalität, was bedeutet, dass die Gesundheitsmaßnahme nicht ursächlich für den sich veränderten Krankenstand war, sondern als Folge dessen eingeführt wurde.

Für die Analyse kausaler Zusammenhänge kann man im Idealfall auf randomisierte Feldexperimente zurückgreifen, in denen Beschäftigtengruppen zufällig einer Treatment- und einer Kontrollgruppe zugeordnet werden (zu randomized controlled trials siehe auch Biemann et al., 2017, 9 ff.). Durch den Vergleich der Krankheitstage von Mitarbeitern, die an der Gesundheitsmaßnahme teilgenommen haben (Treatmentgruppe) und denen, die nicht teilgenommen haben (Kontrollgruppe), können Rückschlüsse auf den Effekt der Maßnahme getroffen werden. Kontrollierte Experimente sind außerhalb von Labortests jedoch kaum anzutreffen, da die zufällige Auswahl der Mitarbeiter in die ein oder andere Gruppe nur schwer umzusetzen ist. Oftmals werden Mitarbeiter für Pilotierungsphasen eher nach Standorten oder auf freiwilliger Basis ausgewählt. Fehlt die zufällige Auswahl der Mitarbeiter, ist die Analyse allerdings möglicherweise verzerrt, da Mitarbeiter an verschiedenen Standorten auch unterschiedlichen Einflüssen auf ihre Gesundheit ausgesetzt sein können. Um trotz fehlender Randomisierung einen kausalen Rückschluss treffen zu können, bieten sich – bestimmte Annahmen vorausgesetzt – eine Reihe von statistischen Verfahren wie Difference-in-Difference Schätzungen, Regressions-Diskontinuitäts-Analysen oder Instrumentenschätzer an, um Ursache-Wirkungsbeziehungen zu untersuchen (Cheng, 2017, 5 f.). Kausale Analysen sind damit um einiges aufwändiger als Analysen reiner Korrelationen und brauchen ein tiefergehendes Verständnis ökonomischer Verfahren.

Evidenzbasierte Entscheidungsfindung birgt das Risiko, sich auf die empirischen Ergebnisse zu verlassen, ohne die zugrundeliegenden Annahmen und Besonderheiten der Daten, wie relevante unbeobachtete Merkmale, zu berücksichtigen. So führte die Verwendung von Kapazitätsplanungstools im Einzelhandel laut einer Studie von Angrave et al. (2016) zu einer einseitigen Fokussierung auf die Reduktion der Personalkosten (Angrave et al., 2016, 7; Ton, 2009). Die darauffolgenden Qualitätsverluste und die damit verbundenen Umsatzeinbußen durch ausbleibende Kunden schaden den Unternehmen langfristig und negierten den vermeintlichen Vorteil durch die Kosteneinsparungen im Personal. Eine weitere Gefahr ist, dass die Analysemodelle so gewählt werden, dass die empirischen die eigenen Vermutungen bestärken und andersartige Erkenntnisse schlicht ignoriert werden. In der Debatte über den Nutzen von Big Data führen Boyd und Crawford sechs allgemeine Kritikpunkte zusammen, die sich als

Fallstricke auch mit Blick auf die Durchführung von People Analytics erweisen können (Boyd/Crawford, 2012, 665 ff.; Scholz, 2017, 25 ff.):

1. Big Data begünstigt einen Forschungsansatz, der explorativ vorgeht, ohne der Empirie eine Hypothese voranzustellen. Die fehlende theoretische Grundlage erhöht die Chance, dass rein zufällige Muster überbewertet werden.
2. Aufgrund der Menge an verfügbaren Daten und der Berücksichtigung von verschiedenen Quellen wird der Analyse irrtümlich per se eine größere Objektivität und Genauigkeit zugesprochen. Auch die statistische Unsicherheit findet nur selten Beachtung in der Interpretation der Ergebnisse.
3. Mehr Daten sind nicht immer besser, denn auch große Datensätze, die nicht die gesamte Grundgesamtheit abbilden, sind möglicherweise selektiv. Verzerrungen in den Schätzungen ergeben sich insbesondere, wenn die Verfügbarkeit oder Auswahl der Daten nicht zufällig sind, sondern Teilgruppen der zu beobachtenden Grundgesamtheit systematisch ausgeschlossen werden.
4. Für die Interpretation der Daten ist der Kontext wichtig, also die Angaben, wer die Daten wann und wie erhoben hat. Im Transformations- und Merging-Prozess von Datenquellen gehen diese Informationen häufig verloren und Beobachtungen können nicht kontextbezogen eingeordnet werden.
5. Die Menge an verfügbaren Daten wächst rasant, doch nicht alle Daten, die verfügbar sind, sollten nach ethischen Maßstäben auch genutzt werden (Hauser et al., 4 ff.). Dies gilt besonders für personenbezogene Daten.
6. Die steigende Bedeutung der Datenanalyse verschiebt das Machtverhältnis zuungunsten derer, die keinen Zugang zu diesen Daten haben bzw. nicht die nötigen Fähigkeiten besitzen, diese auszuwerten oder Analyseergebnisse zu hinterfragen.

Weiterhin zeigt eine Studie von Silberzahn et al. (2018), dass nicht nur die Auswahl des Datensatzes, sondern auch die verwendete Analysemethodik die Ergebnisse beeinflussen kann. In der Studie untersuchten 29 Forscherteams Diskriminierung im Fußball anhand der Fragestellung, ob die Hautfarbe eines Sportlers einen Einfluss auf die Wahrscheinlichkeit hat, eine rote Karte zu bekommen. Obwohl der jeweils gleiche Datensatz verwendet wurde, kamen die Forscher je nach verwendetem Forschungsansatz zu unterschiedlichen Ergebnissen. Statistische Befunde unterliegen zumindest in Teilen immer auch subjektiven Entscheidungen und normativen Annahmen. Gerade für die Analyse, die als Entscheidungsgrundlage dienen soll, ist daher die Güte der Schätzung und die Plausibilität der getroffenen Annahmen zu prüfen. Oftmals fehlen den Entscheidern und auch der HR-Abteilung allerdings die statistischen Kenntnisse, People Analytics in geeigneter Form durchzuführen und die Ergebnisse vor diesem Hintergrund zu interpretieren. Dies mag sich durch die steigende Bedeutung von Daten in Unternehmen langfristig wandeln, bedeutet aber in der gegenwärtigen Debatte, dass es für eine Durchführung und kritische Bewertung der Ergebnisse oftmals externe Partner braucht, die die Ergebnisse vor dem Hintergrund wissenschaftlicher Erkenntnisse einordnen. Ein verantwortungsvoller Umgang mit People Analytics bedeutet somit auch die Entmystifizierung „*[of the] [...] aura of truth, objectivity, and accuracy [...]*“ (Boyd/Crawford, 2012, 664), welche der aktuellen Debatte um Big Data anhaftet.

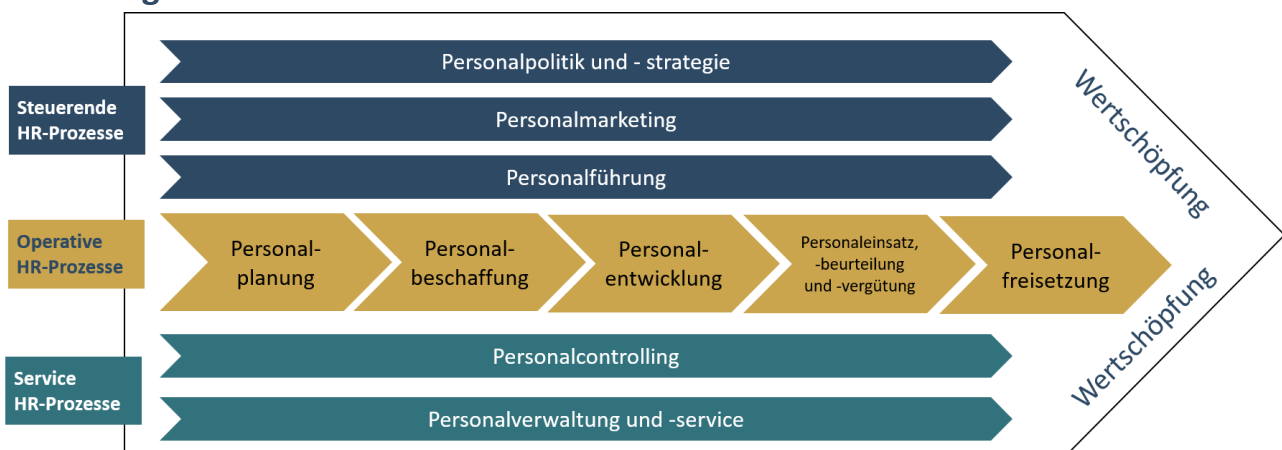
4 Anwendungsfelder im HR

„HR analytics CoE's [Center-of-Expertise] will often use big data to discover insights that they will “push” out to the business. This is a bit like shooting a gun in the air and hoping a bird flies over.“

(Rasmussen/Ulrich, 2015, 237)

Bei der Auswahl des Analysegegenstands sind ökonomische Erwägungen wichtig, die den Aufwand in Relation zu möglichen Effizienzgewinnen stellen. Um den Nutzen eines People Analytics Projekts abschätzen zu können, braucht es eine enge Zusammenarbeit verschiedener Fachbereiche im Betrieb und die Veränderungsbereitschaft der betroffenen Geschäftsbereiche. Die Fragestellungen von People Analytics können an den verschiedenen HR-Prozessen ansetzen oder auch prozessübergreifend laufen. Von der Planung und Beschaffung über die Mitarbeiterentwicklung, den Einsatz, die Beurteilung und Vergütung bis hin zur Personalfreisetzung bilden die HR-Prozesse die gesamte berufliche Laufbahn der Beschäftigten im Betrieb ab (siehe Abbildung 4-1). Begleitend laufen die Prozesse des Personalcontrollings und die Verwaltungs- und Serviceaktivitäten, wie die Gehaltsabrechnungen und die Pflege der Personalakte. Übergeordnet steuert die Personalstrategie die langfristige Ausrichtung der HR-Aktivitäten. Das Personalmarketing ist für die Gestaltung der Arbeitgebermarke und damit für den Aufbau und die gezielte Steuerung der öffentlichen Wahrnehmung des Unternehmens zuständig. Die Personalführung umfasst die Einbindung und Leitung von Mitarbeitern und Führungskräften anhand der unternehmerischen Ziele. In der Literatur zeichnen sich Analyseschwerpunkte in den Bereichen Personalbeschaffung, Personalentwicklung und Personalführung ab. Im Folgenden sollen mögliche Fragestellungen in diesen Bereichen näher untersucht werden.

Abbildung 4-1: HR-Prozesslandkarte



Quelle: Petry/Jäger, 2018, 62

4.1 Schwerpunkt Personalbeschaffung

Eine der wichtigsten Aufgaben des HR-Bereichs ist es, potenzielle Mitarbeiter für das Unternehmen zu interessieren und die richtigen Bewerber aus dem Bewerberpool auszuwählen. Die Personalauswahl eignet sich besonders für statistische Analysen, da viele Daten verfügbar sind und die getroffenen Entscheidungen langfristige Wirkung entfalten und mit hohen Kosten verbunden sind. Darüber hinaus hat der Faktor Arbeit durch die zunehmende Wissensarbeit relativ an Bedeutung für den Unternehmenserfolg gewonnen. Für die Personalbeschaffung ergeben sich eine Reihe interessanter Fragestellungen, denen mithilfe von Daten aus dem Bewerberprozess, Angaben über die betriebliche Laufbahn von Mitarbeitern sowie Daten aus betriebsexternen Quellen (z. B. über Plattformen oder Social Media) auf den Grund gegangen werden kann:

- Was macht das Unternehmen als Arbeitgeber bei bestimmten Zielgruppen attraktiv?
- Wie wird die Arbeitgebermarke wahrgenommen?
- Welche Rekrutierungskanäle sind am erfolgreichsten?
- Welche Kernkompetenzen sind für die Stellenbesetzung ausschlaggebend?
- Welcher Bewerber passt am besten zur Stelle, ins Team und zum Unternehmen?
- Wer eignet sich für Führungsverantwortung und sollte befördert werden?

Mit Methoden der Eignungsdiagnostik werden standardisierte Verfahren bereits seit Jahrzehnten im Personalauswahlprozess angewendet, um die Eignung der Bewerber bestimmen und mehrere Bewerber miteinander vergleichen zu können. Dazu sind in der DIN-Norm 33430 zur Eignungsdiagnostik Standards definiert, die es zu beachten gilt, um die Güte der Testverfahren sicherzustellen. So müssen die verwendeten Kriterien objektiv, reliabel und valide sein, das heißt, die Ergebnisse des Tests dürfen nicht von dem jeweiligen Untersuchungsleiter oder dem Zeitpunkt der Untersuchung abhängen, sie müssen zuverlässig messen (replizierbar sein) und sie müssen das zu messende Kriterium (berufliche Eignung) prognostizieren.

Bei People Analytics zu Fragen der Personalauswahl, die mit traditionellen Verfahren der Eignungsdiagnostik konkurrieren, sind die Gütekriterien häufig nicht transparent und es fehlen Standards. Anbieter wie das Aachener Unternehmen PRECIRE Technologies bieten betriebliche Lösungen, um Bewerber mittels von Sprachanalysen auf ihre Persönlichkeit und ihre Passgenauigkeit für das Unternehmen zu prüfen. Wie genau das Telefon-Interview mit dem Bewerber ausgewertet wird, um das Persönlichkeitsprofil zu erstellen und inwiefern beispielsweise auch Betonungen, Stimmlage und Wortschatz eine Rolle spielen, ist von außen nicht festzustellen. Bei Experten der Eignungsdiagnostik werden diese Analysen kritisch gesehen, nicht zuletzt, weil Persönlichkeitsmessungen häufig nur eine geringe Aussagekraft über berufsrelevante Aspekte wie die Leistungen der Bewerber haben (Müller, 2019). Als stärkste Prädiktoren gelten vielmehr kognitive Leistungstests und Arbeitsproben (Atabaki/Biemann, 2018, 131). Unternehmen wie PRECIRE hingegen versprechen eine Prüfung des Bewerbers, die über fachliche und intellektuelle Kompetenz hinausgeht, also ob ein Bewerber zum Unternehmen und ins Team passt. Ob die verwendete Analysemethode dieses Versprechen halten kann, ist fraglich.

Datenbasierte Entscheidungshilfen in der Personalauswahl werden häufig mit dem Ziel angewendet, mögliche subjektive Verzerrungen zu reduzieren. Stereotypes Denken – sogenannte unconscious bias – führt beispielsweise dazu, dass ein Urteil über eine Person aufgrund einer Gruppenzugehörigkeit getroffen wird. Weiterhin führt der halo effect dazu, dass einzelne herausragende Merkmale einer Person andere Eigenschaften überstrahlen und damit das Gesamturteil dominieren. Auch Merkmale, die denen des Entscheiders ähneln, können die Auswahlentscheidung beeinflussen. Dies ist auch als similarity attraction bias bekannt. Die Liste von irrationalen Entscheidungsparametern ließe sich noch erweitern.

Laut einer Befragung von knapp 200 Unternehmen der Top 1.000 und Top 300 IT-Unternehmen glauben über 60 Prozent, dass eine automatisierte Vorauswahl von Bewerbungen eine diskriminierungsfreie Bewerbervorauswahl fördert (Weitzel et al., 2019). Dass ein Algorithmus Verzerrungen in der Personalauswahl aber nicht notwendigerweise heilen muss, zeigt das Beispiel von Amazon (The Guardian, 2018). Mithilfe eines Algorithmus untersuchte ein Recruiting Team bei Amazon, welche Merkmale im Lebenslauf auf ein Top Talent hindeuten. Für Jobs als Softwareentwickler und andere Stellen mit hohem Anteil an technischen Stellen führte die Auswahl über den Algorithmus zu einer Verzerrung zuungunsten von weiblichen Bewerbern. Dies lag an den Trainingsdaten, die aus Lebensläufen von vor allem männlichen Kandidaten der vergangenen zehn Jahre gewonnen wurden. Dieses Beispiel zeigt, dass auch Algorithmen Muster lernen und daraufhin Prognosen treffen, die Verzerrungen unterliegen und damit im Widerspruch zu Grundsätzen der Gleichbehandlung stehen können (vgl. Kapitel 5.2).

4.2 Schwerpunkt: Personalentwicklung und -führung

Ein weiterer Schwerpunkt von People Analytics ist die Entwicklung und Motivation von Mitarbeitern. Die Arbeitszufriedenheit ist das Ergebnis der individuellen Bewertung und Gewichtung verschiedener Aspekte der Arbeit, die monetärer (z. B. Einkommen), materieller (z. B. Büroausstattung) und immaterieller (z. B. Betriebsklima) Natur sein können (Hammermann/Stettes, 2013). Die Arbeitszufriedenheit hängt wiederum mit der Leistungs- und Veränderungsbereitschaft der Beschäftigten zusammen (Shah et al., 2017), die gemeinsam mit den Fähigkeiten, Fertigkeiten und dem Wissen der Beschäftigten durch die Personalentwicklung und -führung gestärkt und gestaltet werden können. Voraussetzung für eine gute Personalentwicklung ist es, Kompetenzanforderungen für bestimmte Positionen zu definieren und diese mit den vorhandenen Kompetenzen der Beschäftigten optimal zusammenzubringen bzw. frühzeitig Weiterbildungsbedarfe zu erfassen. Neben den vorhandenen Kompetenzen und Leistungen ist auch die Abschätzung des Potenzials der Mitarbeiter für zukünftige Aufgaben wie beispielsweise Führungsverantwortung wichtig. In diesem Zusammenhang stellt sich eine Reihe von Fragen, die anhand von Mitarbeiterdaten untersucht werden können:

- Was macht Mitarbeiter zufrieden?
- Wie lässt sich die Veränderungs- und Leistungsbereitschaft der Beschäftigten erhöhen?
- Worin liegen die Gründe für Mitarbeiterkündigungen?
- Was macht gute Führung aus?
- Welche Kompetenzen der Beschäftigten werden wo gebraucht?
- Wie lassen sich die Kompetenzen am besten aufbauen und erhalten?

Führungskräfte sind mit ihrer Leistungsbeurteilung und der Delegation von Aufgaben und Verantwortungen wichtige Gestalter der Personalentwicklung und haben gleichzeitig auch einen hohen Einfluss auf die Mitarbeiterbindung. Hohe Fluktuationsraten sind nicht selten Ausdruck von enttäuschten Erwartungen der Mitarbeiter hinsichtlich der Anerkennung ihrer Leistungen oder dem fehlenden Gestaltungsspielraum. In ihrer Oxygen-Studie hat Google die Bedeutung von Führung untersucht (Dokuyucu, 2016; Sullivan, 2013). Prasad Setty, Leiter der Abteilung People Analytics bei Google, ging der Frage nach, ob Führungskräfte einen messbaren Einfluss auf den Jobwechsel von Mitarbeitern haben. Dafür untersuchte das Team zuerst die Daten von Mitarbeitern, die das Unternehmen verlassen hatten und fand heraus, dass diese häufig mit den Führungskräften unzufrieden waren. Dennoch wiesen diese Führungskräfte keine schlechten Bewertungen anhand der Mitarbeiterbefragung auf. Daraufhin führte das Team von Setty Interviews durch und zog auch Daten aus dem „Great Manager Award“ heran, um herauszufinden, was eine gute Führungskraft bei Google ausmacht. Als Ergebnis wurden acht Verhaltensweisen identifiziert. Dazu gehört, dass Führungskräfte Mitarbeiter zum eigenverantwortlichen Arbeiten bestärken, statt Micromanagement zu betreiben, sowie eine klare Strategie und Fachkompetenz der Führungskräfte, um den Mitarbeitern beratend zur Seite stehen zu können. Basierend auf diesen Erkenntnissen hat Google das 360-Grad-Feedback für Führungskräfte ausgerichtet und auch die Auswahl und Entwicklung von Führungskräften angepasst (Dokuyucu, 2016; Marr, 2018, 91 f.).

5 Rechtliche und ethische Betrachtung

„I shall try not to use statistics as a drunken man uses lamp-posts, for support rather than for illumination.“

Zitat des schottischen Schriftstellers Andrew Lang, 1937
(Rasmussen/Ulrich, 2015, 238)

5.1 Entscheidungsfindung

Nach der philosophischen Handlungstheorie lässt sich aus der Erkenntnis über einen kausalen Zusammenhang eine Handlungsempfehlung ableiten, die der Logik des pragmatischen Syllogismus folgt. Dieser besagt, dass, „ [...] wenn man einen kausalen Zusammenhang kennt, z. B. dass die Ursache A den Effekt B zur Folge hat, dass man dann gut beraten ist, wenn man B als Zustand möchte, eine geeignete Handlung \mathcal{A} , die A ins Werk setzt, auszuführen und damit diesen kausalen Zusammenhang auszunutzen (Kornwachs, 2018, 83 f.).“ Die so getroffene Entscheidung folgt damit einer logischen Regel und nutzt die aus den Daten destillierte Information zu der vorgegebenen Forschungsfrage. Sprechen die Ergebnisse der Kausalanalyse beispielsweise dafür, dass eine Maßnahme X den Krankenstand in der Belegschaft reduziert, wäre der logische Schluss, die Maßnahme fortzuführen. Dies ist zwar eine plausible, aber nicht notwendigerweise die Entscheidung, die auch in Zukunft zum Erfolg führt. Zu beachten ist, dass der statistische Zusammenhang unter sonst gleichen Bedingungen besteht (ceteris paribus Regel). Ändern sich Einflussgrößen, z. B. durch eine sich ändernde Belegschaftsstruktur oder andere Arbeitsbelastungen, kann die Entscheidung, die Maßnahme fortzuführen, trotzdem „falsch“ sein.

Insbesondere wenn die zugrundeliegenden Algorithmen nur schwer nachvollziehbar sind, wird sich die Person, die die Entscheidung trifft, womöglich unreflektiert auf die Empfehlung verlassen. People Analytics laufen dann Gefahr, Entscheidungen nicht nur zu unterstützen, sondern zu ersetzen. Die datengestützte Entscheidungsfindung lässt sich in drei Stufen gliedern, deren Grenzen jedoch in der Praxis fließend sind (Kornwachs, 2018, 174 ff.):

1. **Unterstützung von Entscheidungen:** Berechnungen bieten anhand von vorgegebenen Kriterien und Bewertungsalgorithmen eine Entscheidungsgrundlage.
2. **Ersetzung von Entscheidungen:** Übernimmt der Mensch die vorgeschlagene Alternative ohne eine weitere Prüfung und Erörterung, z. B. aufgrund von guten Erfahrungen mit vorherigen Lösungsvorschlägen, tritt der Algorithmus de facto an Stelle der menschlichen Entscheidung.
3. **Automatisierung von Entscheidungen:** Eine Exekution der Lösung ohne menschliches Zutun oder Eingreifen.

Entscheidungen sind dann vollständig automatisierbar, wenn es ein eindeutiges Nutzenmaximum gibt. So berechnet das Bayessche Modell, als Grundmodell der Entscheidungstheorie (Kornwachs, 2018, 175), den Erwartungswert des Nutzens U einer jeden Alternative $f_j=1,\dots,m$ über alle Umstände w_i mit $i=1,\dots,n$, die mit einer gewissen Wahrscheinlichkeit $p(w_i)$ eintreten

und im Realisierungsfall eine von den jeweiligen Umständen abhängende Folge c_{ij} = Funktion (w_i, f_j) auslösen würde. Jede dieser Folgen muss mit der Bewertung $V(c_{ij})$ bewertet werden.

$$U(f) = \sum_i V(c_{ij}) * p(w_i) \rightarrow \max$$

Die Bestimmung des Nutzes einer Folge $V(c_{ij})$ hängt von individuellen oder betrieblich festgelegten Werten, Kriterien und Indikatoren ab, wie z. B. stellenspezifischen Kompetenzmodelle bei der Personalauswahl. Algorithmen des maschinellen Lernens leiten diese aus einer Datengrundlage (Trainingsmasse an vergangenen Personalauswahlentscheidungen) ab. Die Nutzenoptimierung ist allerdings komplexer als Modelle basierend auf dem vollständig rational denkenden Menschen vermuten ließen. Die Spieltheorie und Behavioral Economics haben gezeigt, dass neben dem eigenen Nutzen eine Fülle an (sozialen) Präferenzen wie Reziprozität, Altruismus oder Ungleichheitsaversion und im Lebensverlauf stabile wie auch sich ändernde Handlungsparadigmen eine Rolle spielen können. Dies erhöht die Komplexität von Entscheidungen um ein Vielfaches und führt auch dazu, dass nutzenmaximierende Entscheidungen bei verschiedenen Personen (zu verschiedenen Zeitpunkten) sehr unterschiedlich ausfallen können. Ein weiterer entscheidender Aspekt ist, dass Algorithmen nicht die Verantwortung für die getroffenen Entscheidungen und die sich daraus ableitenden Folgen tragen können. Eine Delegation der Verantwortung an Maschinen ist unzulässig. Dies sieht auch der Gesetzgeber so, der automatisierte Entscheidungen, die für andere eine Rechtsfolge entwickeln, verbietet (Art. 22 Nr.1 DSGVO).

5.2 Gleichbehandlung

Nach dem Allgemeinen Gleichbehandlungsgesetz (AGG) sind Benachteiligungen aus Gründen der Rasse, ethnischen Herkunft, des Geschlechts, der Religion oder Weltanschauung, einer Behinderung, des Alters oder der sexuellen Identität im Arbeitskontext unzulässig. Dies ist für die Datenselektion und -auswertung wichtig, aber auch für die sich aus der Analyse ableitenden Implikationen. So wäre es beispielsweise möglich, dass Sprachanalysen Bewerber, die nicht Muttersprachler sind, im Auswahlverfahren benachteiligen. Auch die Bevorzugung von Bewerbern, die sich als Ausdauersportler auszeichnen, ist kritisch, da dadurch behinderte Menschen eine geringere Chance erhalten, eingestellt zu werden (vgl. Dzida, 2018, 2683).

People Analytics beruhen nicht auf Naturgesetzen, sondern auf normativen Grundsätzen: Was einen leistungsstarken Mitarbeiter ausmacht, muss festgelegt und an Messkriterien wie dem Arbeitstempo, der Kundenzufriedenheit oder dem Umsatz geknüpft werden. Die Festlegung in gewünschtes und ungewünschtes Verhalten ist die Basis von maschinellem Lernen (Kim, 2017, 876). Gleichbehandlungsgrundsätze müssen bereits in der Trainingsmasse selbstlernender Programme bzw. im Algorithmus selbst verankert sein, ansonsten können sich Diskriminierungstendenzen aus der Vergangenheit in die Zukunft fortschreiben. Systematische Verzerrungen können sich zudem ergeben, wenn der verfügbare Datensatz Teilgruppen aufgrund von fehlenden Beobachtungen nicht berücksichtigt. Ist die vorliegende Datenstruktur selektiv,

systematisch verzerrt oder gar manipuliert (z. B. durch Suggestivfragen in Fragebögen), sind die daraus resultierenden Zusammenhänge fragwürdig und nicht repräsentativ (Kim, 2017, 865).

Problematisch ist, wenn kausale Ableitungen auf Basis von Korrelationen getroffen werden. Damit besteht die Gefahr von sogenannten „Scheinkorrelationen“, also Zusammenhängen, die nur indirekt oder zufällig aufgrund von unvollständigen Daten zustande kommen. Ein bekanntes Beispiel für eine Scheinkorrelation ist der statistisch nachweisbare positive Zusammenhang der Geburtenraten und der Storchpopulation. Beides steht in keinem kausalen Zusammenhang, sondern hängt jeweils von der wirtschaftlichen Entwicklung der Regionen ab. Mit dem wirtschaftlichen Aufschwung einer Region steigen auch die Berufschancen von Frauen und damit sinkt die Geburtenrate. Gleichzeitig nehmen aber auch die Umweltbelastungen mit negativen Folgen für die Storchpopulation zu. Wenn der eigentliche Treiber beider Entwicklungen, in diesem Fall die Wirtschaftskraft der Region, nicht Teil der Analyse ist, entsteht ein statistischer Zusammenhang, dem keine Ursache-Wirkungsbeziehung zugrunde liegt. Unvollständige Informationen können zu verzerrten Ergebnissen und damit zu falschen Schlussfolgerungen führen.

Auch Personalverantwortliche verfügen nur über wenig Informationen, anhand derer sie im Einstellungsverfahren entscheiden müssen, ob der Bewerber passt oder nicht. Die Auswahl beruht daher oftmals auf dem Bauchgefühl der Entscheider. Aber nicht nur subjektive Beurteilungen unterliegen Verzerrungen: Werden statistische Unterschiede in der Verteilung von Merkmalen zwischen Gruppen (z. B. in der durchschnittlichen Produktivität) als Auswahlkriterium anstelle des häufig nicht beobachtbaren tatsächlichen Merkmals des Einzelnen herangezogen, spricht man von statistischer Diskriminierung (Arrow, 1971, 25 ff.). Statistische Analyseverfahren können zudem Stigmatisierung begünstigen, da Entscheidungsvorlagen auf mathematischen Verfahren beruhen, die Merkmalsunterschiede zwischen Gruppen identifizieren. Problematisch bei der Textanalyse können auch tradierte Festlegungen im Sprachgebrauch sein wie z. B. dominante männliche oder weibliche Formen von Berufsbezeichnungen (wie beispielsweise Krankenschwester oder Zimmermann).

Algorithmen sind also nicht per se diskriminierungsfrei, aber sie können Verzerrungen zu Ungunsten von Minderheiten aufdecken. Dies zeigt der Fall des Technologie- und Dienstleistungsunternehmens Xerox, die Evolv Solution damit beauftragten, den Recruitingprozess ihrer Call Center zu analysieren. Dabei kam heraus, dass der stärkste Prädiktor für eine gute Mitarbeiterbindung eine geringe Distanz des Wohnortes zum Unternehmensstandort war. Der Zusammenhang zwischen den Pendelzeiten und der Wahrscheinlichkeit, den Arbeitgeber zu wechseln, erschien plausibel, barg aber auch die Gefahr der Diskriminierung. In dem genannten Fall lag das Call Center im Stadtzentrum und Mitarbeiter aus den Randgebieten der Stadt gehörten häufiger ethnischen Minderheiten an. Als Auswahlkriterium für Bewerber wurde die Distanz zwischen Arbeit und Wohnort daher trotz guter statistischer Vorhersagekraft nicht weiter genutzt (Kim, 2017, 873).

5.3 Datenschutz

Rechtsgrundlagen für den Datenschutz in Deutschland sind die Europäische Datenschutzgrundverordnung (DSGVO) und das Bundesdatenschutzgesetz (BDSG). Die DSGVO formuliert europaweit einheitliche Datenschutzregelungen im Umgang mit personenbezogenen Daten und ersetzt die europäische Richtlinie aus dem Jahr 1995 „zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr“. Die DSGVO gilt seit dem 25. Mai 2018 unmittelbar in allen EU-Mitgliedstaaten und wird in Deutschland durch ein 2018 novelliertes BDSG konkretisiert sowie ergänzt, dort wo es Öffnungsklausel der DSGVO zulassen. Auch wenn People Analytics in der Regel darauf abzielen, statistische Zusammenhänge und Muster unabhängig von einzelnen Personen zu erkennen und zu erklären, sind hohe Datenschutzstandards zu beachten, die aufmerksam und sorgfältig umgesetzt werden müssen (vgl. Reindl/Krügl, 2017, 71 ff.).

BDSG und DSGVO beziehen sich auf personenbezogene Daten. Darunter werden alle Angaben verstanden, die eine natürliche Person identifizieren. Als „[...] identifizierbar wird eine natürliche Person angesehen, die direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einem Namen, zu einer Kennnummer, zu Standortdaten, zu einer Online-Kennung oder zu einem oder mehreren besonderen Merkmalen identifiziert werden kann, die Ausdruck der physischen, physiologischen, genetischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität dieser natürlichen Person sind“ (§ 46 Nr. 1 BDSG; Art. 4 Nr. 1 DSGVO).

Für die Datenverarbeitung zum Zweck des Beschäftigungsverhältnisses ist geregelt, dass personenbezogene Daten der Beschäftigten verarbeitet werden dürfen, [...] wenn *die Entscheidung über die Begründung eines Beschäftigungsverhältnisses oder nach Begründung des Beschäftigungsverhältnisses für dessen Durchführung oder Beendigung oder zur Ausübung oder Erfüllung der sich aus einem Gesetz oder einem Tarifvertrag, einer Betriebs- oder Dienstvereinbarung (Kollektivvereinbarung) ergebenden Rechte und Pflichten der Interessenvertretung der Beschäftigten erforderlich ist.*“ (§ 6 Nr. 1 BDSG). Die Verarbeitung von besonderen Kategorien personenbezogener Daten wie z. B. genetische oder biometrische Daten sind grundsätzlich untersagt und nur nach ausdrücklicher Einwilligung bezogen auf einen vorab festgelegten Zweck zulässig (Art. 9 DSGVO mit Ausnahmen nach § 22 sowie § 26 Nr. 3 BDSG beispielsweise zur Ausübung von Rechten und der Erfüllung rechtlicher Pflichten aus dem Arbeitsrecht).

Die Einwilligung zur Datenverarbeitung muss freiwillig, explizit und nachweislich sein und jederzeit widerrufen werden können (Art. 21 DSGVO). Aufgrund des Machtgefälles im Beschäftigungsverhältnis zwischen Arbeitgeber und Arbeitnehmer wird die Zulässigkeit einer freiwilligen Einwilligung immer wieder angezweifelt, auch wenn diese im Erwägungsgrund 155 DSGVO verankert ist (Dzida, 2018, 2682; Gola, 2015, 138). Verboten ist zudem eine Totalüberwachung, in der das Verhalten permanent aufgezeichnet wird. Weiterhin dürfen Entscheidungen, die eine rechtliche Wirkung entfalten oder in ähnlicher Weise eine betroffene Person beeinflussen, nicht ausschließlich auf einer automatisierten Verarbeitung beruhen (Art. 22 Nr. 1 DSGVO). Letztere Regelung wirft heute noch nicht eindeutig beantwortete rechtliche Fragen auf, inwieweit beispielsweise eine evidenzbasierte Grundlage zur Entgeltfindung oder auch

eine auf Scores beruhende Arbeitnehmerbeurteilung diese Regelung bereits verletzen (Maschmann, 2019, 636).

Zu den Grundsätzen für die Verarbeitung personenbezogener Daten gehören die Zweckbindung und die Datenminimierung, das bedeutet, dass die Erhebung auf das notwendige Maß an Daten beschränkt und dem Zweck angemessen sein muss (Art. 5 DSGVO). Unzulässig ist eine Datensammlung auf Vorrat. Die Definition des Zwecks der Datenerhebung ist jedoch zu Beginn eines People Analytics Projekts nicht immer einfach, da das Analyseinteresse häufig vage ist und sich im zeitlichen Ablauf wandeln kann. Eine Änderung des Zwecks bedarf jedoch einer erneuten Einwilligung der betroffenen Person. Weiterhin sind umfangreiche Informationspflichten zu beachten, um das Recht des Einzelnen, über die Preisgabe und Verwendung seiner personenbezogenen Daten selbst zu entscheiden, zu gewährleisten. Dieses Recht auf informationelle Selbstbestimmung hat das Bundesverfassungsgericht als Grundrecht anerkannt (Gola, 2015, 128).

Häufig reichen für die Untersuchung von Zusammenhängen im Rahmen der People Analytics anonymisierte Daten. Zu diesem Zweck können personenbezogene Daten „pseudoanonymisiert“ werden, sodass diese „[...] ohne Hinzuziehung zusätzlicher Informationen nicht mehr einer spezifischen betroffenen Person zugeordnet werden können, sofern diese zusätzlichen Informationen gesondert aufbewahrt werden und technischen und organisatorischen Maßnahmen unterliegen, die gewährleisten, dass die personenbezogenen Daten nicht einer identifizierten oder identifizierbaren natürlichen Person zugewiesen werden“ (Art. 4 Nr. 5 DSGVO). Auf anonymisierte oder aggregierte Daten findet das Datenschutzgesetz zunächst keine Anwendung, kann jedoch dann relevant werden, wenn die Auswertungsergebnisse auf Einzelne angewandt werden und daraus Entscheidungen resultieren, die das Arbeitsverhältnis betreffen, wie Kündigungen oder Beförderungen (Reindl/Krügl, 2017, 76 f.). Bei kleineren Einheiten aggregierter Daten (beispielsweise kleine Abteilungen) ist zudem zu prüfen, inwieweit eine Re-identifizierung möglich wäre.

5.4 Betriebliche Mitbestimmung

Nach § 75 Abs. 2 des Betriebsverfassungsgesetzes (BetrVG) sind Arbeitgeber und Betriebsrat gemeinsam dafür verantwortlich, „die freie Entfaltung der Persönlichkeit der im Betrieb beschäftigten Arbeitnehmer zu schützen und zu fördern“ und damit auch die Privatsphäre des Einzelnen zu garantieren. Analysen mit Mitarbeiterdaten sind daher grundsätzlich mitbestimmungspflichtig. Zur Wahrung der Aufgaben des Betriebsrates, insbesondere der Überwachungspflichten zur Einhaltung des Datenschutzes nach § 80 Abs. 1 Nr. 1 BetrVG, sind für People Analytics Unterrichtungspflichten und Mitbestimmungsrechte zu beachten. Dies gilt vor allem mit Blick auf die „Einführung und Anwendung von technischen Einrichtungen, die dazu bestimmt sind, das Verhalten oder die Leistung der Arbeitnehmer zu überwachen“ (§ 87 Abs. 1 Nr. 6 BetrVG). Für die Anwendung des Paragraphen ist es irrelevant, ob der Arbeitgeber die Technik zur Kontrolle der Mitarbeiter verwenden möchte. Es reicht, wenn diese dafür nach objektiven Maßstäben geeignet wäre. Eine Datenverarbeitung, der nicht durch die Mitbestimmung zugestimmt wurde, ist unrechtmäßig und die daraufhin ergangenen Maßnahmen und Anweisungen sind unwirksam (Gola, 2015, 147).

Des Weiteren sind je nach Analyseschwerpunkte auch Unterrichtungspflichten in Fragen der Personalplanung (§ 92 BetrVG) relevant. Personalfragebögen und Richtlinien über die personelle Auswahl bei Einstellungen, Versetzungen, Umgruppierungen und Kündigungen bedürfen der Zustimmung des Betriebsrates (§ 94 und § 95 BetrVG). Für die Förderung der Berufsbildung kann der Betriebsrat zudem verlangen, dass der Arbeitgeber den Berufsbildungsbedarf ermittelt und kann selbst beratend tätig werden (Vorschlagsrecht des Betriebsrates nach § 96 Abs. 1 BetrVG). Für eine verbesserte Erfassung des Qualifizierungsbedarfs, aber auch in vielfältigen anderen Anwendungsbereichen, sind gemeinsame Interessen der Arbeitgeber- und Betriebsratsseite denkbar, statistische Verfahren im Rahmen von People Analytics zur datengestützten Entscheidungsfindung einzusetzen. Eine vertrauensvolle Zusammenarbeit in enger Absprache miteinander ist eine wichtige Voraussetzung für die Nutzung der Potenziale, die sich durch die Datenanalyse ergeben können.

6 Handlungsfelder und Ausblick

Hinter dem Begriff People Analytics steckt mehr als eine Modeerscheinung. Datenbasierte Entscheidungshilfen können die HR-Arbeit professionalisieren und mithilfe von prädiktiver und präskriptiver Analyse stärker strategisch ausrichten. Die Datenbasis dafür wird immer umfangreicher, doch der teils verklärte Blick auf die Potenziale des neuen Trends in der Managementliteratur verstellt, dass es mit Ausnahmen von prominenten Einzelfallbeispielen kaum belastbare Empirie gibt, dass People Analytics in der Praxis auch tatsächlich erfolgreich angewendet werden. Damit aus dem Modebegriff eine wertschöpfende Praxis wird, muss eine Reihe an hemmenden Faktoren überwunden werden.

Zu diesen gehört in erster Linie mehr Rechtssicherheit im Umgang mit Daten. Durch die Datenschutzgrundverordnung ist die Unsicherheit im Umgang mit personenbezogenen Daten eher größer geworden. In regelmäßigen Abständen machen Themen zum Datenschutz Schlagzeilen, wie der Fall eines Bewohners in Wien, der sich wegen der Namensnennung auf Klingelschildern an Hauseingängen beklagte oder auch die aktuelle Debatte auf der jüngsten Innenministerkonferenz um die Verwendung von Daten der Assistenzsysteme Siri, Alexa und Co. zur Strafverfolgung. Gerade für kleine und mittlere Unternehmen wirkt die Rechtsunsicherheit und die Angst vor den hohen Strafen bei Verstößen gegen den Datenschutz abschreckend, ihr Datenpotenzial stärker zu nutzen.

Ein weiteres Hemmnis stellt die Emotionalität der geführten Debatte dar. Einzelne Fälle, in denen Mitarbeiterdaten missbräuchlich gesammelt und ausgewertet wurden, machen Schlagzeilen und bleiben in Erinnerung. Um nicht in den Verdacht derartiger Methoden zu geraten und das Misstrauen ihrer Belegschaft zu riskieren, sind Unternehmen daher häufig sehr vorsichtig in der Verwendung interner Datensätze. Auch wenn viele Mitarbeiter in ihrem Privatleben mit Daten eher sorglos umgehen, ist die Akzeptanz zur Datennutzung im Arbeitskontext häufig recht gering. Mehr Transparenz im Umgang mit den gespeicherten Daten und ein stärkeres Bekennen zu ethischen Werten der Datennutzung hinsichtlich des Schutzes der Privatsphäre, der Gleichbehandlung und Informationellen Selbstbestimmung (vgl. Hauser et al., 4 ff.) der betrieblichen Sozialpartner könnten helfen, mehr Vertrauen hierzulande in Datenanalysen zu entwickeln.

Als drittes Hemmnis sind die fehlenden Analysekenntnisse in den Unternehmen zu nennen. Zum einen steigen die Anforderungen an die Datenaufbereitung mit der zunehmenden Komplexität und Menge an Daten an. Zum anderen sind für die Gewinnung verwertbarer Informationen aus den Daten statisches Know-how und Kenntnisse über den wissenschaftlichen Forschungsstand nötig, um die Ergebnisse auch einordnen zu können. Eine Kooperation von Wissenschaft und Praxis bietet eine Möglichkeit, die nötigen Kompetenzen für die Analyse in die Betriebe zu tragen. Allerdings impliziert das Kompromissbereitschaft auf beiden Seiten, um sowohl den betrieblichen Zielen als auch den wissenschaftlichen Anforderungen gerecht werden zu können.

Abstract

The data volume is growing fast. In 2018 33 Zettabytes of data were generated and the amount is expected to grow to 44 Zettabytes in 2020. Only a fraction of data in companies is used for improving decision-making. This holds true specifically for the HR department. Decisions in recruiting and talent management are nowadays mainly based on intuition and experience rather than hard facts.

A more evidence-based approach for decision making in human resource departments is currently hyped in the management literature and picked-up in the academic research literature using the terms People Analytics and HR Analytics. However, assumed advantages like non-discriminating and objective decisions are seen more critically from a scientific point of view.

Our study analyses the risks and opportunities of People Analytics and its use of data sources and analytical approaches from an economic, legal and ethical perspective. We point out several pitfalls and obstacles which must be addressed thoughtfully before a People Analytics project is recommended. The three main obstacles in our regard are the lack of analytical skills, the reservation of data usage in workforces and legal uncertainty.

Tabellenverzeichnis

Tabelle 2-1: People Analytics versus Personalcontrolling	6
Tabelle 2-2: Cluster	9
Tabelle 3-1: Datentypen zur Messung des Verhaltens.....	12
Tabelle 3-2: Analyseformen.....	14

Abbildungsverzeichnis

Abbildung 2-1: Entwicklung der Datenorientierung im HR.....	5
Abbildung 2-2: Auswertung der wissenschaftlichen Fachliteratur	8
Abbildung 4-1: HR-Prozesslandkarte	17

7 Literaturverzeichnis

- AGG, Allgemeines Gleichbehandlungsgesetz. Ausfertigungsdatum: 14.08.2006; zuletzt geändert durch Art. 8 G v. 3.4.2013 I 610, <http://www.gesetze-im-internet.de/agg/> [22.8.2019]
- Angrave, David et al., 2016, HR and analytics: why HR is set to fail the big data challenge, in: Human Resource Management Journal, 26. Jg., Nr. 1, S. 1–11
- Aral, Sinan / Brynjolfsson, Erik / Wu, Lynn, 2012, Three-Way Complementarities: Performance Pay, Human Resource Analytics, and Information Technology, in: Management Science, 58. Jg., Nr. 5, S. 913–931
- Arrow, Kenneth, 1971, The theory of discrimination. Working Papers 403, Princeton University, Department of Economics, Industrial Relations Section, Princeton
- Atabaki, Armita / Biemann, Torsten, 2018, Potenziale der Datenanalyse für HR (People Analytics), in: Petry, Thorsten / Jäger, Wolfgang (Hrsg.), Digital HR. Smarte und agile Systeme, Prozesse und Strukturen im Personalmanagement, Freiburg, München, Stuttgart
- BDSG, Bundesdatenschutzgesetz. Gesetz zur Anpassung des Datenschutzrechts an die Verordnung (EU) 2016/679 und zur Umsetzung der Richtlinie (EU) 2016/680 (Datenschutz-Anpassungs- und -Umsetzungsgesetz EU- DSAnpUG-EU), <https://dsgvo-gesetz.de/bdsg/> 22.08.2019]
- Becker, Brian E. / Huselid, Mark A. / Ulrich, David, 2007, The HR scorecard. Linking people, strategy, and performance, Boston, Mass.
- Berufenet, 2019, Data Scientist – Die Tätigkeit im Überblick, <https://berufenet.arbeitsagentur.de/berufenet/faces/index?path=null/kurzbeschreibung&dkz=129987&such=Data+Scientist> [4.6.2019]
- BetrVG, Betriebsverfassungsgesetz. Betriebsverfassungsgesetz in der Fassung der Bekanntmachung vom 25. September 2001 (BGBl. I S. 2518), das zuletzt durch Artikel 4e des Gesetzes vom 18. Dezember 2018 (BGBl. I S. 2651) geändert worden ist, <https://www.gesetze-im-internet.de/betrvg/BetrVG.pdf> [22.8.2019]
- Bhattacharyya, Dipak Kumar, 2017, HR analytics. Understanding theories and applications, New Delhi, India, Thousand Oaks, California
- Biemann, Torsten / Englmaier, Florian / Sliwka, Dirk / Weller, Ingo, 2017, People Analytics – Personaldata als Erfolgsfaktor, in: Personalquarterly – Wissenschaftsjournal für die Personalpraxis, 69. Jg., Nr. 3, S. 3–15
- Borowczak, Mike / Yang, Richard, 2018, Predictive Liability Models and Visualizations of High Dimensional Retail Employee Data, ICAI '18 Proceedings of the 2nd International Conference on Innovation in Artificial Intelligence, S. 186-190
- Boudreau, John / Lawler, Edward E., 2015, Making Talent Analytics and Reporting Into a Decision Science, CEO Publication, G15-16(663), <https://ceo.usc.edu/making-talent-analytics-and-reporting-into-a-decision-science/> [11.6.2019]
- Boyd, Danah / Crawford, Kate, 2012, Critical questions for big data, in: Information, Communication & Society, 15. Jg., Nr. 5, S. 662–679

Cheng, Minghui, 2017, Causal Modeling in HR Analytics: A Practical Guide to Models, Pitfalls, and Suggestions, in: Academy of Management Annual Meeting Proceedings, 1. Jg., S. 1-6

Dokuyucu, Kenan, 2016, Analysis of Google's Project Oxygen, MGMT, Nr. 6351, University of Houston-Victoria

DSGVO, Datenschutz-Grundverordnung. Verordnung (EU) 2016/679 des europäischen Parlaments und des Rates zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG, <https://dsgvo-gesetz.de/> [22.8.2019]

Dzida, Boris, 2018, Der neue Beschäftigtendatenschutz – Erste Erfahrungen aus der Praxis, in: Betriebs-Berater, Nr. 45, S. 2677–2684

Engels, Barbara, 2018, Ein unbekannter Schatz – Wie bestimmen Unternehmen in Deutschland den Wert ihrer Daten?, in: IW-Trends, 45. Jg., Nr. 4, S. 41–59

Engels, Barbara / Goecke, Henry, 2019, Big Data in Wirtschaft und Forschung – Eine Bestandsaufnahme, IW-Analysen Nr. 130. Institut der deutschen Wirtschaft Köln Medien GmbH. Köln

Finlay, Steven, 2014, Predictive analytics, data mining and big data. Myths, misconceptions and methods, Houndmills, Basingstoke, Hampshire

Fitz-enz, Jac, 2001, How to measure human resources management, New York, London

Gola, Peter, 2015, HR Intelligence und Analytics – Datenschutzrechtliche Grenzziehungen, in: Strohmeier, Stefan / Piazza, Franca (Hrsg.), Human Resource Intelligence und Analytics, Springer Fachmedien Wiesbaden, S. 127–158

Hammermann, Andrea / Stettes, Oliver, 2013, Qualität der Arbeit - zum Einfluss der Arbeitsplatzmerkmale auf die Arbeitszufriedenheit im europäischen Vergleich, in: IW-Trends, Nr. 2, S. 93–109

Hauser, Christian / Blumer, Helene / Christen, Markus, / Hilty, Lorenz / Huppenbauer, Lorenz / Kaiser, Tony, 2017, Ethische Herausforderungen für Unternehmen im Umgang mit Big Data. Schweizerische Akademie der Technischen Wissenschaften, Zürich.

Heupel, Thomas / Lange, Victoria W., 2019, Wird der Controller zum Data Scientist? Herausforderungen und Chancen in Zeiten von Big Data, Predictive Analytics und Echtzeitverfügbarkeit, in: Hermeier, Burghard / Heupel, Thomas / Fichtner-Rosada, Sabine (Hrsg.), Arbeitswelten der Zukunft, Wiesbaden, S. 201–220

Hill, Kashmir, 2012, How Target figured out teen girl was pregnant before her father did, <https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#110ae2626668> [22.8.2019]

Holthaus, Christian / Park, Young-kul / Stock-Homburg, Ruth, 2015, People Analytics und Datenschutz–Ein Widerspruch?, in: Datenschutz und Datensicherheit – DuD, 39. Jg., Nr. 10, S. 676–681

Jain, Neetu / Maitri, 2018, Big Data and Predictive Analytics: A Facilitator for Talent Management, in: Munshi, Usha Mujoo / Verma, Neeta (Hrsg.), Studies in big data, Data science landscape. Towards research standards and protocols, Singapore, S. 199–204

- Kaplan, Robert S. / Norton, David P., 2009, The balanced scorecard. Translating strategy into action, Boston, Mass.
- Kim, Pauline, 2017, Data-Driven Discrimination at Work, in: William & Mary Law Review, 48. Jg., S. 857–936
- Kornwachs, Klaus, 2018, Arbeit 4.0 – People Analytics – Führungsinformationssysteme: Soziologische, psychologische, wissenschaftsphilosophisch–ethische Überlegungen zum Einsatz von Big Data in Personalmanagement und Personalführung, Gutachten für die Universität Münster, Vergabenummer 2017_59_BS. Büro für Kultur und Technik, Argenbühl-Eglofs
- Kosinski, Michal / Stillwell, David / Graepel, Thore, 2013, Privat traits and attributes are predictable from digital records of human behavior, in: PNAS, 110. Jg., Nr. 15, S. 5802–5805
- kununu, Online-Plattform für Arbeitgeberbewertungen, <https://engage.kununu.com/de/> [7.6.2019]
- Lewis, Michael, 2013, Moneyball. The art of winning an unfair game, New York
- Madsen, Dag / Slåtten, Kåre, 2017, The Rise of HR Analytics: A Preliminary Exploration, in: Conference: Global Conference on Business and Finance Proceedings, 12. Jg., Nr. 1, S. 148–159
- Marler, Janet H. / Boudreau, John W., 2017, An evidence-based review of HR Analytics, in: The International Journal of Human Resource Management, 28. Jg., Nr. 1, S. 3–26
- Marler, Janet / Cronemberger, Felipe / Tao, Carson, 2017, HR Analytics: Here to stay or short lived management fashion?, in: Bondarouk, Tanya / Ruël, Huub / Parry, Emma (Hrsg.), The changing context of managing people, Electronic HRM in the smart era, Bingley, UK, S. 59–86
- Marr, Bernard, 2018, Data-driven HR. How to use analytics and metrics to drive performance, New York
- Maschmann, Frank, 2019, Verarbeitung personenbezogener Entgeltdata und neuer Datenschutz, in: Betriebs-Berater, Nr. 11, S. 628–636
- Müller, Martin, 2019, Plaudern zum Job, in: Der Spiegel, Nr. 3, S. 66–67
- Paszczka, Bartosz, 2016, Comparison of Microsoft Academic (Graph) with Web of Science, Scopus and Google Scholar, https://www.researchgate.net/publication/313240234_Comparison_of_Microsoft_Academic_Graph_with_Web_of_Science_Scopus_and_Google_Scholar [22.8.2019]
- Petry, Thorsten / Jäger, Wolfgang, 2018, Digital HR – Ein Überblick, in: Petry, Thorsten / Jäger, Wolfgang (Hrsg.), Digital HR. Smarte und agile Systeme, Prozesse und Strukturen im Personalmanagement, Freiburg, München, Stuttgart, S. 27–99
- Rasmussen, Thomas / Ulrich, Dave, 2015, Learning from practice: how HR analytics avoids being a management fad, in: Organizational Dynamics, 44. Jg., Nr. 3, S. 236–242
- Reindl, Cornelia, 2016, People Analytics: Datengestützte Mitarbeiterführung als Chance für die Organisationspsychologie, in: Gruppe.Interaktion.Organisation., Zeitschrift für Angewandte Organisationspsychologie, 47. Jg., Nr. 2, S. 193–197
- Reindl, Cornelia / Krügl, Stefanie, 2017, People Analytics in der Praxis – inkl. Arbeitshilfen online, Freiburg im Breisgau

- Roberts, Margaret E. / Stewart, Brandon M. / Tingley, Dustin, 2018, stm: R Package for Structural Topic Models, <http://www.structuraltopicmodel.com> [22.8.2019]
- Roberts, Margaret E. / Stewart, Brandon M. / Tingley, Dustin / Airoldi, Edoardo M., 2013, The Structural Topic Model and Applied Social Science, Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation
- Scholz, Tobias M., 2017, Big Data in Organizations and the Role of Human Resource Management. A Complex Systems Theory-Based Conceptualization, Frankfurt am Main
- Shah, Naimatullah / Irani, Zahir / Sharif, Amir M., 2017, Big data in an HR context: Exploring organizational change readiness, employee attitudes and behaviors, in: Journal of Business Research, 70. Jg., S. 366–378
- Silberzahn, R. et al., 2018, Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results, in: Advances in Methods and Practices in Psychological Science, 1. Jg., Nr. 3, S. 337–356
- Simón, Cristina / Ferreiro, Eva, 2018, Workforce analytics: A case study of scholar-practitioner collaboration, in: Human Resource Management, 57. Jg., Nr. 3, S. 781–793
- Sinha, Arnab et al., 2015, An Overview of Microsoft Academic Service (MAS) and Applications, in: Gangemi, Aldo / Leonardi, Stefano / Panconesi, Alessandro (Hrsg.), Proceedings of the 24th International Conference on World Wide Web – WWW '15 Companion, New York, New York, USA, S. 243–246
- Statista Digital Economy Outlook, 2019, <https://static2.statista.com/download/pdf/DigitalEconomyCompass2019.pdf> [28.8.2019]
- Strohmeier, Stefan, 2015, Analysen der Human Resource Intelligence and Analytics, in: Strohmeier, Stefan / Piazza, Franca (Hrsg.), Human Resource Intelligence und Analytics, Springer Fachmedien Wiesbaden, S. 3–48
- Strohmeier, Stefan / Piazza, Franca, 2015, Prozesse der Human Resource Intelligence and Analytics, in: Strohmeier, Stefan / Piazza, Franca (Hrsg.), Human Resource Intelligence und Analytics, Springer Fachmedien Wiesbaden, S. 49–75
- Sullivan, John, 2013, How Google is using people analytics to completely reinvent HR, TLNT, <https://www.tlnt.com/how-google-is-using-people-analytics-to-completely-reinvent-hr/> [4.6.2019]
- The Guardian, 2018, Amazon ditched AI recruiting tool that favored men for technical jobs, <https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine> [22.8.2019]
- Ton, Zeynep, 2009, The effect of labor on profitability: the role of quality, Harvard Business School Technology & Operations Mgt. Unit Research Paper, Nr. 09-040
- Waber, Ben, 2013, People analytics. How social sensing technology will transform business and what it tells us about the future of work, Upper Saddle River, N.J.
- Weitzel, Tim et al., 2019, Digitalisierung und Zukunft der Arbeit. Ausgewählte Ergebnisse der Recruiting Trends 2019, einer empirischen Unternehmens-Studie mit den Top-1.000-

Unternehmen aus Deutschland sowie den Top-300-Unternehmen aus der Branche IT und der Bewerbungspraxis 2019, einer empirischen Kandidaten-Studie mit Antworten von über 3.300 Kandidaten, Bamberg