

Zeitschrift für Schweizerisches Recht
Revue de droit suisse
Rivista di diritto svizzero
Revista da dretg svizzer

Gegründet 1852, Neue Folge seit 1882

Band 138 (2019), ISSN 0254-945X

Band 138 I: 5 Hefte jeweils Ende der Monate März, Mai, Juli, September und Dezember

Band 138 II: Referate und Mitteilungen des Schweizerischen Juristenvereins (SJV)

Zitiervorschlag: ZSR 2019 I **Proposition de citation:** RDS 2019 I

Herausgeber: Prof. Pascal Pichonnaz, Prof. Ruth Arnet, Prof. Samantha Besson,
Dr. Peter Isler, Prof. Vincent Martenet, Dr. Andreas Zünd

Redaktion: Alessandra Beeler, RA lic. iur., E-Mail: alessandra.beeler@helbing.ch

Verlag: Helbing Lichtenhahn Verlag, Elisabethenstrasse 8, CH-4051 Basel
Telefon 061 228 90 70, Telefax 061 228 90 71, E-Mail: zeitschriften@helbing.ch

Beiträge sind per E-Mail (zeitschriften@helbing.ch) an den Verlag (Lektorat Zeitschriften) einzureichen. Richtlinien zur Textfassung sind beim Verlag anzufordern.

© 2019 by Helbing Lichtenhahn Verlag, Basel

Alle Rechte vorbehalten. Die Zeitschrift und ihre Teile sind urheberrechtlich geschützt. Jede Verwertung in anderen als den gesetzlich zugelassenen Fällen bedarf deshalb der vorherigen schriftlichen Einwilligung des Verlages.

Vertrieb: Helbing Lichtenhahn Verlag, Abo-Service, Elisabethenstrasse 8, CH-4051 Basel,
Telefon +41 (0)61 228 90 20, Telefax +41 (0)61 228 91 50, E-Mail: zeitschriften@helbing.ch

Jahresabonnement inkl. Onlinezugang (1 Benutzerlizenz)

Regulär CHF 298.– / EUR 298.– Studentenabonnement CHF 108.– / EUR 108.–

Versandkosten für Jahresabonnement

Europa CHF 26.– / EUR 26.– Übersee CHF 35.– / EUR 35.–

Einzelausgabe CHF 68.– / EUR 68.–, exklusive Versandkosten

Einbanddecke (Band I + II) CHF 85.– / EUR 85.–, exklusive Versandkosten

Alle Preise verstehen sich inkl. MwSt. Preisänderungen vorbehalten.

Kündigungen für die neue Abonnementsperiode sind schriftlich bis 8 Wochen vor Ablauf der Abonnementsperiode mitzuteilen (zeitschriften@helbing.ch).

La résiliation pour une nouvelle période d'abonnement doit être communiquée par écrit au plus tard 8 semaines précédant la nouvelle période (zeitschriften@helbing.ch).

Inserate und Beilagen (Verkauf): bw medien, Zürichstrasse 57, 8840 Einsiedeln, Tel. 055 418 82 00, Fax 055 418 82 22, E-Mail: info@bwmedien.ch. Verlangen Sie die Mediadaten.

Beihefte zur ZSR: Vergünstigung für ZSR-Abonnenten: 20% Rabatt auf Verkaufspreis.

SJV: Sämtliche Mitglieder des SJV erhalten die ZSR im Rahmen ihrer Mitgliedschaft.

SSJ: Les membres du SSJ reçoivent la RDS dans le cadre de leur cotisation.

Unterstützt durch die
Schweizerische Akademie
der Geistes- und Sozialwissenschaften
www.sagw.ch



Zeitschrift für Schweizerisches Recht
Revue de droit suisse
Rivista di diritto svizzero
Revista da dretg svizzer

Band 138 (2019) I · Heft 5

Inhaltsverzeichnis

Abhandlungen

- ISABELLE WILDHABER/MELINDA F. LOHMANN/GABRIEL KASPER,
Diskriminierung durch Algorithmen – Überlegungen zum
schweizerischen Recht am Beispiel prädiktiver Analytik
am Arbeitsplatz 459
- JOËLLE VUILLE/ALEX BIEDERMANN, Une preuve scientifique suffit-elle
pour fonder une condamnation pénale? 491
- MICHAEL KLING, Bezweckte Wettbewerbsbeschränkung und
Erheblichkeit – Ein neuer Revisionsvorschlag aus der Schweiz
im Vergleich zum unionskartellrechtlichen Regelungsmodell 513
- FEDERICA DE ROSSA GISIMUNDO, Pour un revenu équitable
(mais non inconditionnel) 539

Entscheidbesprechung

- KURT PÄRLI/LAURA KUNZ, Besprechung des Urteils des Bundesgerichts
2C_499/2015 vom 6.9.2017, BGE 144 I 50 561

Herausgeberchronik

1852

Gegründet von:

F. OTT, D. RAHN, J. SCHNELL, F. VON WYSS

1882–1920

A. HEUSLER, E. HUBER, P. SPEISER

1921–1948

E. HIS

1949–1970

M. GUTZWILLER, leitender Hrsg.; M. BATELLI, G. BROGGINI, P. CAVIN, H. FRITZSCHE,
O. A. GERMANN, A. GRISEL, H. HUBER, W. KÄGI, P. LIVER, K. SPIRO, W. YUNG

1971–1993

A. MEIER-HAYOZ, leitender Hrsg.; B. DUTOIT, P. SALADIN; ab 1992: B. SCHNYDER, D. THÜRER
Besprechungen: Dr. P. ISLER

1994–1997

B. DUTOIT, leitender Hrsg.; P. SALADIN, A. K. SCHNYDER, B. SCHNYDER, D. THÜRER
Besprechungen: Dr. P. ISLER

1998–1999

B. DUTOIT, leitender Hrsg.; A. K. SCHNYDER, B. SCHNYDER, D. THÜRER
Besprechungen: Dr. P. ISLER

2000–2003

B. DUTOIT, leitender Hrsg.; A. K. SCHNYDER, B. SCHNYDER, D. THÜRER, H. P. WALTER
Besprechungen: Dr. P. ISLER

2004

B. DUTOIT, leitender Hrsg.; P. ISLER, A. K. SCHNYDER, B. SCHNYDER, D. THÜRER, H. P. WALTER

2005–2006

B. DUTOIT, leitender Hrsg.; S. V. BERTI, P. ISLER, P. PICHONNAZ, A. K. SCHNYDER, D. THÜRER,
H. P. WALTER

2007–2015

B. DUTOIT, leitender Hrsg.; S. V. BERTI, P. ISLER, P. PICHONNAZ, D. THÜRER, H. P. WALTER

2016

B. DUTOIT, leitender Hrsg.; P. ISLER, P. PICHONNAZ, D. THÜRER, H. P. WALTER

2017

B. DUTOIT, leitender Hrsg.; S. BESSON, P. ISLER, P. PICHONNAZ, D. THÜRER

2018

P. PICHONNAZ, leitender Hrsg.; R. ARNET, S. BESSON, P. ISLER, V. MARTENET, A. ZÜND

Diskriminierung durch Algorithmen – Überlegungen zum schweizerischen Recht am Beispiel prädiktiver Analytik am Arbeitsplatz

ISABELLE WILDHABER*/MELINDA F. LOHMANN**/
GABRIEL KASPER***

Schlagwörter: Big Data, Predictive Analytics, Prädiktive Analytik,
People Analytics, HR Analytics, Workforce Analytics,
Diskriminierung, Bias, Algorithmen, Datenschutzrecht,
Algorithm Governance, Regulierung

A. Einleitung

2019 – Wir leben in einer Ära von Big Data. Unser zunehmender Verlass auf digitale Kommunikation, gepaart mit der technischen Möglichkeit, grosse Datenvolumen zu sammeln und zu analysieren, haben zum Einsatz von Techniken der sog. prädiktiven Analytik geführt. Diese werden in den verschiedensten Bereichen unseres Lebens eingesetzt, um Vorhersagen über die Zukunft zu machen, z.B. im Gesundheitswesen und am Arbeitsplatz. Der Einsatz von prädiktiver Analytik in der Arbeitswelt wird unter dem Titel «People Analytics» diskutiert. People Analytics unterstützt den Arbeitgeber dabei, mittels Datenanalysen die vielversprechendsten Mitarbeiter zu identifizieren, einzustellen, zu binden und zu belohnen. Dies gelingt mit Hilfe von Algorithmen, die darauf abzielen, eine grosse Menge Daten zu schneiden und zu würfeln, um bestimmte Informationen zu extrahieren. So können Muster und Beziehungen aufgezeigt werden, welche in einer kleineren Datenmenge nicht ersichtlich wären.

* Prof. Dr. iur., LL.M., Ordinaria für Privat- und Wirtschaftsrecht an der Universität St. Gallen, Direktorin am Forschungsinstitut für Arbeit und Arbeitswelten (FAA-HSG).

** Prof. Dr. iur., Assistenzprofessorin für Wirtschaftsrecht mit Schwerpunkt Informationsrecht an der Universität St. Gallen, Direktorin an der Forschungsstelle für Informationsrecht (FIR-HSG).

*** Wissenschaftlicher Assistent und Doktorand am Forschungsinstitut für Arbeit und Arbeitswelten der Universität St. Gallen (FAA-HSG).

Denkbar ist beispielsweise, dass ein Algorithmus einen Zusammenhang zwischen der Postleitzahl von Arbeitnehmern und ihrer Verweildauer im Unternehmen feststellen kann. Sofern Arbeitnehmer mit einer bestimmten Postleitzahl und einem entsprechend langen Anfahrtsweg eine höhere Fluktuation aufweisen, wird der Algorithmus Bewerbungen von Arbeitnehmern, deren Arbeitsweg eine bestimmte Länge überschreitet, aussortieren. In der Tat bleiben Arbeitnehmer, die 0–5 Meilen von ihrem Arbeitsplatz entfernt wohnen, 20% länger in ihren Positionen als diejenigen von weiter weg, wie das Softwareunternehmen Evolv in einer Analyse herausgefunden hat.¹ Die Anwendung eines solchen «Redlining» durch den Arbeitgeber im Bewerbungsprozess, d.h. das Ziehen einer roten Linie bspw. entlang der ermittelten 5-Meilen-Linie, mag auf den ersten Blick nicht weiter schlimm erscheinen; denn es ist verständlich, dass der Arbeitgeber wissen will, wo sich der beste Pool an potenziellen Talenten befindet. Aber was ist, wenn der Algorithmus Bewerber ausschliesslich anhand ihrer Postleitzahl aussortiert und in diesen Nachbarschaften mehrheitlich Minderheiten leben?² Durch die Nutzung eines solchen Algorithmus würde ein Arbeitgeber u.U. gewisse Minderheiten vollständig unberücksichtigt lassen, im Zweifel ohne es überhaupt zu merken. Diskriminierung wird so in den Bewerbungsprozess eingearbeitet und wird damit zu einem strukturellen Problem.

Es gibt verschiedene Gründe, warum Algorithmen diskriminieren können. Festzulegen, wann und wo die Grenze zur Diskriminierung überschritten wird und was dagegen unternommen werden kann, ist in unseren Augen eine grosse Herausforderung bei der Nutzung von Algorithmen am Arbeitsplatz. Arbeitgeber, die das Potenzial von Big Data nutzen wollen, betreten Neuland.³ Big Data passt nicht sauber in den juristischen Rahmen, der unsere Arbeitsverhältnisse prägt. Der bestehende juristische Rahmen ist auf eine analoge Welt abgestimmt und bietet wenige Wegweiser für eine digitale Welt. Es ist deshalb an der Zeit, über die rechtlichen Schutzmechanismen bei datengesteuerter, algorithmischer Diskriminierung nachzudenken. In den USA hat das Executive Office des Präsidenten die Bundesregierung 2014 ermahnt, auf die Gefahr zu achten, dass Big Data Diskriminierungen erleichtert, die mit den Gesetzen und Werten des Landes unvereinbar seien.⁴ Das Europäische Parlament hat im Feb-

ruar 2017 betont, wie wichtig es sei, mögliche Massnahmen zu identifizieren und zu ergreifen, um Diskriminierungen durch Algorithmen zu minimalisieren.⁵

Ziel dieses Aufsatzes ist es aufzuzeigen, wie Algorithmen in der Arbeitswelt genutzt werden (nachfolgend B.), wie sie diskriminieren können (nachfolgend C.), und wie das Recht mit einer möglichen Diskriminierung durch Algorithmen am Arbeitsplatz umgehen kann und soll (nachfolgend D.), abgeschlossen von einem Fazit (nachfolgend E.). Dabei scheint es uns wichtig, die Grenzen des Diskriminierungsschutzrechts aufzuzeichnen und zu erkennen, dass sich dieses Konzept bei datengesteuerten, algorithmischen Diskriminierungen als wenig hilfreich erweist.⁶

B. Nutzung von Algorithmen am Arbeitsplatz

I. Prädiktive Analytik

Vorhersagen sind das Herzstück von Big Data.⁷ Mit «prädiktiver Analytik»⁸ (engl.: «Predictive Analytics») werden bisher nicht zusammenhängende Daten kombiniert, um präzise Prognosen zu machen. Als Methode werden «Machine Learning»-Modelle genutzt, um Trends, Muster und Beziehungen zwischen Daten zu identifizieren. Anhand der aufgedeckten Muster werden Dinge oder Aktivitäten klassifiziert, der Wert von unbeobachteten Variablen eingeschätzt und – basierend auf Wahrscheinlichkeiten – Verhalten vorausgesehen. Es geht bei der prädiktiven Analytik darum, eine Grundlage dafür zu schaffen, einem Individuum Eigenschaften zuzuschreiben, die diejenigen besitzen, die statistisch ähnlich erscheinen.⁹ Im selben Atemzug werden diejenigen Individuen abgesondert, die statistisch verschieden sind. Prädiktive Analytik ist somit per Definition immer eine Form der statistischen (und damit scheinbar rationalen)

1 ALEX ROSENBLAT/KATE WIKELIUS/DANAH BOYD/SEETA PEÑA GANGADHARAN/CORRINE YU, Data & civil rights: employment primer, 30.10.2014, <<http://www.datacivilrights.org/pubs/2014-1030/Employment.pdf>>, S. 2.

2 BORIS DZIDA/NAEMI GROH, Diskriminierung nach dem AGG beim Einsatz von Algorithmen im Bewerbungsverfahren, NJW 2018, S. 1917–1922, S. 1919 (mittelbare Benachteiligung nach AGG möglich).

3 Sog. «uncharted territory», REBECCA J. WILSON/KILEY M. BELLIVEAU/LEIGH ELLEN GRAY, Busting the black box: big data, employment and privacy, Def. Counsel J. 2017 1, S. 2–34, S. 9.

4 Executive Office of the President, Big data: seizing opportunities, preserving values, 01.05.2014, S. 46; Federal Trade Commission, Big data: a tool for inclusion or exclusion? Event Webpage, 15.09.2014, <<http://www.ftc.gov/news-events/events-calendar/2014/09/big-data-tool-inclusion-or-exclusion>>.

5 European Parliament, Report on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law enforcement, 2016/2225(INI), A8–0044/2017, E. 20.

6 DANIELLE KEATS CITRON/FRANK PASQUALE, The scored society: due process for automated predictions, Wash.L. Rev. 2014 1 S. 1–33; TAL Z. ZARSKY, Understanding discrimination in the scored society, Wash.L. Rev. 2014, S. 1375–1412.

7 VIKTOR MAYER-SCHÖNBERGER/KENNETH CUKIER, Big data: a revolution that will transform how we live, work and think, Boston MA 2013, S. 55. Ähnlich SOLON BAROCAS, Progress and continuing challenges in eradicating employment discrimination, written testimony, Meeting of the Equal Employment Opportunity Commission, 01.07.2015, <<https://www.eeoc.gov/eeoc/meetings/7-1-15/barocas.cfm>>.

8 Mit der zunehmenden Diskussion der mit Algorithmen verbundenen Rechtsprobleme im deutschen Sprachraum verbreitet sich die deutsche Übersetzung von «Predictive Analytics», so z.B. in: YOANN HERMSTRÜWER, Die Regulierung der prädiktiven Analytik: eine juristisch-verhaltenswissenschaftliche Skizze, in: Wolfgang Hoffmann-Riem (Hrsg.), Big Data – regulative Herausforderungen, Baden-Baden 2018, S. 99–116.

9 SOLON BAROCAS/ANDREW B. SELBST, Big data's disparate impact, Calif. L. Rev. 2016, S. 671–732, S. 677.

Diskriminierung – dies übrigens ganz im etymologischen Wortsinn von «discriminare» (lat.: trennen, absondern).

Die Vorhersagen der prädiktiven Analytik basieren auf *Korrelationen*. Es geht in der Regel darum, herauszufinden, welche Eigenschaften eines Mitarbeiters oder Bewerbers mit dem Erfolg des Geschäfts korrelieren.¹⁰ Nehmen wir ein Beispiel: Besucher japanischer Mangawebsites scheinen häufig gute Programmierer zu sein.¹¹ Eine Affinität zu einer bestimmten japanischen Mangawebsite ist somit ein scheinbar zuverlässiger Prädiktor für starke Programmierfähigkeiten. Zwei Merkmale, die Wertschätzung von Manga und das Kodiergeschick, korrelieren, aber keines von beiden verursacht das andere (im Sinne einer nachweisbaren Kausalität).

Korrelationen sind nur nützlich, wenn sie konstant sind. Wenn im Januar die besten Programmierer zu einer bestimmten Mangawebsite strömen, aber im Juli eine andere Website hip ist, wird ein Algorithmus, der sich weiterhin auf die Besuche der ersten Website verlässt, die Bewerber falsch einstufen. Wichtig für die Vorhersagegenauigkeit ist somit die Stabilität der Korrelation. Das Aufdecken *stabiler Korrelationen* soll als Kompensation für den Umstand dienen, dass die prädiktive Analytik keine grundlegenden Ursache-Wirkungs-Beziehungen nachweisen kann.¹²

Die prädiktive Analytik wird nicht-kausale gegenüber kausalen Analysen stärken. Wie MAYER-SCHÖNBERGER und CUKIER sagen: «Causality won't be discarded, but it is being knocked off its pedestal as the primary fountain of meaning.»¹³ Dies kann zu potenziell unfairen und diskriminierenden Folgen führen.¹⁴ Unser Rechtssystem, das vom Kausalitätsprinzip geprägt ist, steht somit vor der Herausforderung, sich in einer von Korrelationen bestimmten Welt der prädiktiven Analytik und deren möglicherweise diskriminierenden Folgen weiterzuentwickeln.¹⁵

10 Siehe die Beispiele in ROSENBLAT et al. (Fn. 1), S. 2.

11 DON PECK, They're watching you at work, The Atlantic 12.2013, <<http://www.theatlantic.com/magazine/archive/2013/12/theyre-watching-you-at-work/354681/>>.

12 MICHAEL J. MAUBOUSSIN, The true measures of success, Harv. Bus. Rev., 10.2012, <<https://hbr.org/2012/10/the-true-measures-of-success>>; ALLAN G. KING/MARKO MRKONICH, «Big data» and the risk of employment discrimination, Okla. L. Rev. 2016, S. 555–584, S. 555.

13 MAYER-SCHÖNBERGER/CUKIER (Fn. 7), S. 68; eine wissenschaftliche Gegenbewegung will jedoch vermehrt Kausalitätsbeziehungen statt statistischer Korrelationen abbilden: JUDEA PEARL, Causality, 2. Aufl., Cambridge 2009 und JUDEA PEARL/DANA MACKENZIE, The book of why: the new science of cause and effect, New York City 2018. Z.B. sollen Maschinen künftig erklären, dass Malaria Fieber verursacht, statt bloss Fieber und Malaria zu korrelieren: Judea Pearl im Interview mit KEVIN HARTNETT, How a pioneer of machine learning became one of its sharpest critics, The Atlantic 19.05.2018, <<https://www.theatlantic.com/technology/archive/2018/05/machine-learning-is-stuck-on-asking-why/560675/>>.

14 PAULINE T. KIM/ERIK A. HANSON, People analytics and the regulation of information under the Fair Credit Reporting Act, St. Louis U. L. J. 2016, S. 17–33, S. 19.

15 The Littler Report, The big move toward big data in employment, 08.2015, <<https://www.littler.com/publication-press/publication/big-move-toward-big-data-employment>>, S. 563.

II. People Analytics

Prädiktive Analytik kommt auch am Arbeitsplatz zur Anwendung; wir sprechen dann von People Analytics.¹⁶ Mittels technologie-gestützter Analysen kombiniert sie traditionelle Beschäftigungsdaten (z.B. Leistungsbeurteilungen, Krankheitstage oder Saläre) und neue Daten (z.B. Social Media-Aktivitätsprotokolle, Sensordaten, Verbraucherdaten aus GPS- oder Tracking-Systemen), um Prozesse zur Identifizierung, Einstellung, Bindung und Belohnung von Stellenbewerbern und Mitarbeitern zu schaffen.¹⁷

Als moderne Form der Mitarbeiterüberwachung unterscheidet sich People Analytics von herkömmlichen elektronischen Überwachungen durch mindestens drei Faktoren: Erstens steigt die Vielfalt der über den Arbeitnehmer erhobenen Daten, z.B. fliessen in die Analyse nicht mehr nur arbeits-, sondern auch kontextbezogene Daten (Einbindung des Mitarbeiters im Unternehmen, Gesundheitsergebnisse, Verhalten der Mitarbeiter ausserhalb des Arbeitsplatzes, u.ä.).¹⁸ Zweitens steigt die Interoperabilität der Systeme, wodurch Daten, die traditionell getrennt voneinander lagerten, zusammengeführt werden können, z.B. werden Daten aus dem Personalbereich mit anderen Unternehmensdaten oder externen Daten verknüpft.¹⁹ Drittens steigt die analytische Leistungsfähigkeit der Systeme: Neben deskriptiven (beschreibenden) Algorithmen kommen neu auch prädiktive (voraussagende) und präskriptive (Handlungen empfehlende) zum Einsatz. Gleichzeitig werden die den Algorithmen zugrunde liegenden versteckten Annahmen aufgrund der steigenden Komplexität der Systeme undurchsichtiger.²⁰

III. Anwendungsarten von People Analytics in der Schweiz

People Analytics kann in vielen Bereichen benutzt werden, um die Produktivität, Effizienz und Innovationskraft zu steigern. Diese Nutzungen sind in der Literatur schon vielfach beschrieben worden.²¹ Ein aktuelles Forschungsprojekt

16 ROBERT SPRAGUE, Welcome to the machine: privacy and workplace implications of predictive analytics, 21 Rich. J. L. & Tech. 2015, S. 1–46, S. 31; STEVE LOHR, Trying to build better workers, The New York Times 21.04.2013, <<http://www.nytimes.com/2013/04/21/technology/big-data-trying-to-build-better-workers.html>>.

17 KELLY TRINDEL, Big data in the workplace: examining implications for equal employment opportunity law, Written Testimony, Meeting of the Equal Employment Opportunity Commission 13.10.2016, <<https://www.eeoc.gov/eeoc/meetings/10-13-16/trindel.cfm>>.

18 DAVID ANGRAVE/ANDY CHARLWOOD/IAN KIRKPATRICK/MARK LAWRENCE/MARK STUART, HR and analytics: why HR is set to fail the big data challenge, Human Resource Management Journal 2016, S. 1–11; ALEX ROSENBLAT/TAMARA KNEESE/DANAH BOYD, Workplace surveillance, 08.10.2014 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2536605>.

19 BORIS DZIDA, Big Data und Arbeitsrecht, NZA 2017, S. 541–546, S. 541; ULRICH LEICHT-DEOBALD/THORSTEN BUSCH/CHRISTOPH SCHANK/ANTOINETTE WEIBEL/SIMON SCHAFHEITEL/ISABELLE WILDHABER/GABRIEL KASPER, The challenges of algorithm-based HR decision-making for personal integrity, J. Bus. Ethics 07.06.2019 (Online-Veröffentlichung), S. 3.

20 LEICHT-DEOBALD et al. (Fn. 19), S. 3 ff.

21 Statt vieler SPRAGUE (Fn. 16), S. 33; Littler Report (Fn. 15), S. 12.

des Schweizerischen Nationalfonds untersucht den Einsatz von People Analytics am Arbeitsplatz.²² Die beteiligten Forscher haben gestützt auf eine Analyse der entsprechenden Fachliteratur und 27 Experteninterviews fünf Gruppen von Anwendungen identifiziert: Eine erste Gruppe betrifft die Rekrutierung und Auswahl in der Bewerbungsphase (Recruitment and Selection). Sodann kann People Analytics während des laufenden Arbeitsverhältnisses zur Leistungssteuerung (Performance Management), zur Überwachung der Einhaltung der Gesetze (Compliance Management) und zur Gestaltung der Arbeit und Arbeitsplätze (Work and Workplace Design) dienen. Verbreitet ist schliesslich die Gruppe der Anwendungen zur Mitarbeiterbindung und Analysen bei Mitarbeiterwechseln (Retention and Transition).

C. Ursachen der Diskriminierung durch Algorithmen

Der Hype um das Potential von Big Data und prädiktiver Analytik führt bei vielen Menschen und Unternehmen zur Überzeugung, dass Daten stets die objektive Wahrheit reflektieren würden. Der frühere Chefredaktor der amerikanischen Computerzeitschrift Wired, CHRIS ANDERSON, begeisterte sich in einem Kommentar im Jahr 2008 für diese Idee der einzigen objektiven Wahrheit. Er bezeichnete die wissenschaftliche Methode der Identifikation und Überprüfung von Ursache-Wirkungs-Ketten als überholt: «with enough data, the numbers speak for themselves.»²³ Aber können Zahlen immer für sich selbst sprechen? Unserer Meinung nach ist ein unkritischer Zahlenglaube gefährlich, denn Zahlen werden von Menschen kreiert und interpretiert – und daraus können sich Diskriminierungen ergeben. Schon die Definition des Gewünschten, z.B. was unter einer guten Arbeitsleistung zu verstehen ist,²⁴ wird vom Menschen bestimmt.

Das Problem möglicher Diskriminierungen durch Algorithmen am Arbeitsplatz wurde in den USA schon vor einigen Jahren erkannt. Die amerikanische Equal Employment Opportunity Commission (EEOC) beschäftigt sich seit

2014²⁵ mit dieser Frage.²⁶ Kritiker der prädiktiven Analytik am Arbeitsplatz monieren u.a. die Tatsache, dass Arbeitgeber – ohne das Wissen der Betroffenen – Daten aus digitalen Aktivitäten verwenden, um Annahmen über das Verhalten dieser Personen zu treffen, die sich wiederum auf deren Chancen und Möglichkeiten bezüglich ihres wirtschaftlichen Fortkommens auswirken.²⁷

Diskriminierungen können beabsichtigt oder unbeabsichtigt geschehen. Denkbar und technisch machbar ist eine *vorsätzliche Diskriminierung*, indem die prädiktive Analytik zur Verschleierung diskriminierender Motive verwendet wird.²⁸ Es wird sich jedoch für Arbeitgeber vielfach nicht lohnen, komplexe und kostspielige Methoden des Machine Learning zu nutzen, nur um ihre diskriminierenden Absichten hinter einem scheinbar neutralen Datenmodell zu verbergen.²⁹ Hierfür gibt es einfachere Wege; die meisten Fälle von Diskriminierungen am Arbeitsplatz – auch analoge Diskriminierungen ausserhalb der Welt von Big Data und Algorithmen – sind schwierig nachzuweisen. In unseren Augen ist die *unbeabsichtigte Diskriminierung* bei der prädiktiven Analytik das drängendere Problem, da sie wahrscheinlich häufiger vorkommt und leichter zu übersehen ist.³⁰ Häufig fliessen Diskriminierungen unbeabsichtigt, als Produkte unbewusster menschlicher Verzerrungen («implicit bias», z.B. fehlendes Bewusstsein für Lücken im Datensatz) oder als einfache Fehler (z.B. manuelle Tippfehler), in den algorithmischen Prozess ein.³¹

Es gibt viele verschiedene *Ursachen für algorithmische Diskriminierungen*. Die folgende Grafik³² zeigt vereinfacht das Verfahren der Entscheidungsfindung mit Hilfe eines Algorithmus. Zunächst können sich diskriminierende Fehler in der

- 22 Im Projekt «Big Brother in Schweizer Unternehmen? Vertrauen, Daten und Privatsphäre im Job» arbeiten Prof. Dr. Isabelle Wildhaber, LL.M., Prof. Dr. Antoinette Weibel, Gabriel Kasper und Simon Schafheide (alle vom Forschungsinstitut für Arbeit und Arbeitswelten, Universität St. Gallen) sowie Prof. Dr. Christoph Schank und Isabel Ebert (beide vom Institut für Wirtschaftsethik, Universität St. Gallen).
- 23 CHRIS ANDERSON, The end of theory: the data deluge makes the scientific method obsolete, Wired 23.06.2008, <<https://www.wired.com/2008/06/pb-theory>>; ähnlich MATT RICHTEL, How big data is playing recruiter for specialized workers, The New York Times 27.04.2013; kritisch KATE CRAWFORD, The hidden biases in big data, Harv. Bus. Rev. 01.04.2013, <<https://hbr.org/2013/04/the-hidden-biases-in-big-data>>.
- 24 PHILIPP HACKER, Teaching fairness to artificial intelligence: existing and novel strategies against algorithmic discrimination under EU law, Common Market Law Review 2018, S. 1143–1186, S. 1147; ISABELLE WILDHABER/MELINDA LOHMANN, Video «Understanding Algorithms», 11.12.2018, <<https://www.youtube.com/watch?v=XKSenhb6Sww>>.

- 25 Aussage Chair EEOC: JACQUELINE A. BERRIEN, Social media in the workplace: examining implications for equal employment opportunity law, Meeting of the Equal Employment Opportunity Commission, 12.03.2014, <<https://www.eeoc.gov/eeoc/meetings/3-12-14/transcript.cfm>>; Big data in the workplace: examining implications for equal employment opportunity law, Meeting of the Equal Employment Opportunity Commission, 13.10.2016, <<https://www.eeoc.gov/eeoc/meetings/10-13-16/>>.
- 26 EEOC research and data plan for 2016–2019, <https://www.eeoc.gov/eeoc/plan/research_data_plan.cfm>; U.S. Equal Employment Opportunity Commission Strategic Enforcement Plan Fiscal Years 2017–2021, <<https://www.eeoc.gov/eeoc/plan/sep-2017.cfm>>.
- 27 IAN KERR/JESSICA EARLE, Prediction, preemption, presumption: how big data threatens big picture privacy, Stan. L. Rev. Online 2013, S. 65–72, S. 71.
- 28 HACKER (Fn. 24), S. 1149; MATTHEW T. BODIE/MIRIAM A. CHERRY/MARCIA L. McCORMICK/JINTONG TANG, The law and policy of people analytics, U. Colo. L. Rev. 2017, S. 961–1042, S. 1014.
- 29 BAROCAS/SELBST (Fn. 9), S. 692; BART CUSTERS, Data dilemmas in the information society: introduction and overview, in: Bart Custers/Toon Calders/Bart Schermer/Tal Zarsky (Hrsg.), Discrimination and privacy in the information society, Berlin/Heidelberg 2013, S. 3–26, S. 9 f.
- 30 BAROCAS/SELBST (Fn. 9), S. 693.
- 31 ANDREA ROMEI/SALVATORE RUGGERI, A multidisciplinary survey on discrimination analysis, The Knowledge Engineering Review 2014, S. 582–638, S. 585 ff.; TAL Z. ZARSKY (Fn. 6), S. 1389; BAROCAS/SELBST (Fn. 9), S. 712 ff.
- 32 Die vorliegende übersetzte und leicht veränderte Grafik stützt sich auf: algo:aware, State-of-the-art report: algorithmic decision-making, 12.2018, <<https://www.algoaware.eu/wp-content/uploads/2018/08/AlgoAware-State-of-the-Art-Report.pdf>>.

Input-Phase bei der Datenerhebung und -aufbereitung ergeben. Ursachen können sodann im Modell selber liegen (z.B. Wahl der Modellarchitektur und des Algorithmus und der relevanten Messwerte) oder sich aus der Output-Phase ergeben (z.B. unsensibler Umgang mit den Resultaten).³³ Wir wollen nun diese Quellen der algorithmischen Diskriminierung erläutern.

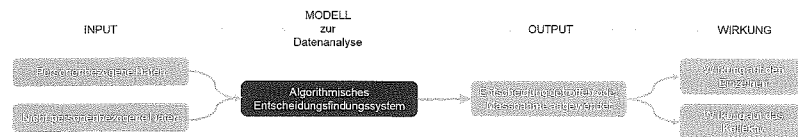


Abb. 1: Verfahren der Entscheidungsfindung mit Hilfe eines Algorithmus

Die Inputphase beschreibt den Prozess, bei dem die Daten aufbereitet werden. Der grösste Teil der Input-Daten wird dazu verwendet, den (Roh-)Algorithmus zu trainieren (sog. Trainingsdaten), während ein kleinerer Datensatz aufbewahrt wird, um später zu testen, ob der trainierte Algorithmus richtig funktioniert (sog. Testdaten), bevor er in der Praxis eingesetzt wird. Die Inputdaten können rechtlich geschützte Gruppen, z.B. Frauen, Behinderte oder ethnische Minoritäten systematisch benachteiligen (sog. «statistical bias»).34 Die Daten können schlecht ausgewählt, unvollständig, falsch, veraltet oder nicht-repräsentativ sein.³⁵ Wenn beispielsweise der Gesichtserkennungsalgorithmus von Google fälschlicherweise schwarze Menschen als Gorillas bezeichnete, lag dies an einem unzureichenden Training mit schwarzen Gesichtern.³⁶ KI-Schönheitswettbewerbe haben mehrheitlich weisse Gewinner erkoren, da in den Trainingsdaten mehr Weisse als Nicht-Weisse vertreten waren.³⁷ Solche Fehler bei den Inputdaten können statistisch verzerrte Datenmodelle nach sich ziehen.

Das Phänomen der falschen Handhabung von Trainingsdaten, bei welchem ein Teil der Bevölkerung unterrepräsentiert ist, nennt sich «*sampling bias*».³⁸

Bei Big Data gibt es sog. *Signalprobleme*, d.h. es gibt Schattenzonen, in denen einige Bürger und Gemeinschaften übersehen oder unterrepräsentiert werden.³⁹ Von ihnen kommen keine oder wenige Signale. Die Qualität und Repräsentativität der Trainingsdaten kann in einer Weise variieren, die mit der Zugehörigkeit gewisser geschützter Klassen korreliert, z.B. könnten Institutionen systematisch weniger genaue, aktuelle und vollständige Aufzeichnungen für die unterrepräsentierten Personengruppen führen. Auf diese Art und Weise könnten Gemeinschaften, deren Leben weniger datafiziert ist als das der allgemeinen Bevölkerung, sei es aufgrund von Armut, Geographie oder Lebensstil, diskriminiert werden.⁴⁰ Korreliert ein Algorithmus effizientes Arbeiten mit dem Fahren schneller Autos, so werden Stellenbewerber ausgeschlossen, die zu arm sind, um ein Auto zu besitzen, die den öffentlichen Verkehr benutzen, oder die eine Behinderung haben, welche sie am Autofahren hindert. Bei Big Data müssen wir also stets fragen, welche Personen im Datensatz fehlen und welche Orte weniger sichtbar sind.

Ein weiterer Subtyp der voreingenommenen Trainingsdaten ergibt sich, wenn die *Trainingsdaten mit historischer Verzerrung* behaftet sind. Machine Learning kann bestehende Diskriminierungsmuster vertiefen.⁴¹ Ein Algorithmus, der Kandidaten identifizieren soll, die in die existierende Unternehmenskultur passen, kann unbeabsichtigt vergangene Anstellungspraktiken perpetuieren, welche historisch benachteiligte Gruppen ausgeschlossen haben.⁴² Die leistungsstarke Gruppe kann homogen sein, und daher widerspiegeln die Merkmale dieser Gruppe mehr ihre demografischen Privilegien als die Fähigkeiten, die zur Erfüllung der Aufgabe erforderlich sind. Beispielsweise befanden sich auf der Führungsetage von Unternehmen bisher mehrheitlich Männer. Das Resultat des Sortieralgorithmus droht dann, auf den Eigenschaften der Personen (Geschlecht), nicht auf den Anforderungen des Arbeitsplatzes (Führungsstärke) zu beruhen.⁴³

In der Modellphase kann eine sog. *Proxy-Diskriminierung* erfolgen: Der Algorithmus beobachtet gerade *nicht* ein geschütztes Merkmal wie Geschlecht, Behinderung oder Rasse, sondern eine vordergründig harmlose Stellvertretervariable. Diese leicht beobachtbare Variable kann mit der Zugehörigkeit zu einer geschützten Gruppe korrelieren, weshalb sie als Stellvertreter (Proxy) eines ge-

33 Vgl. PAULINE T. KIM, Data-driven discrimination at work, Wm. & Mary L. Rev. 2017, S. 857–936, 890 f.; BAROCAS/SELBST (Fn. 9), S. 677 ff.; Executive Office of the President, Big data: a report on algorithmic systems, opportunity, and civil rights, 05.2016, http://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf; BODIE/CHERRY/McCORMICK/TANG (Fn. 28), S. 2015.

34 KIM (Fn. 33), S. 887.

35 HACKER (Fn. 24), S. 1147 f.; BAROCAS/SELBST (Fn. 9), S. 681.

36 ALISTAIR BARR, Google Mistakenly Tags black people as «gorillas», showing limits of algorithms, The Wall Street Journal 01.07.2015, <https://blogs.wsj.com/digits/2015/07/01/google-mistakenly-tags-black-people-as-gorillas-showing-limits-of-algorithms/>.

37 SAM T. LEVIN, A beauty contest was judged by AI and the robots didn't like dark skin, The Guardian 08.09.2016, <https://www.theguardian.com/technology/2016/sep/08/artificial-intel-ligence-beauty-contest-doesnt-like-black-people>.

38 JOSHUA A. KROLL/JOANNA HUEY/SOLON BAROCAS/EDWARD W. FELTEN/JOEL R. REIDENBERG/DAVID G. ROBINSON/HARLAN JU, Accountable algorithms, U. Pa. L. Rev. 2017, S. 633–705, S. 681; MICHAEL VEALE/REUBEN BINNS, Fairer machine learning in the real world: mitigating discrimination without collecting sensitive data, Big Data & Society 2017 (2), S. 1–17, S. 2.

39 KATE CRAWFORD (Fn. 23).

40 JONAS LERMAN, Big data and its exclusions, 66 Stan. L. Rev. Online 2013, S. 55–63, S. 57; BAROCAS/SELBST (Fn. 9), S. 684–685; KATE CRAWFORD, Big data: why the rise of machines isn't all it's cracked up to be, Foreign Policy 10.05.2013, <http://foreignpolicy.com/2013/05/10/think-again-big-data/>; BODIE/CHERRY/McCORMICK/TANG (Fn. 28), S. 1015.

41 SPRAGUE (Fn. 16), S. 46; BAROCAS/SELBST (Fn. 9), S. 681.

42 Executive Office of the President, 2016 (Fn. 33); BAROCAS/SELBST (Fn. 9), S. 682.

43 Vgl. dazu die Erfahrungen mit einem algorithmischen Bewertungssystem von Amazon: MARTIN HOLLAND, Amazon: KI zur Bewerbungsprüfung benachteiligte Frauen, Heise Online 11.10.2018, <https://www.heise.de/newsticker/meldung/Amazon-KI-zur-Bewerbungspruefung-benachteiligte-Frauen-4189356.html>.

geschützten Merkmals fungiert.⁴⁴ Auch wenn also die geschützten Informationen nicht explizit erfasst werden, kann die Zugehörigkeit zu geschützten Kategorien dennoch abgeleitet werden.⁴⁵ So können relativ triviale Informationen wie Postleitzahl oder Facebook-Likes eng mit geschützten Merkmalen wie Behinderung, Geschlecht oder Alter korrelieren oder die politischen oder religiösen Ansichten einer Person offenbaren.⁴⁶ Dies erklärt, warum die blosser Eliminierung sensibler Attribute aus dem Modell keine Nichtdiskriminierung garantiert.⁴⁷ Ein Beispiel für eine Proxy-Diskriminierung ist das bereits erwähnte «Redlining»: Nur Personen mit einer bestimmten, privilegierten Postleitzahl (kein rechtlich geschütztes Merkmal) sehen beim Surfen eine Stellenanzeige im Browser. Dadurch bekommen gewisse geschützte Minderheiten, die anderswo wohnen, die Anzeige nie zu Gesicht, womit ihnen die Chance auf eine Bewerbung entgeht.

In der Outputphase ergibt sich eine mögliche Ursache für algorithmische Diskriminierungen daraus, dass Algorithmen auf verallgemeinerten Statistiken basieren und nicht auf der individuellen Situation des vom Algorithmus betroffenen Menschen. Aus Gruppenwahrscheinlichkeiten werden Rückschlüsse auf Einzelne gezogen. Diese Rückschlüsse können diskriminierend wirken.⁴⁸ Betreibt jemand z.B. für ein bestimmtes Geschenk eine Produkterecherche, heisst das nicht, dass diese Person für sich selber am Produkt interessiert ist. Googelt etwa jemand auf der Suche nach einem Hochzeitsgeschenk nach Fritteusen, lebt diese Person selber nicht zwingend ungesund.⁴⁹ Wenn nun aus der Korre-

lation etwas Konkretes abgeleitet wird, z.B. eine Nichteinstellung oder eine höhere Versicherungsprämie, so kann der Output der prädiktiven Analytik hier diskriminierend wirken.

Aus der (rechtlich relevanten) Wirkung der algorithmischen Entscheide auf Einzelne und Gruppen ergeben sich Rechtsfragen, mit denen wir uns im folgenden Kapitel befassen.

D. Rechtliche Überlegungen für die Schweiz

Der potentiell diskriminierende Effekt von Algorithmen wäre kein regulatorisches Problem, wenn sichergestellt wäre, dass der Markt algorithmische Diskriminierung im Laufe der Zeit eliminieren würde. Diese Hoffnung scheint aber nicht berechtigt zu sein. Die Arbeitsmärkte sind so beschaffen, dass Arbeitgeber keine zuverlässigen Signale empfangen, wenn die Anwendung von People Analytics zu Diskriminierungen gegenüber Minderheitengruppen führt. Ausserdem fehlen denjenigen, die People Analytics einsetzen, oft ausreichende Anreize, um Diskriminierungen zu minimieren.⁵⁰ Der Wettbewerb wird die algorithmische Diskriminierung somit kaum beseitigen, womit als Nächstes regulatorische Ansätze zu prüfen sind.

I. Diskriminierungsschutzrecht am Arbeitsplatz

Im Folgenden prüfen wir zunächst, wie der bestehende arbeitsrechtliche Diskriminierungsschutz in der Schweiz mit dem Problem möglicher Diskriminierungen durch Algorithmen umgeht. Sodann fragen wir, ob sich das schweizerische Diskriminierungsschutzrecht de lege ferenda anders ausgestalten liesse, um möglichen Diskriminierungen durch Algorithmen vorzubeugen. Auf der Suche nach künftigen Lösungsrezepten blicken wir auch auf das US-amerikanische Antidiskriminierungsrecht und auf das europäische Gegenstück.

1. Schweizerisches Diskriminierungsschutzrecht

Es gibt in der Schweiz ein *allgemeines arbeitsrechtliches Diskriminierungsverbot*, welches sich auf den zivilrechtlichen Persönlichkeitsschutz von Art. 328 OR stützt.⁵¹ Eine diskriminierende Persönlichkeitsverletzung liegt vor, wenn die Ungleichbehandlung eines Arbeitnehmers an diskriminierungssensible Persönlichkeitsmerkmale anknüpft. Solange der Gesetzgeber kein umfassendes arbeitsrechtliches Diskriminierungsschutzgesetz erlässt, kommt den Gerichten

44 HACKER (Fn. 24), S. 1148; ROMEL/RUGGIERI (Fn. 31), S. 592 f., S. 606 f.; PHILIPP HACKER/BILYANA PETKOVA, Reining in the big promise of big data, transparency, inequality, and new regulatory frontiers, Nw. J. Tech. & Intell. Prop. 2017, S. 1–42, S. 9 ff.

45 BAROCAS/SELBST (Fn. 9), S. 692.

46 MICHAL KOSINSKI/DAVID STILLWELL/THORE GRAEPEL, Private traits and attributes are predictable from digital records of human behavior, Proceedings of the Nat'l Acad. of Sci. of the U.S. 2013, S. 5802–5805, S. 5802 (aufzeigend, dass Facebook-Likes einer Person verwendet werden können, um persönliche Merkmale wie Rasse, Geschlecht und sexuelle Orientierung, religiöse und politische Ansichten und Persönlichkeitsmerkmale wie Intelligenz genau vorherzusagen); SOLON BAROCAS, Data mining and the discourse on discrimination, 2014, <<https://pdfs.semanticscholar.org/abb7/235fcf3b163afd74e1967f7d3784252b44fa.pdf>>. Die neuen, digital erstellten Diskriminierungsmerkmale sind gesetzlich nicht geschützt: ROMY DADELOW, Wenn Algorithmen (unfair) über Menschen entscheiden ..., Jusletter 26.11.2018, Rz. 1–39, Rz. 38.

47 VEALE/BINNS (Fn. 38), S. 4; INDRE ŽLIOBAITE/BART CUSTERS, Using sensitive personal data may be necessary for avoiding discrimination in data-driven decision models, Artificial Intelligence and Law 2016, S. 183–201, S. 194.

48 BAROCAS/SELBST (Fn. 9), S. 680; MIREILLE HILDEBRANDT/BERT-JAAP KOOPS, The challenges of ambient law and legal protection in the profiling era, Mod. L. Rev. 2010, S. 428–460, S. 433 ff.

49 STEVE LOHR, Sizing up big data, broadening beyond the internet. The New York Times 19.06.2013, <<https://bits.blogs.nytimes.com/2013/06/19/sizing-up-big-data-broadening-beyond-the-internet/>> (er gibt David C. Vladeck, den ehemalige Direktor der Verbraucherschutzbehörde der Federal Trade Commission, einer unabhängigen Behörde der US-Regierung, wieder).

50 HACKER (Fn. 24), S. 1150.

51 KURT PÄRLI, Vertragsfreiheit, Gleichbehandlung und Diskriminierung im privatrechtlichen Arbeitsverhältnis, Habil. Bern 2009, Rz. 1385 f., 1594. Siehe auch Leading Case BGE 129 III 276.

die Aufgabe zu, den Diskriminierungsschutz im Rahmen des bestehenden Persönlichkeitsschutzes zu verwirklichen. Bis anhin hat der Bundesrat den Erlass eines allgemeinen Antidiskriminierungsgesetzes stets abgelehnt.⁵² In Deutschland gibt es hingegen ein Allgemeines Gleichbehandlungsgesetz (AGG), in dem für Private ein Diskriminierungsverbot verankert ist (§ 7 Abs. 1 AGG).⁵³

Nach dem allgemeinen arbeitsrechtlichen Diskriminierungsverbot in Art. 328 OR dürfen Algorithmen im Arbeitskontext nicht so programmiert sein, dass sie *direkt* diskriminieren. Sie dürfen auch nicht *indirekt* diskriminieren, d.h. trotz neutraler Regelung benachteiligende Auswirkungen für verschiedene Gruppen von Arbeitnehmern (beruhend auf Rasse, Alter, Geschlecht, Nationalität usw.) entfalten, es sei denn, dies sei sachlich begründet und verhältnismässig.⁵⁴ Diskriminierungsklagen im Zusammenhang mit der Nutzung von People Analytics werden eher indirekte Diskriminierungen betreffen.⁵⁵

Durch die Anknüpfung des allgemeinen arbeitsrechtlichen Diskriminierungsverbots an den Begriff der Persönlichkeit greift das schweizerische Diskriminierungsschutzrecht sehr weit. Es schützt nicht (nur) eine bestimmte gesellschaftliche Gruppe vor Diskriminierung, sondern dem Grundsatz nach jede Person.⁵⁶ Damit könnte es vom *Anwendungsbereich* her eigentlich ein geeignetes Instrument zur Regulierung algorithmischer Diskriminierungen sein. Trotzdem sollte man unseres Erachtens die Hoffnung auf die Lösung des Problems hierzulande nicht auf das allgemeine Diskriminierungsschutzrecht setzen.

Das geltende schweizerische Diskriminierungsschutzrecht genügt aus zwei Gründen nicht: Erstens ist ein indirekter Diskriminierungsschutz *schwierig zu beweisen*;⁵⁷ eine Beweislastumkehr zugunsten des Diskriminierungsopfers

fehlt. Zweitens bestehen *kaum abschreckende Sanktionen* gegen diskriminierendes Verhalten.⁵⁸ Der Bundesrat hat erst im Mai 2016 festgestellt, dass der Diskriminierungsschutz im Privatrecht lückenhaft ist.⁵⁹ Er hat berichtet, die geringe Zahl der Gerichtsfälle zu Diskriminierungsproblemen könnte darauf hindeuten, dass die bestehenden Rechtsinstrumente für Betroffene entweder zu wenig bekannt oder zu kompliziert sind. Zudem könnten die verfahrensrechtlichen Hindernisse dazu führen, dass der bestehende Diskriminierungsschutz seine Wirksamkeit zu wenig entfalten kann.

Das allgemeine arbeitsrechtliche Diskriminierungsverbot wird ergänzt durch *spezialgesetzliche Diskriminierungsverbote*. Diese schützen aber nur sehr punktuell:

1. Im Gleichstellungsgesetz ist jede direkte oder indirekte Diskriminierung aufgrund des *Geschlechts* untersagt (Art. 3 GIG) und eine Beweislastleichterung ist vorgesehen (Art. 6 GIG), sodass eine Glaubhaftmachung genügt.⁶⁰
2. Das Bundesgesetz über die Beseitigung von Benachteiligungen von Menschen mit *Behinderungen* (BehiG) gilt nur für Arbeitsverhältnisse des Bundes und sieht dort ein Diskriminierungsverbot und die Pflicht zum Ergreifen angemessener Vorkehrungen vor.⁶¹ Das BehiG klammert den Bereich privatrechtlicher Arbeitsverhältnisse ausdrücklich aus. In der Schweiz besteht deshalb für private Arbeitgeber keine ausdrückliche Regelung zum Schutz von Behinderten vor Diskriminierung.⁶² Ganz allgemein gibt es in der Schweiz kein Recht auf Beseitigung einer Benachteiligung mit Blick auf Behinderte, es sei denn, es bestehe eine Pflicht zum Ergreifen von Massnahmen zum Schutz der Persönlichkeit (Art. 328 OR und Art. 6 Abs. 1 ARG).⁶³
3. Das Gesetz über genetische Untersuchungen beim Menschen (GUMG) bezweckt u.a. den Schutz vor genetischer Diskriminierung (Art. 4 GUMG). Um *genetische Diskriminierung* handelt es sich, wenn eine Person einer

52 Zuletzt Schweizerischer Bundesrat, Recht auf Schutz vor Diskriminierung: Bericht des Bundesrates in Erfüllung des Postulats Naef 12.3543 vom 14.06.2012, Bern 25.05.2016, S. 16. Auch der National- und der Ständerat haben Vorstösse zur Schaffung eines Gleichbehandlungsgesetzes/Antidiskriminierungsgesetzes bisher stets abgelehnt. Vgl. Ip. Bea Heim (09.3242 «Diskriminierungsschutz»), Pa. Iv. Rechsteiner (07.422 «Allgemeines Gleichbehandlungsgesetz»).

53 DZIDA/GROH (Fn. 2), S. 1917 ff.; BAGE 137, 80 = NZA 2011, S. 1361.

54 BGE 142 V 316, E. 6.1.2; BVerwGer A-6157/2014 vom 19.12.2016 E. 4.2.3; KGer LU 1C 17 41 vom 03.04.2018, 2018 I Nr. 3 (sachliche Rechtfertigung bei einer Anstellungsdiskriminierung).

55 KING/MRKONICH (Fn. 12), S. 564; IFEOMA AJUNWA/SORELLE A. FRIEDLER/CARLOS E. SCHEIDEGGER/SURESH VENKATASUBRAMANIAN, Hiring by algorithm: predicting and preventing disparate impact, 01.04.2016, <http://sorelle.friedler.net/papers/SSRN-id2746078.pdf>; BAROCAS/SELBST (Fn. 9), S. 693.

56 ULLIN STREIFF/ADRIAN VON KAENEL/ROGER RUDOLPH, Arbeitsvertrag, Praxiskommentar zu Art. 319–362 OR, Art. 328 N7; WOLFGANG PORTMANN/ROGER RUDOLPH, in: Heinrich Honnelli/Nedim Peter Vogt/Wolfgang Wiegand, Basler Kommentar, Obligationenrecht I, Art. 1–529 OR, 6. Aufl., Basel 2015, Art. 328 N 43 und 45; vgl. zum Persönlichkeitsschutz nach Art. 328 Abs. 1 OR: MANFRED REHBINDER/JEAN-FRITZ STÖCKLI, Berner Kommentar zu den Art. 319–330b OR, Bern 2010, Art. 328 N 4 ff.

57 Siehe dazu ausführlich PÄRLI (Fn. 51), Rz. 1256 ff.; JAKOB UEBERSCHLAG, Die Anstellungsdiskriminierung aufgrund des Geschlechts im privatrechtlichen Arbeitsverhältnis (Art. 3 Abs. 2 GIG), Diss. Zürich 2009, S. 49 ff. Zur Anstellungsdiskriminierung siehe ArbGer ZH vom 13.01.2006, AN 050401/U1; BVerwGer T304.021563 vom 01.06.2005, E. B.c.; vgl. BVerwGer A-6157/2014 vom 19.12.2016 E. 4.2.5. Gerichtsfälle indirekter Anstellungsdiskriminierung

gen sind äusserst selten, was u.a. auf die fehlende Beweislastleichterung zurückzuführen ist, siehe PÄRLI (Fn. 51), Rz. 1429 ff.; UEBERSCHLAG (Fn. 57), S. 250 ff.

58 PÄRLI (Fn. 51), Rz. 1597; Schweizerischer Bundesrat (Fn. 52), S. 17 ff.

59 Schweizerischer Bundesrat (Fn. 52), S. 2.

60 JEAN-PHILIPPE DUNAND, L'interdiction de la discrimination à l'embauche dans la loi fédérale sur l'égalité (LEg), in: Jean-Philippe Dunand/Karine Lempen/Pascal Mahon (Hrsg.), L'égalité entre femmes et hommes dans les relations de travail 1996–2016: 20 ans d'application de la LEg, Genf 2016, S. 35–72, S. 69.

61 Büro für Arbeits- und sozialpolitische Studien BASS AG, Evaluation des Bundesgesetzes über die Beseitigung von Benachteiligungen von Menschen mit Behinderungen – BehiG, Bern 08.2015, S. 50 f.; BVerwGer A6550/2007 vom 29.04.2008, E. 7.

62 BASS (Fn. 61), S. 63 f.; kritisch dazu Kurt PÄRLI/ANNETTE LICHTENAUER/ALEXANDRA CAPLAZI, Literaturanalyse Integration in die Arbeitswelt durch Gleichstellung, 28.07.2008, https://www.edi.admin.ch/dam/edi/de/dokumente/gleichstellung/bericht/studie_integrationsindiarbeitsweltdurchgleichstellung.pdf.download.pdf/studie_integrationsindiarbeitsweltdurchgleichstellung.pdf, Rz. 4.20 und 4.21.

63 BGer 4A.102/2008 vom 27.5.2008; BGer 4A.2/2014 vom 19.2.2014. Siehe auch PÄRLI/LICHTENAUER/CAPLAZI (Fn. 62), Rz. 4.21.

negativen Behandlung ausgesetzt ist und zwar nicht als Ergebnis der physischen Manifestation einer Krankheit oder einer Behinderung, sondern allein wegen ihres Erbguts.⁶⁴ Das Verbot, eine Person wegen ihres Erbguts zu diskriminieren, richtet sich sowohl an staatliche Organe wie auch an Private. Das GUMG wird derzeit revidiert; mit einem Inkrafttreten der Revision ist nicht vor 2021 zu rechnen.⁶⁵

4. Das Freizügigkeitsabkommen der Schweiz mit der EU und ihren Mitgliedsstaaten (FZA) verbietet die Diskriminierung europäischer *Wanderarbeitnehmer* bei den Beschäftigungs- und Arbeitsbedingungen einschliesslich der Anstellung (Art. 2, Art. 7 lit. a FZA sowie Anhang I Art. 9 Abs. 1 und 4 FZA).⁶⁶ Das FZA sieht aber weder Beweislastleichterungen noch Sanktionsnormen vor.

Der arbeitsrechtliche Diskriminierungsschutz in der Schweiz bietet keine befriedigende Lösung, um das Problem möglicher Diskriminierungen durch Algorithmen anzugehen. Zum einen fehlen griffige Regelungen, zum andern bestehen Durchsetzungshindernisse. Es gibt zwar vereinzelte gesetzliche Diskriminierungs-Sondertatbestände, aber diese sind nur sehr punktuell nützlich.

2. Diskriminierungsschutzrecht als geeignetes Instrument?

Wäre es denn grundsätzlich denkbar, das schweizerische Diskriminierungsschutzrecht anders auszugestalten und das Problem möglicher Diskriminierungen durch Algorithmen mittels eines (modifizierten) Diskriminierungsschutzrechts zu lösen?

In den USA ist das Diskriminierungsschutzrecht weit entwickelt, weshalb wir als Erstes einen Blick darauf werfen wollen. Um eine Klage wegen indirekter Diskriminierung im Zusammenhang mit prädiktiver Analytik erfolgreich vorzubringen, muss ein Kläger in den USA aufzeigen, dass ein spezifischer Algorithmus in der Anwendung dazu führt, dass eine nach amerikanischem Diskriminierungsrecht geschützte Gruppe schlechter abschneidet als die Mehrheit, dass also z.B. schwarze oder behinderte Bewerber bei der Selektion für eine

64 ISABELLE WILDHABER, Genetische und medizinische Informationen in der Arbeitswelt, Jusletter 06.12.2010, Rz. 1–48; IFEOMA AJUNWA, Genetic testing meets big data: tort and contract law issues, Ohio St. L. J. 2014, S. 1225–1262.

65 Bundesamt für Gesundheit, Revision Gesetz über genetische Untersuchungen beim Menschen, <<https://www.bag.admin.ch/bag/de/home/medizin-und-forschung/genetische-untersuchungen/aktuelle-rechtsetzungsprojekte1.html>>.

66 In der Lehre wird vertreten, dass das FZA für private Arbeitgeber direkt verbindlich sei: KURT PÄRLI, Eurolohn: Berufung auf das Diskriminierungsverbot ist rechtsmissbräuchlich, Jusletter 20.05.2019, Rz. 28–30. Das Bundesgericht liess die Frage der Drittwirkung jüngst offen: Urteil des Bundesgerichts 4A.230/2018 vom 15.01.2019, E. 2.6.

ausgeschriebene Stelle schlechter abschneiden als die Mehrheit.⁶⁷ In den USA befürworten die meisten Gerichte statistische Methoden zur Feststellung einer indirekten Diskriminierung.⁶⁸ Die Anwendung von Statistik stösst in den USA jedoch gerade hinsichtlich prädiktiver Analytik in Fachkreisen auf Kritik.⁶⁹ Über die wichtige Frage, wann eine indirekte Diskriminierung zu bejahen ist, sollte nicht pauschal eine Statistik entscheiden. Die schweizerische⁷⁰ und auch die europäische⁷¹ Rechtsprechung ziehen zur Feststellung, ob eine geschützte Gruppe wesentlich stärker benachteiligt wird, zwar auch Statistiken heran, legen jedoch keine starren Grenzwerte fest.

Kann ein Kläger den Nachweis der indirekten Diskriminierung erbringen, so kann sich der Arbeitgeber in den USA entlasten, indem er beweist, dass der angefochtene Algorithmus «*job-related and consistent with business necessity*» war.⁷² Hier stellen sich ähnliche Probleme wie unter Art. 328b OR im schweizerischen Recht (nachfolgend D.II.1.), wonach der Arbeitgeber Daten nur bearbeiten darf, soweit sie dessen Eignung für das Arbeitsverhältnis betreffen oder zur Durchführung des Arbeitsvertrages erforderlich sind. Es ist fraglich, ob die *berufliche Bezogenheit* gegeben ist, wenn der Algorithmus Daten berücksichtigt, deren Informationsgehalt in der Korrelation zwischen nicht-

67 Ricci v. DeStefano, 557 U.S. 557, 587 (2009) (Vergleich der Erfolgsrate zwischen schwarzen und weissen Feuerwehrleuten, die die Beförderung zum Hauptmann anstreben); Jones v. City of Boston, 752 F.3d 38, 47 n.9 (1st Cir. 2014); Watson v. Ft. Worth Bank & Trust, 487 U.S. 977, 995 (1988).

68 Verbreitet ist einerseits die Methode der statistischen Signifikanz. Littler Report (Fn. 15), S. 8; siehe etwa die Entscheide Castaneda v. Partida, 430 U.S. 482 (1977); Griffin v. Bd. of Regents of Regency Univs., 795 F.2d 1281, 1291 n.19 (7th Cir. 1986); Hazelwood School District v. United States, 433 U.S. 299, 308 n.14 (1977); Watson v. Fort Worth Bank & Trust, 487 U.S. 977 (1988); vgl. auch Bennett v. Total Minatome Corp., 138 F.3d 1053, 1062 (5th Cir. 1998); Contreras v. City of Los Angeles, 656 F.2d 1267 (9th Cir. 1981); Waison v. Port Authority, 948 F.2d 1370, 1376 (2d Cir. 1991); Ottaviani v. State Univ. of N.Y. at New Paltz, 875 F.3d 365, 370–371 (1989); Sobel v. Yeshiva Univ., 839 F.2d 18 (1988). Andererseits gibt es die 4/5-Regel. Nach ihr erachten die US-Bundesbehörden es im Allgemeinen als Nachweis für eine indirekte Diskriminierung, wenn die Selektionsrate für eine geschützte Gruppe weniger als vier Fünftel (oder achtzig Prozent) derjenigen für die Gruppe mit der höchsten Rate beträgt (Section 4 lit. D Uniform Guidelines on Employee Selection Procedures).

69 Die Methode der statistischen Signifikanz wird kritisiert, weil bei prädiktiver Analytik selbst kleinste Unterschiede «statistisch signifikant» sein können. Im Entscheid Wal-Mart Stores, Inc. v. Dukes, 131 S. Ct. 2541 (2011) wurde ein Unterschied in der Zahl der beförderten Männer und Frauen von nur sieben Zehntel eines Prozentsatzes als statistisch signifikant eingestuft. Siehe dazu ALLAN G. KING, «Cross statistical disparities» as evidence of a pattern and practice of discrimination: statistical versus legal significance, The Labor Lawyer 2007, S. 271–292, S. 280. Kritiker stellen fest, dass im Zeitalter von Big Data die statistische Signifikanz die Norm und damit ein schlechter Indikator für die rechtliche Relevanz ist, Littler Report (Fn. 15), S. 11.

70 Vgl. BGE 143 II 366 E. 3.6 und 141 II 411 E. 6.2; STÉPHANIE PERRENOUD, Durées du travail et discrimination, AJP 2017, S. 657–682.

71 Vgl. Urteile des EuGH vom 08.05.2019 C-161/18, Villar Láziz, ECLI:EU:C:2019:382, Rz. 39 ff. und vom 28.02.2013 C-427/11 Kenny, ECLI:EU:C:2013:122, Rz. 42 ff.; ROMÉI/RUGGIÉRI (Fn. 31), S. 591.

72 Title VII of the Civil Rights Act of 1964, 42 U.S.C. Sec. 2000e-2(k)(1)(A)(i).

arbeitsbezogenen Daten und der Arbeitsleistung liegt,⁷³ z.B. in der Korrelation zwischen Manga-Affinität und Programmierfähigkeiten. Zwar betrifft die Liebe zu Manga die Eignung für das Arbeitsverhältnis im Sinne einer Korrelation, sie ist aber nicht kausal für gutes Programmieren und sie bezieht sich auf eine Freizeitbeschäftigung. Ausserdem ist nicht klar, wie lange die Korrelation anhalten wird.

Wenn es dem Arbeitgeber gelingt, legitime Geschäftsgründe für die Verwendung des Algorithmus aufzuzeigen, so muss der Kläger eine *alternative Praxis* darlegen, die ebenso effizient wäre, den legitimen Geschäftsanforderungen des Arbeitgebers zu entsprechen, die aber *weniger negative Auswirkungen* auf seine geschützte Gruppe hätte.⁷⁴ Es liegt in den USA somit nicht am Arbeitgeber, die Verhältnismässigkeit des eingesetzten Tools darzulegen und zu beweisen. Dieser Schritt wird für den klagenden Arbeitnehmer schwierig sein, da er sich mit alternativen Praktiken wenig auskennen wird, und Arbeitgeber, die People Analytics verwenden, argumentieren werden, dass deren Nutzung an legitime Geschäftszwecke gebunden ist – zur Verbesserung von Produktivität, Bindung und Rentabilität.⁷⁵

Das amerikanische Recht verbietet, wie die meisten Rechtsordnungen, Diskriminierungen basierend auf gewissen Kriterien, wie Rasse, Geschlecht, Behinderung, Alter, sexuelle Ausrichtung oder Religion.⁷⁶ Diskriminierungen können im Zeitalter von Big Data aber auch viele andere Eigenschaften betreffen. So können z.B. Arbeitgeber in Versuchung geraten, über Fitness-Armbänder kostenverursachende Mitarbeiter zu identifizieren und dann gezielt abzubauen. Viele der von Arbeitgebern gesponserten Wellnessprogramme konzentrieren sich auf Gewichtsabnahme und Raucherentwöhnung; Adipositas und Raucherstatus sind nach den Antidiskriminierungsgesetzen jedoch keine geschützten Kategorien.⁷⁷ In den USA ist eine Diskriminierung aufgrund des Ge-

sundheitszustandes eines potenziellen Arbeitnehmers erlaubt, solange er nicht an einer Behinderung im Sinne des Americans with Disabilities Act (ADA) leidet. Algorithmische Diskriminierungen sprengen somit die Kategorien der gesetzlich geschützten Personengruppen.

Ähnliches kann für das *europäische Diskriminierungsschutzrecht* gesagt werden. Auszugehen ist ebenso davon, dass eine direkte Diskriminierung selten sein wird, vielmehr wird es sich um indirekte und versehentliche Diskriminierungen handeln. Für die gerichtliche Geltendmachung einer indirekten Diskriminierung liegt der erste Schritt – wie im amerikanischen Recht – im Nachweis eines besonderen Nachteils für die geschützte Gruppe. Auch im europäischen Recht wird dabei oftmals die Statistik entscheidend sein.⁷⁸ Teilweise wird in der Lehre argumentiert, dass ein besonderer Nachteil als nachgewiesen gelten kann, wenn die Wahrscheinlichkeit, dass ein Mitglied einer geschützten Gruppe durch den Algorithmus positiv bewertet wird, also z.B. für eine Anstellung selektioniert wird, höchstens 75% der Wahrscheinlichkeit eines Mitglieds der Mehrheit beträgt.⁷⁹

Das EU-Antidiskriminierungsrecht ist aus drei Gründen *schlecht gerüstet*, um mit Diskriminierung bei algorithmischen Entscheidungen umzugehen: Erstens werden algorithmische Prozesse vielfach den materiellen und persönlichen Anwendungsbereich der bestehenden EU-Antidiskriminierungsgesetze sprengen.⁸⁰ Zweitens wird die Diskriminierung einfach zu rechtfertigen sein, denn die mögliche Rechtfertigung einer indirekten Diskriminierung ist in ihrer eigentlichen Definition enthalten: Eine indirekte Diskriminierung ist eine ungünstige Behandlung einer geschützten Gruppe durch eine per se neutrale Bestimmung, ein neutrales Kriterium oder eine neutrale Praxis «unless that provision, criterion or practice is objectively justified by a legitimate aim and the means of achieving that aim are appropriate and necessary».⁸¹ Die angebliche prädiktive Genauigkeit von Modellen bietet also eine einfache Rechtfertigung für diskriminierende Effekte.⁸² Bei der direkten Diskriminierung im Arbeitskontext hingegen ist eine Rechtfertigung wesentlich schwieriger, das Unterscheidungsmerkmal muss eine massgebliche berufliche Eigenschaft dar-

73 Littler Report (Fn. 15), S. 19; GABRIEL KASPER/ISABELLE WILDHABER, Big Data am Arbeitsplatz, Datenschutz- und arbeitsrechtliche Herausforderungen von People Analytics in Schweizer Unternehmen, in: Ueli Kieser/Kurt Pärli/Ursula Uttinger (Hrsg.), Datenschutztagung 2018, Zürich/St. Gallen 2019, S. 189–232, S. 202.

74 Littler Report (Fn. 15), S. 36.

75 WILSON/BELLIVEAU/GRAY (Fn. 3), S. 20.

76 USA: Title VII of the Civil Rights Act of 1964, 42 U.S.C. Sec. 2000e-2(k); Americans with Disabilities Act of 1990 (ADA); Genetic Information Nondiscrimination Act of 2008 (GINA), Pregnancy Discrimination Act of 1978 (PDA), Age Discrimination in Employment Act of 1967 (ADEA), Fair Credit Reporting Act of 1970 (FCRA). EU: Art. 1 Richtlinie 2000/78/EG des Rates vom 27.11.2000 zur Festlegung eines allgemeinen Rahmens für die Verwirklichung der Gleichbehandlung in Beschäftigung und Beruf, L 303/16. Deutschland: § 1 Abs. 1 Allgemeines Gleichbehandlungsgesetz vom 14.08.2006, BGBl. I S. 1897 (AGG; erfasst werden Rasse, ethnische Herkunft, Geschlecht, Religion, Weltanschauung, Behinderung, Alter und sexuelle Identität, nicht aber z.B. der Wohnort oder die soziale Herkunft).

77 IFOMA AJUNWA/KATE CRAWFORD/JASON SCHULTZ, Limitless worker surveillance, California Law Review 2017, S. 735–776; JESSICA L. ROBERTS, Healthism and the law of employment discrimination, Iowa L. Rev. 2014, S. 571–636, S. 595 ff.

78 SANDRA FREDMAN, Discrimination Law, Oxford 2011, S. 154 ff., S. 183 ff.

79 HACKER (Fn. 24), S. 1153.

80 Dazu detailliert HACKER (Fn. 24), S. 1154 ff.

81 Art. 2 lit. b Richtlinie 2004/113/EG des Rates vom 13.12.2004 zur Verwirklichung des Grundsatzes der Gleichbehandlung von Männern und Frauen beim Zugang zu und bei der Versorgung mit Gütern und Dienstleistungen, L 373/37; siehe auch Art. 2 Abs. 1 lit. b Richtlinie 2000/43/EG des Rates vom 29.06.2000 zur Anwendung des Gleichbehandlungsgrundsatzes ohne Unterschied der Rasse oder der ethnischen Herkunft, L 180/22; Art. 2 Abs. 1 lit. b Richtlinie 2006/54/EG des Europäischen Parlaments und des Rates vom 05.07.2006 zur Verwirklichung des Grundsatzes der Chancengleichheit und Gleichbehandlung von Männern und Frauen in Arbeits- und Beschäftigungsfragen (Neufassung), L 204/23.

82 HACKER (Fn. 24), S. 1160 ff.

stellen.⁸³ Das ist nur dann der Fall, wenn die Stelle nicht hätte besetzt werden können, wenn kein Bewerber die erforderliche Eigenschaft besessen hätte.⁸⁴ Im Gegensatz zur indirekten Diskriminierung genügt es also nicht, wenn die Bewerber mit der geforderten Eigenschaft besser abschneiden; vielmehr müssen Bewerber, die die Eigenschaft nicht besitzen, für die Stelle völlig ungeeignet sein.⁸⁵

Drittens wird der Kläger mit Beweisproblemen zu kämpfen haben. Der europäische Gesetzgeber hat diese Beweisprobleme vorausgesehen und spezifische Bestimmungen aufgenommen, welche die Beweislast teilweise umkehren und so die Schwierigkeiten des Klägers mildern, aber nicht lösen.⁸⁶ Der Kläger muss nur Fakten nachweisen, von denen angenommen werden kann, dass sie zu einer direkten oder indirekten Diskriminierung geführt haben; es obliegt dann dem Beklagten zu beweisen, dass keine Verletzung des Gleichbehandlungsgrundsatzes vorliegt.⁸⁷ Aber selbst dieser erleichterte Beweis setzt voraus, dass der Kläger eine statistische Diskrepanz zwischen den Auswirkungen des algorithmischen Prozesses auf die geschützte und eine vergleichbare Gruppe nachweist.⁸⁸ Das wiederum verlangt Kenntnisse der algorithmischen Ergebnisse und der Gruppenzugehörigkeit der untersuchten Kandidaten – mithin Informationen, die dem Kläger im Allgemeinen fehlen. Ausserdem muss der Kläger die eventuelle Rechtfertigung des Beklagten widerlegen und die Fehlerhaftigkeit des Modells nachweisen. Dies erweist sich als eine fast unmögliche Aufgabe angesichts von Modellen, die selbst für ihre Entwickler oft unergründlich sind, zumal der Kläger keinen Zugriff auf die Trainingsdaten und Modellspezifikationen hat. Diese Faktoren schaffen ein beachtliches *Rechtsdurchsetzungsproblem*, das wiederum die Anreize für Compliance seitens der Entwickler und Anwender verringert.

Ein naheliegender Weg für die Zukunft wäre die *Änderung des europäischen Antidiskriminierungsrechts*, um Instrumente zur Schliessung dieser Lücken aufzunehmen. So könnte z.B. die öffentliche Durchsetzung, die in Umfang und Kompetenzen stark eingeschränkt ist⁸⁹ und die aufgrund der Unterausstat-

tung der Antidiskriminierungsbehörden bis anhin noch keine spürbaren Auswirkungen zeigte,⁹⁰ verstärkt werden. Es ist jedoch eher unwahrscheinlich, dass in absehbarer Zeit die notwendigen legislativen Schritte unternommen werden.⁹¹

Unser Blick auf das US-amerikanische Antidiskriminierungsrecht und auf das europäische Gegenstück ergibt, dass beide keine einfache Lösung für Fälle von Diskriminierung durch Algorithmen bieten. Die Erfassung algorithmischer Diskriminierungen durch bestehende Antidiskriminierungsbestimmungen erfordert oft eine weitreichende und wohlwollende Interpretation der Anwendbarkeit der Bestimmungen. Insgesamt erweist sich das Diskriminierungsschutzrecht allein als ungeeignetes Instrument zur Verhinderung von Ungleichbehandlungen durch den Einsatz von Algorithmen am Arbeitsplatz.

II. Datenschutzrecht am Arbeitsplatz

Zwischen dem Diskriminierungs- und dem Datenschutz besteht ein *Zusammenhang*:⁹² Das Diskriminierungsschutzrecht bezweckt, Menschen vor Herabwürdigung, Ausgrenzung und Benachteiligung aufgrund von Persönlichkeitsmerkmalen zu schützen. Datenschutzrechtliche Bestimmungen stellen ebenfalls Instrumente zur Verwirklichung der Ziele des Diskriminierungsschutzes dar.⁹³ Diese Gemeinsamkeit wird vermehrt auch im anglo-amerikanischen Raum anerkannt.⁹⁴ Der Datenschutz ist eine spezifische Antwort des Gesetzgebers auf die besonderen Bedrohungen der Persönlichkeit, die durch die technologische Entwicklung in der Datenverarbeitung möglich wurden.⁹⁵ In der digitalen Welt ist ein integriertes Verständnis dieser Bestimmungen notwendig, weshalb wir den Blick über das Diskriminierungsschutzrecht hinaus auf das Datenschutzrecht erweitern.

83 Art. 4 Abs. 1 Richtlinie 2000/43/EG (Fn. 81); Art. 4 Abs. 1 Richtlinie 2000/78/EG (Fn. 76); Art. 14 Abs. 2 Richtlinie 2006/54/EG (Fn. 81); zu beachten ist, dass dies auch für Fälle von rassistischer und ethnischer Diskriminierung gilt; die Beschränkung auf die Beschäftigung ist ausdrücklich nur in der zuletzt genannten Bestimmung über die Diskriminierung aufgrund des Geschlechts enthalten.

84 GREGOR THÜSING, § 8 AGG, in: Franz J. Säcker/Roland Rixecker/Hartmut Oetker, Münchener Kommentar zum Bürgerlichen Gesetzbuch (BGB), 7. Aufl., München 2015, Rz. 10.

85 BAROCAS/SELBST (Fn. 9), S. 711.

86 Art. 8 Richtlinie 2000/43/EG (Fn. 81); Art. 10 Richtlinie 2000/78/EG (Fn. 83); Art. 9 Richtlinie 2004/113/EG (Fn. 81); Art. 19 Richtlinie 2006/54/EG (Fn. 81).

87 HACKER (Fn. 24), S. 1167 ff.

88 Urteil des EuGH vom 17.10.1989 109/88 Danfoss, ECLI:EU:C:1989:383, Rz. 16; vgl. auch Urteil des EuGH vom 27.10.1993 C-127/92 Enderby, ECLI:EU:C:1993:859, Rz. 19.

89 Art. 13 Richtlinie 2000/43/EG (Fn. 81); Art. 20 Richtlinie 2006/54/EG (Fn. 81); Art. 12 Richtlinie 2004/113/EG (Fn. 81).

90 EVELYN ELLIS/PHILIPPA WATSON, EU anti-discrimination law, 2. Aufl., Oxford 2012, S. 507.

91 HACKER (Fn. 24), S. 1171; ELLIS/WATSON (Fn. 90), S. 503.

92 KURT PÄRLI, Datenschutz, in: Wolfgang Portmann/Adrian von Kaenel (Hrsg.), Fachhandbuch Arbeitsrecht, Zürich 2018, S. 685–724, Rz. 17.10; KARINE LEMPEN, Protection des données et discrimination lors de la procédure de recrutement, in: Jean-Philippe Dunand/Pascal Mahon (Hrsg.), La protection des données dans les relations de travail, Genf 2017, S. 269–286; RAPHAËL GELLERT/KATJA DE VRIES/PAUL DE HERT/SERGE GUTWIRTH, A comparative analysis of anti-discrimination and data protection legislations, in: Bart Custers/Toon Calders/Bart Schermer/Tal Zarsky (Hrsg.), Discrimination and privacy in the information society, Berlin/Heidelberg 2013, S. 61–90, S. 61.

93 BVerwGer vom 10.4.2012, A-4457/2011, E. 9.3; HACKER (Fn. 24), S. 1171.

94 Z.B. MARK BURDON/PAUL HARPUR, Re-conceptualising privacy and discrimination in an age of talent analytics, U.N.S.W.L.J. 2014, S. 679–712; GELLERT/DE VRIES/DE HERT/GUTWIRTH (Fn. 92), S. 81.

95 FRANK SEETHALER, Entstehungsgeschichte DSG, in: Urs Maurer-Lambrou/Gabor-Paul Blechta (Hrsg.), Basler Kommentar zum Datenschutzgesetz/Öffentlichkeitsgesetz, 3. Aufl., Basel 2014, N 13 ff.; PÄRLI (Fn. 92), Rz. 17.7.

Wenn nun also das Diskriminierungsschutzrecht algorithmische Diskriminierungen nicht zuverlässig eliminieren kann, wie sieht es dann mit der datenschutzrechtlichen Regulierung aus? Können Beschränkungen bei der Erhebung und Bearbeitung personenbezogener Daten den Risiken einer Diskriminierung entgegenwirken? Selbst ein starker Datenschutz wird die Bedenken hinsichtlich einer algorithmischen Diskriminierung am Arbeitsplatz nicht vollständig zerstreuen.⁹⁶

1. Schweizerisches Datenschutzrecht

Das Datenschutzrecht ist ein wichtiges Instrument zum Schutz der Privatsphäre. Das *schweizerische Datenschutzrecht de lege lata* regelt den Bereich Arbeitsplatz insbesondere in den Art. 328 und 328b OR, flankiert von den allgemeinen Datenschutzprinzipien im Bundesgesetz vom 19.06.1992 über den Datenschutz (DSG, SR 235.1), wie Verhältnismässigkeit, Transparenz und Zweckbindung sowie von Spezialgesetzen, z.B. dem Bundesgesetz vom 08.10.2004 über genetische Untersuchungen beim Menschen (GUMG, SR 810.12). Relevant bei prädiktiver Analytik am Arbeitsplatz ist auch Art. 26 Abs. 1 der Verordnung 3 vom 18.08.1993 zum Arbeitsgesetz (ArGV3, SR 822.113). Diese Bestimmung schränkt die Nutzung von Überwachungs- und Kontrollsystemen ein, welche das Verhalten der Arbeitnehmer am Arbeitsplatz überwachen und welche nicht durch betriebliche oder andere anerkannte Zwecke gerechtfertigt werden können.⁹⁷ Dabei handelt es sich aber primär um ein Instrument der Arbeitssicherheit und Gesundheitsvorsorge.

Gemäss Art. 328b OR darf der Arbeitgeber Daten über den Arbeitnehmer nur bearbeiten, soweit sie dessen *Eignung für das Arbeitsverhältnis betreffen oder zur Durchführung des Arbeitsvertrages erforderlich* sind. Der Begriff der Eignung kann auch Charakter, Weltanschauung und Ehrgeiz umfassen.⁹⁸ Allerdings ist eine generelle Persönlichkeitsdurchleuchtung im Sinne einer Ergründung des Wesens des Arbeitnehmers unzulässig.⁹⁹ Art. 328b OR beschränkt somit die zulässige Datenbearbeitung für den Arbeitgeber im Sinn des Verhältnismässigkeitsgrundsatzes, wie er auch in Art. 4 DSG verankert ist.¹⁰⁰ Die Verhältnismäs-

sigkeit erfordert, dass die Betriebsinteressen und die unternehmerische Freiheit des Arbeitgebers das Persönlichkeits- und Gesundheitsschutzinteresse des Arbeitnehmers überwiegen. Es ist in unseren Augen fraglich, ob die von Art. 328b OR geforderte *berufliche Bezogenheit* gegeben ist, wenn der Algorithmus Daten berücksichtigt, deren Informationsgehalt in der Korrelation zwischen nicht-arbeitsbezogenen Daten und der Arbeitsleistung liegt.¹⁰¹

Des Weiteren fällt People Analytics nicht in den *Geltungsbereich des DSG*, wenn keine Bearbeitung von *Personendaten* stattfindet (Art. 3 lit. a DSG). Grundsätzlich zulässig erscheinen Datenbearbeitungen, die aufgrund technischer Massnahmen nicht personenbezogen erfolgen; ebenso Bearbeitungen aggregierter Daten oder Bearbeitungen, bei denen Personen bloss singularisierbar, nicht identifizierbar sind.¹⁰²

Durch eine Beschränkung der Datenbearbeitung kann der Datenschutz jedenfalls gewisse Formen von People Analytics verhindern, wie wir an anderer Stelle bereits analysiert haben.¹⁰³ Aber ist der Datenschutz an sich geeignet, Diskriminierungen durch Algorithmen zu verhindern? Bis zu einem gewissen Grad können Datenschutzrechte direkte Diskriminierungen verhindern. So verbietet z.B. das GUMG Arbeitgebern, genetische Informationen über Bewerber und Arbeitnehmer zu erfragen oder anderweitig bewusst zu erwerben (Art. 21 GUMG), und verbindet dies mit einem expliziten Diskriminierungsverbot (Art. 4 GUMG). Das DSG bietet einen erhöhten Schutz vor der Bearbeitung von Daten zu diskriminierungsrelevanten Tatbeständen (besonders schützenswerte Personendaten, Art. 3 lit. c DSG).¹⁰⁴ Diesen Schutz kann ein Analytisches System jedoch auf dem Weg der beschriebenen Proxy-Diskriminierung unterlaufen.¹⁰⁵

2. Datenschutzrecht als geeignetes Instrument?

Ein Blick auf das europäische Datenschutzrecht, insbesondere die europäische Datenschutz-Grundverordnung (DSGVO), lohnt sich. Die DSGVO enthält heute schon Bestimmungen, welche den regulatorischen Schwerpunkt verlagern, und zwar von der Ex-post-Korrektur auf Ex-ante-Regeln, welche eine

⁹⁶ KIM (Fn. 33), S. 898.

⁹⁷ ISABELLE WILDHABER, Die Roboter kommen – Konsequenzen für Arbeit und Arbeitsrecht, ZSR 2016, S. 315–352, S. 346 ff.; ISABELLE WILDHABER, Robotik am Arbeitsplatz: Robo-Kollegen und Robo-Bosse, AJP 2017, S. 213–224; MELINDA F. MÜLLER (heute: Lohmann), Roboter und Recht, AJP 2014, S. 595–609, S. 606 f.

⁹⁸ GABRIEL AUBERT, Art. 328b CO, in: Luc Thévenoz/Franz Werro (Hrsg.), Commentaire romand du code des obligations I, Art. 1–529 CO, 2. Aufl., Basel 2012, N 3; STREIFF/VON KAENEL/RUDOLPH (Fn. 56), Art. 328b N 5.

⁹⁹ DZIDA (Fn. 19), S. 544–545; STREIFF/VON KAENEL/RUDOLPH (Fn. 56), Art. 328b N 10.

¹⁰⁰ Mit der Verankerung dieses Grundsatzes in Art. 4 Abs. 2 DSG wurde der Verhältnismässigkeitsgrundsatz für die Datenbearbeitung ins Bundeszivilrecht übernommen, siehe URS MAURER-LAMBROU/RETO STEINER, Art. 4 DSG, in: Urs Maurer-Lambrou/Gabor-Paul Blechta (Hrsg.), Basler Kommentar zum Datenschutzgesetz/Öffentlichkeitsgesetz, 3. Aufl., Basel 2014, N 10;

zur Konkretisierung in Art. 328b OR siehe WOLFGANG PORTMANN/JEAN-FRITZ STÖCKLI, Schweizerisches Arbeitsrecht, 3. Aufl., Zürich 2013, N 438.

¹⁰¹ KIM/HANSON (Fn. 14), 17; JEANNE MEISTER, The year social HR matters, Forbes 06.01.2014, <<http://www.forbes.com/sites/jeannemeister/2014/01/06/2014-the-year-social-hrmatters/#4b005ef362dc>>; PECK (Fn. 11).

¹⁰² Vgl. DZIDA (Fn. 19), S. 543; KASPER/WILDHABER (Fn. 73), S. 213 f.; die Problematik der Singularisierbarkeit ist jedoch, soweit ersichtlich, noch nicht höchstrichterlich geklärt worden.

¹⁰³ KASPER/WILDHABER (Fn. 73), S. 208 f.

¹⁰⁴ Die entsprechende europäische Bestimmung (Art. 9 DSGVO) wird «informationelles Diskriminierungsverbot» genannt, DAEELOW (Fn. 46), S. 4.

¹⁰⁵ CUSTERS (Fn. 29), S. 9–10; ein einfaches Verbot der Bearbeitung von sensiblen Daten verhindert Diskriminierungen nicht vollständig; es kann sie sogar noch verschlimmern: KIM (Fn. 33), S. 898 f.

unfaire Datenverarbeitung von Anfang an verhindern sollen.¹⁰⁶ In Verbindung mit dem europäischen Antidiskriminierungsrecht bietet die DSGVO deshalb einen *vielfersprechenden Weg* zur Bekämpfung der Diskriminierung durch Algorithmen. Dies gilt nicht nur auf der materiellen Ebene, sondern auch auf der Vollzugsebene, wo der Vollstreckungsapparat der DSGVO aktiv genutzt werden könnte und sollte.

Das europäische Datenschutzrecht verfügt schon über einige *Instrumente zur Öffnung der algorithmischen Blackbox*. Wir verstehen unter einer algorithmischen Blackbox ein komplexes Entscheidungsfindungssystem, von welchem nur das äussere Verhalten, nicht aber die innere Struktur sichtbar ist, oder – mit Bezug auf unsere obige Abb. 1 – es sind nur Input und Output bekannt, das Modell bleibt undurchsichtig. Zur Öffnung der algorithmischen Blackbox könnten z.B. die individuellen Auskunftsrechte von Art. 15 Abs. 1 DSGVO genutzt werden, um Informationen über das Vorhandensein algorithmischer Diskriminierungen zu verlangen. Im Entscheid *Meister* stellte der EuGH fest, dass sich in Fällen von Diskriminierung am Arbeitsplatz ein solches Auskunftsrecht nicht aus dem Antidiskriminierungsrecht ergebe.¹⁰⁷ Anstatt auf Antidiskriminierungsrecht könnte sich ein solcher Auskunftsanspruch nun auf Datenschutzrecht (Art. 15 Abs. 1 DSGVO) stützen. Diese unmittelbaren Auskunftsrechte sind auf die personenbezogenen Daten des einzelnen Nutzers beschränkt. Um einen Anschein von Diskriminierung zu etablieren, muss der potenzielle Kläger Informationen zu den Ergebnissen der algorithmischen Auswertung verschiedener Gruppen haben.

Ein Ansatz für eine Steuerung algorithmenbasierter Verfahren und zur Verhinderung von algorithmischen Diskriminierungen könnte ein *Abwehrrecht gegen automatisierte Entscheidungsverfahren* sein, wie es sich in Art. 22 DSGVO findet.¹⁰⁸ Der normative Radius dieser Bestimmung ist aber insgesamt eng: Erstens lässt er zahlreiche Durchbrechungen des Grundsatzes zu, und zweitens erfasst er nur Entscheidungen, die ohne jeglichen menschlichen Einfluss zustande gekommen sind (nicht aber die algorithmenbasierte Unterstützung menschlicher Entscheidungen).¹⁰⁹ Art. 22 DSGVO stellt uns noch vor ein weiteres Problem. Findet die prädiktive Analyse in einer Blackbox statt, können Transparenzpflichten ein taugliches Regulierungsinstrument sein, um Licht ins Dunkel zu bringen. Das den Algorithmen zu Grunde liegende Modell ist regelmässig undurchsichtig, nicht intuitiv und für Menschen schwer verständlich.¹¹⁰ Nur wer eine Rechtsverletzung, die ein algorithmenbasiertes Verfahren verursacht, zu erkennen und zu beweisen vermag, kann sich wirksam wehren. Transparenz

106 HACKER (Fn. 24), S. 1171.

107 Urteil des EuGH vom 09.04.2012 C-415/10 *Meister*, ECLI:EU:C:2012:217, Rz. 46.

108 KROLL et al. (Fn. 38), S. 685–690.

109 MARIO MARTINI, Algorithmen als Herausforderung für die Rechtsordnung, *JuristenZeitung* 2017, S. 1017–1025, S. 1020.

110 Vgl. Littler Report (Fn. 15), S. 36.

kann durch ein Recht auf Erklärung gewährleistet werden. Wir sind aber der Meinung, dass sich weder aus Art. 13–15 DSGVO noch aus Art. 22 DSGVO ein Recht auf eine spezifische Einzelfallerklärung ableiten lässt.¹¹¹ Vielmehr sieht die DSGVO nur eine allgemeine Erklärung der Systemfunktionalität und gewisse flankierende Massnahmen vor.

Die nationalen Datenschutzbehörden sollten ausserdem auf *algorithmische Audits und Folgenabschätzungen* zum Datenschutz gemäss Art. 58 Abs. 1 (b) und Art. 35 ff. DSGVO zurückgreifen, um die Ursachen von Diskriminierungen aufzudecken und angemessene Massstäbe der algorithmischen Fairness durchzusetzen. Fortschritte in der Informatik ermöglichen die Erkennung und Minimierung von Diskriminierung auf Daten- und Code-Ebene¹¹². Weitere nicht-individuelle Ansätze der DSGVO, wie z.B. Datenschutz durch Technik und Datenschutz-Zertifizierungen (Art. 25 und Art. 42 DSGVO), können helfen, Algorithmen verantwortungsvoll, erklärbar und menschenzentriert zu gestalten.¹¹³ Solche kollektiv ausgerichtete Rechtsbehelfe sind erforderlich, da die Vorstellung, dass Gleichbehandlung, Privatsphäre und Transparenz individuelle Rechte darstellen, heute nur noch relativiert gilt. In diesem Sinn kann das europäische Gebot des Datenschutzes durch Design einer «Gleichbehandlung durch Design» dienen. Diese technischen Ex-ante-Lösungen allein genügen jedoch nicht. Ex-post-Überprüfungen der Algorithmen auf ihre Auswirkungen auf geschützte Gruppen sind ebenfalls erforderlich.¹¹⁴

Auch das europäische Datenschutzrecht konnte das Problem der Diskriminierung durch Algorithmen bis anhin nicht vollständig auffangen. Denn es verhindert viele der in diesem Aufsatz beschriebenen Ursprünge von algorithmischen Diskriminierungen nicht, z.B. die indirekte Proxy-Diskriminierung.¹¹⁵ Deshalb werden Neuerungen vorgeschlagen, wie das Konzept eines «Right to a Reasonable Inference» von WACHTER und MITTELSTADT.¹¹⁶ Demnach sollten Arbeit-

111 SANDRA WACHTER/BRENT MITTELSTADT/LUCIANO FLORIDI, Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation, *International Data Privacy Law* 2017, S. 76–99, S. 83 f.

112 HACKER (Fn. 24), S. 1184; vgl. Executive Office of the President, 2016 (Fn. 33), S. 18.

113 LILIAN EDWARDS/MICHAEL VEALE, Slave to the algorithm? Why a «right to an explanation» is probably not the remedy you are looking for, *Duke L. & Tech. Rev.* 2017, S. 18–84, S. 18; EDWARDS fordert anlässlich der von Prof. Dr. Isabelle Wildhaber und Prof. Dr. Melinda Lohmann organisierten Tagung *Algocracy* vom 09.11.2018 auch die systemischen, nicht-individuellen Rechtsinstrumente aus Arbeits- und Gleichbehandlungsrecht sowie die Informationsfreiheit zu berücksichtigen.

114 PAULINE T. KIM, Auditing algorithms for discrimination, *U. Pa. L. Rev. Online* 2017, S. 189–203, S. 202.

115 «Technical tools alone cannot solve the problem. [...] these tools do not and cannot address many of the reasons that automated decision processes may be biased in the first places»: KIM (Fn. 114), S. 194.

116 SANDRA WACHTER/BRENT MITTELSTADT, A right to reasonable inferences: re-thinking data protection law in the age of big data and AI, *Colum. Bus. L. Rev.* 2019, S. 494–620, S. 617 f.; DAEDELLOW (Fn. 46), Rz. 32.

geber im Voraus nachweisen, dass eine mit Big-Data-Analytik zu ziehende Schlussfolgerung, die für Einzelpersonen hohe Risiken birgt, angemessen ist. Diese Ex-ante-Begründung würde durch einen zusätzlichen Ex-post-Mechanismus unterstützt, der es erlaubt, unangemessene Schlussfolgerungen anzufechten.

Das U.S.-Datenschutzrecht bietet auf Bundesstaatsebene kaum Schutz für Arbeitnehmer in Bezug auf Überwachungen in privatrechtlichen Arbeitsverhältnissen.¹¹⁷ Die entsprechenden Regeln finden sich einerseits im Fall- und Richterrecht, andererseits im Gesetzesrecht. Das Fall- und Richterrecht des Common Law setzt für eine erfolgreiche arbeitnehmerseitige Klage eine berechnete Erwartung von Privatsphärenschutz voraus, welche sich aus den äusseren Umständen und den Auffassungen der Gesellschaft oder einer vernünftigen Person ergibt.¹¹⁸ Arbeitgeber können aber alle Erwartungen durch einseitige Erklärung zerstören, indem sie etwa in Datenschutzerklärung auf die Überwachung hinweisen.¹¹⁹ Das Gesetzesrecht auf Bundesstaatsebene regelt den Datenschutz sektorspezifisch, d.h. die bundesstaatlichen Normen gelten nur für bestimmte Unternehmen und bestimmte Arten von Daten.¹²⁰ Dadurch unterscheidet sich die amerikanische Rechtsordnung grundlegend von den europäischen Datenschutzgesetzen, die sämtliche Personendaten schützen.¹²¹ Generell sind die U.S.-Bestimmungen über Arbeitnehmer-Gesundheitsdaten restriktiver,¹²² während bezüglich der Bearbeitung aller anderen Arbeitnehmerdaten (z.B. Leistungsdaten) grosse Freiheiten bestehen.¹²³

Trotz der fragmentierten und unvollständigen Datenschutzgesetzgebung weist die U.S.-Praxis im Vergleich zur europäischen ein robustes Datenschutz-Management auf, wie eine 2015 veröffentlichte Fallstudie ergeben hat.¹²⁴ Un-

ternehmen in den Vereinigten Staaten verstünden den Datenschutz als generelle Fairness gegenüber Mitarbeitern und als Risikomanagement.¹²⁵ Amerikanische Datenschutzbeauftragte seien professionell tätig und hätten eine einflussreiche und relativ unabhängige Stellung im Unternehmen.¹²⁶ Der *hohe gelebte Datenschutz-Standard* sei darauf zurückzuführen, dass das regulatorische Umfeld sich durch offene, unklare gesetzliche Bestimmungen auszeichne, die Mitsprache von verschiedenen, auch unternehmens-externen Interessenvertretern (Gewerkschaften, Aufsichtsbehörden, Öffentlichkeit) beim Datenschutz fördere und erhöhte Transparenzpflichten vorsehe.¹²⁷ Trotz dieses robusten Datenschutz-Managements scheint das amerikanische Gesetzesrecht Personendaten im Angesicht der immer komplexer werdenden Analysen am Arbeitsplatz allerdings ungenügend zu schützen. Diesbezüglich besteht eine Parallele zum ungenügenden Schutz des schweizerischen Diskriminierungsrechts vor algorithm-basierten Diskriminierungen. Aus der amerikanischen Datenschutz-Praxis lässt sich für den Diskriminierungsschutz in der Schweiz jedoch mitnehmen, dass eine Sensibilisierung im ganzen Unternehmen für eine weit verstandene Gerechtigkeit beim Einsatz von Algorithmen wichtig ist und genügend Ressourcen für die verantwortlichen Stellen geschaffen werden müssen.

Das schweizerische Datenschutzrecht konzentriert sich auf die Regulierung der *Datenbearbeitung* (vgl. Art. 4 DSG). Es sieht die Zweckbeschränkung und Transparenzpflicht (Art. 4 Abs. 3 DSG) sowie Zugriffsbeschränkungen (vgl. Art. 7 Abs. 1 DSG) vor. So wirkt das Datenschutzrecht durchaus prohibitiv für die Anwendung von People Analytics, es limitiert aber auch die Nutzung der Vorteile der Datenanalytik. Der Ansatz über die Datenbearbeitung ist in einer Welt automatisierter und vernetzter Datenbanken und Informationsnetze immer schwieriger durchzusetzen. Im Fokus des Gesetzgebers sollte unseres Erachtens nicht mehr (nur) die Beschaffung und Bearbeitung von Daten stehen, sondern die Auswirkung der Bearbeitung auf die Persönlichkeit der Betroffenen. Es ist an der Zeit, im Datenschutzrecht die Frage zu stellen, *wie Daten genutzt werden können und sollen*.¹²⁸ Als Reaktion auf die Herausforderungen, die Big Data mit sich bringt, schlagen Datenschutz-Wissenschaftler Ansätze vor, die über die Regelung der klassischen Datenbearbeitung hinausgehen. Der Datenschutz

117 LOTHAR DETERMANN/ROBERT SPRAGUE, *Intrusive monitoring: employee privacy expectations are reasonable in Europe, destroyed in the United States*, Berkeley Tech.L. J. 2011, S. 979–1036, S. 1016; ungenügender Schutz durch den Electronic Communications Privacy Act (ECPA): ebd., S. 1001.

118 DETERMANN/SPRAGUE (Fn. 117), S. 981.

119 DETERMANN/SPRAGUE (Fn. 117), S. 981.

120 DETERMANN/SPRAGUE (Fn. 117), S. 993–994; Executive Office of the President, 2014 (Fn. 4), S. 18; ANN CAVOUKIAN, *Privacy by design in law, policy and practice. A white paper for regulators, decision-makers and policy-makers*, Toronto 08.2011, <http://www.ontla.on.ca/library/repository/mon/25008/312239.pdf>, S. 23. Es gibt unzählige bereichsspezifische U.S.-Gesetze: CAVOUKIAN (Fn. 120), S. 22; allen voran müssen Arbeitgeber den Americans with Disabilities Act (ADA), Genetic Information Nondiscrimination Act (GINA) und Health Insurance Portability and Accountability Act (HIPPA) berücksichtigen (PAV AKHTAR/PHOEBE MOORE, *The psychosocial impacts of technological change in contemporary workplaces, and trade union responses*, Internat. J. of Labour Research 2016, S. 101–131, S. 117) sowie den Health Information Technology for Economic and Clinical Health (HITECH) Act und Fair Credit Reporting Act (FCRA) einhalten.

121 DETERMANN/SPRAGUE (Fn. 117), S. 980; vgl. Littler Report (Fn. 15), S. 17.

122 Littler Report (Fn. 15), S. 12.

123 Littler Report (Fn. 15), S. 12.

124 KENNETH A. BAMBERGER/DEIRDRE K. MULLIGAN, *Privacy on the ground*, Cambridge, Massachusetts 2015, S. 6.

125 BAMBERGER/MULLIGAN (Fn. 124), S. 6.

126 BAMBERGER/MULLIGAN (Fn. 124), S. 12.

127 BAMBERGER/MULLIGAN (Fn. 124), S. 15.

128 World Economic Forum, *Unlocking the value of personal data: from collection to usage*, Cologny/Genf 02.2013, <http://perma.cc/74VK-9AGZ>, S. 3; KATE CRAWFORD/JASON SCHULTZ, *Big data and due process: toward a framework to redress predictive privacy harms*, Boston College L. Rev. 2014, S. 93–128, S. 109; ROGER ALLAN FORD, *Unilateral invasions of privacy*, Notre Dame L. Rev. 2016, S. 1075–1115, S. 1105–1106; vgl. auch die Aussagen von Prof. Dr. Florent Thouvenin im Interview mit: SERMIN FAKI/ANJA WURM, *Daten-Professor Florent Thouvenin verteidigt Facebook, Google und Co.: «Die beißen doch nicht!», Blick* 30.05.2018; WACHTER/MITTELSTADT (Fn. 116), S. 615f.; BART CUSTERS, *Data mining with discrimination sensitive and privacy sensitive attributes*, SSRN 2010, S. 31–38.

sollte nicht nur den Zugang zu personenbezogenen Daten einschränken, sondern vielmehr den nachfolgenden *Informationsfluss als solchen regeln*.¹²⁹ Zum Beispiel empfehlen CRAWFORD und SCHULTZ einen Prozess, der Einzelpersonen berechtigt zu wissen, wann prädiktive Analysen verwendet werden, und die Fairness des Prozesses in Frage zu stellen.¹³⁰ CITRON und PASQUALE plädieren dafür, dass die Regulierungsaufsicht die Fairness der Bewertungssysteme gewährleisten sollte.¹³¹ Aus diesem Grund wollen wir in einem dritten Teil unserer rechtlichen Überlegungen (nachfolgend III.) über mögliche Ansätze zu einer «Algorithm Governance» nachdenken, welche über das Diskriminierungsschutzrecht (oben I.) und das Datenschutzrecht (oben II.) hinausgehen.

III. Denkbare Ansätze zu einer Governance von Algorithmen

Bei der sog. «Algorithm Governance» geht es darum, Instrumente zu finden, um die mit der Verwendung von Algorithmen verbundenen Risiken und Gefahren zu minimieren. Da algorithmische Entscheidungen nicht nur Wirkungen auf den Einzelnen, sondern auch auf das Kollektiv haben können, sollten Governance-Instrumente auch die gesamtgesellschaftlichen Risiken adressieren. Ein einheitliches Allheilmittel ist nicht zu erwarten, vielmehr wird eine Kombination verschiedener Ansätze sinnvoll sein.

Technische Ansätze könnten dabei helfen, algorithmische Risiken zu minimieren.¹³² Analog zu der in Art. 25 DSGVO verankerten Strategie des «Datenschutzes durch Technikgestaltung» (sog. Privacy by Default und Privacy by Design) wäre ein Diskriminierungsschutz durch Technikgestaltung wünschenswert, um algorithmische Systeme möglichst nichtdiskriminierend zu konzipieren. Ferner gibt es bereits technische Möglichkeiten wie das «Data Repair», um indirekte Diskriminierungen bei Algorithmen, die auf diesen Daten basieren, auszuschließen.¹³³

129 NEIL RICHARDS/JONATHAN KING, Big data ethics, Wake Forest L. Rev. 2014, S. 393–432, S. 411–412.

130 CRAWFORD/SCHULTZ (Fn. 128), S. 109.

131 CITRON/PASQUALE (Fn. 6), S. 20 ff.

132 FLORIAN SAURWEIN/NATASCHA JUST/MICHAEL LATZER, Governance of algorithms: options and limitations, info 2015, S. 35–49, S. 39.

133 MICHAEL FELDMAN/SORELLE A. FRIEDLER/JOHN MOELLER/CARLOS SCHEIDEGGER/SURESH VENKATASUBRAMANIAN, Certifying and removing disparate impact, Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM 2015, Section 4, S. 259–268, S. 261 ff.; IAN AYRES/JENNIFER GERARDA BROWN, Mark(et)ing nondiscrimination: privatizing ENDA with a certification mark, Mich.L. Rev. 2006, S. 1639–1712, S. 1639 ff.; KIM, Data-driven (Fn. 33), S. 935; SARA HAJIAN/JOSEF DOMINGO-FERRER, Direct and indirect discrimination prevention methods, in Bart Custers/Toon Calders/Bart Schermer/Tal Zarsky (Hrsg.), Discrimination and privacy in the information society, Berlin/Heidelberg 2013, S. 241–254, 247–251; FAISAL KAMIRAN/TOON CALDERS/MYKOLA PECHENIZKIY, Techniques for discrimination-free predictive models, in Bart Custers/Toon Calders/

Die Machine-Learning-Community arbeitet sodann seit Längerem daran, das Problem der mangelnden Nachvollziehbarkeit algorithmischer Entscheidungen zu lösen, etwa im Rahmen des 2016 lancierten Projekts zu «*Explainable AI*» (XAI) der amerikanischen Defense Advanced Research Projects Agency (DARPA).¹³⁴ Unter dem Titel «Explainable AI» wird diskutiert, wie undurchsichtige Machine-Learning-Modelle erklärbar gemacht werden können.¹³⁵ Dem Nutzer sollen konkret einzelne Entscheidungsschritte erläutert, die allgemeinen Stärken und Schwächen des Modells vermittelt und ein Verständnis dafür erweckt werden, wie sich das System in Zukunft verhalten wird.¹³⁶ Herausfordernd ist speziell der Trade-Off zwischen der Vorhersagegenauigkeit eines Systems und dessen Nachvollziehbarkeit: Künstliche neuronale Netze und Deep Learning Verfahren weisen tendenziell eine hohe Vorhersagegenauigkeit auf, sind aber (noch) schwer nachvollziehbar.¹³⁷

Nicht jede Form der Erklärung ist somit (technisch) umsetzbar und/oder zielführend. Fraglich ist bspw., wie eine Erklärung konkret ausgestaltet sein müsste, um tatsächlich einen Mehrwert für den Empfänger zu schaffen – die Antwort auf diese Frage wird wesentlich vom Empfängerkreis abhängen, der jedoch in vielen Fällen nicht homogen sein wird. Selbstverständlich müssen bei der Diskussion um Transparenz auch berechnete Interessen der Unternehmen berücksichtigt werden, dies ist bereits aus volkswirtschaftlichen Überlegungen zur Innovationsförderung angezeigt. Nicht zu vernachlässigen sind schliesslich Anwendungsfelder, in denen Einblicke in die Systemfunktionalität geradezu kontraproduktiv wären, weil damit das bewusste Ausnutzen der Funktionalität im eigenen Interesse (sog. «Algorithm Gaming») ermöglicht würde.¹³⁸ Zu denken ist etwa an einen Algorithmus, der entscheidet, welche Steuerveranlagungen genauer überprüft oder welcher Fluggast einer weiteren Sicherheitskontrolle unterzogen werden soll.

Ein weiterer Governance-Ansatz ist die (freiwillige) *Selbstregulierung*, d.h. die Einigung auf branchenweite Regelungen, denen sich die Algorithmen verwendenden Unternehmen unterstellen.¹³⁹ Der Vorteil dieses Ansatzes ist, dass branchenspezifisches Wissen nutzbar gemacht und auf Entwicklungen flexibel

Bart Schermer/Tal Zarsky (Hrsg.), Discrimination and privacy in the information society, Berlin/Heidelberg 2013, S. 223–240, 229 ff.

134 MATT TUREK, Explainable artificial intelligence (XAI), <https://www.darpa.mil/program/explainable-artificial-intelligence>.

135 MELINDA F. LOHMANN, Black Box öffne dich – Transparenzpflichten bei Algorithmen, F.A.Z. Einspruch 22.08.2018, abrufbar unter <http://einspruch.faz.net/einspruch-magazin>.

136 LOHMANN (Fn. 135).

137 LOHMANN (Fn. 135); differenzierend zur sog. «Interpretierbarkeit» ZACHARY C. LIPTON, The mythos of model interpretability, Commun. ACM 2018, S. 36–43.

138 Vgl. MARIJN JANSSEN/GEORGE KUK, The challenges and limits of big data algorithms in technocratic governance, Government Information Quarterly 2016, S. 371–377, S. 372.

139 MARTINI (Fn. 109), S. 1022 f.

reagiert werden kann.¹⁴⁰ Beispiele der Selbstregulierung¹⁴¹ sind die Schaffung von Verhaltenskodizes oder Ethik-Richtlinien¹⁴², die Formulierung technischer und organisatorischer Minimalstandards und Zertifizierungsanforderungen¹⁴³, die Einführung eines Auditing¹⁴⁴ oder eines hippokratischen Eids für Entwickler von Algorithmen¹⁴⁵. Die Erfahrung mit freiwilliger Selbstverpflichtung zeigt allerdings, dass diese ohne steuernde Anreize und Sanktionierungsmechanismen kaum die erhoffte Wirkung entfaltet und nicht selten als Feigenblatt dient.¹⁴⁶ Griffiger ist deshalb vielfach die regulierte Selbstregulierung bzw. die Ko-Regulation,¹⁴⁷ womit wir beim Instrument der staatlichen Intervention an- gelangt wären.

Der Staat kann zwecks Minimierung der Risiken und Gefahren algorithmischer Entscheidungsprozesse *intervenieren*, etwa indem er Gesetze erlässt.¹⁴⁸ So wurde in Deutschland kurzzeitig die Schaffung eines digitalen Antidiskriminierungsgesetzes diskutiert, dessen Nutzen und Praktikabilität angesichts des diffusen Regelungsgegenstandes und der rapiden technologischen Entwicklung indes fraglich gewesen wäre.¹⁴⁹ Denkbar wären ferner die Einführung einer Kennzeichnungspflicht für algorithmenbasierte Verfahren in persönlichkeits- sensiblen Feldern¹⁵⁰, die Einführung von Risikomanagementsystemen¹⁵¹ und die Einführung eines staatlichen Kontrollverfahrens für bestimmte algorithmen- basierte Verfahren, wobei ein Ex-ante-Regulierungssystem zur Steuerung von Algorithmen entwickelt werden könnte, wie es derzeit für die Zulassung von Medikamenten vor dem Markteintritt verwendet wird.¹⁵² Wie wir vorangehend aus datenschutzrechtlicher Perspektive gesehen haben, lassen sich die mit der

140 MARTINI (Fn. 109), S. 1022 f.

141 Vgl. SAURWEIN/JUST/LATZER (Fn. 132), S. 40.

142 Siehe etwa die Ethik-Richtlinien von Google (<https://ai.google/principles>) oder der Telekom (<https://www.telekom.com/de/konzern/digitale-verantwortung/details/ki-leitlinien-der-telekom-523904>).

143 Siehe etwa das TÜV Süd und DFKI-Projekt openGENESIS zur Schaffung einer Zertifizierung für KI im Kontext des autonomen Fahrens, <https://www.tuvsud.com/en/industries/mobility-and-automotive/automotive-and-oem/autonomous-driving/opengenesi>.

144 KIM (Fn. 114), S. 189 ff.; CITRON/PASQUALE (Fn. 6), S. 21.

145 LORI SHERER, Data scientists, take a hippocratic oath, Forbes Online 13.06.2018, <https://www.forbes.com/sites/baininsights/2018/06/13/data-scientists-take-a-hippocratic-oath>.

146 MARTINI (Fn. 109), S. 1023.

147 WOLFGANG HOFFMANN-RIEM, Rechtliche Rahmenbedingungen für und regulative Herausforderungen durch Big Data, in: Wolfgang Hoffmann-Riem (Hrsg.), Big Data – Regulative Herausforderungen, Baden-Baden 2018, S. 11–80, 68; MARTINI (Fn. 109), S. 1023.

148 SAURWEIN/JUST/LATZER (Fn. 132), S. 40 f.

149 Vgl. PATRICK BEUTH, Maas schlägt digitales Antidiskriminierungsgesetz vor, Zeit Online 03.07.2017, <https://www.zeit.de/digital/internet/2017-07/heiko-maas-algorithmen-regulierung-antidiskriminierungsgesetz>.

150 MARTINI (Fn. 109), S. 1020.

151 MARTINI (Fn. 109), S. 1022.

152 ANDREW TUTT, An FDA for algorithms, Admin. L. Rev. 2017, S. 83–123 (für eine Bundesaufsichtsbehörde plädierend, die die Sicherheit und Wirksamkeit von Algorithmen gewährleistet, bevor sie auf den Markt kommen).

Verwendung von Algorithmen verbundenen Risiken durch gesetzliche Vorschriften bereits aus Praktikabilitätsgründen nur begrenzt bewältigen.¹⁵³ Eine sinnvolle staatliche Intervention wäre jedenfalls das Lancieren von Informations- und Sensibilisierungsmassnahmen für die Bürger sowie Arbeitgeber und Arbeitnehmer (sog. Governance durch Information).¹⁵⁴

E. Fazit

Prädiktive Analytik am Arbeitsplatz (People Analytics) hilft Arbeitgebern, den Betrieb zu optimieren. Die im Analyseprozess eingesetzten Algorithmen können jedoch zu Diskriminierungen gewisser Arbeitnehmer führen.¹⁵⁵ Die Ursachen der Diskriminierung können sich auf die Inputdaten, das Analysemodell oder den Output beziehen. Soweit systeminhärente Fehler die Chancen für bereits benachteiligte geschützte Gruppen verringern, sollte dies die gleichen Bedenken hinsichtlich der Gleichbehandlung am Arbeitsplatz wecken, die in der Vergangenheit gewisse Antidiskriminierungsgesetze motivierten.¹⁵⁶

Das traditionelle Diskriminierungsschutzrecht vermag Diskriminierungen durch Algorithmen nicht zu greifen: Zum einen erschweren Beweisschwierigkeiten und fehlende abschreckende Sanktionen die Durchsetzung des allgemeinen arbeitsrechtlichen Diskriminierungsverbots, zum andern vermitteln die vorhandenen Diskriminierungs-Sondertatbestände nur punktuellen Schutz. Auch nach Betrachtung des amerikanischen und europäischen Regelwerks ergibt sich, dass das Anti-Diskriminierungsrecht allein nicht dazu geeignet ist, Diskriminierungen durch Algorithmen zu mildern.

Das mit dem Diskriminierungsschutzrecht zusammenhängende Datenschutzrecht bietet vielversprechende Ansätze, um den Diskriminierungsrisiken von Algorithmen zu begegnen. So hält es etwa einen wirksame(re)n Vollstreckungsapparat bereit, der auch kollektiv ausgerichtete Rechtsbehelfe umfasst. Aber das Datenschutzrecht bietet nur limitierte Auskunfts- und Erklärungsrechte, da es sich tendenziell auf einen bestimmten «Täter» und einzelne «Opfer» konzentriert. Zudem orientiert es sich an den Bearbeitungsprozessen statt an den Auswirkungen auf die Persönlichkeit der Betroffenen.

Aufgrund des mangelhaften Rechtsschutzes *de lege lata* sind neue Formen der Regulierung (Governance) von Algorithmen erforderlich. Unseres Erachtens kommt die gesetzliche Verankerung eines Diskriminierungsschutzes durch Technikgestaltung in bestimmten Konstellationen als Lösung infrage. Zu drängen ist allerdings auch auf Selbstregulierung und Eigenverantwortung: Arbeit-

153 SAURWEIN/JUST/LATZER (Fn. 132), S. 43.

154 SAURWEIN/JUST/LATZER (Fn. 132), S. 40 f.

155 KIM (Fn. 33), S. 865.

156 KIM (Fn. 33), S. 857.

geber sollten in der Unternehmenspolitik dem Thema Fairness beim Einsatz von Algorithmen das nötige Gewicht verleihen und genügend Ressourcen für die verantwortlichen Stellen schaffen. Prädiktive Analytik sollte angesichts der verschiedenen Risiken sorgfältig verwendet werden. So sollte sichergestellt werden, dass die Auswahl von Daten repräsentativ ist. Dazu muss u.a. die Frage gestellt werden, ob es Gruppen von Arbeitnehmern gibt, die ihre Daten nicht teilen, oder die keinen Zugang zu Technologien wie Social Media haben. Sodann sollten Verzerrungen im Datenmodell gesucht und eliminiert werden. Die prädiktiven Systeme müssen ausserdem kontinuierlich überprüft werden, um einen möglichst fairen Output zu garantieren. Mit einer gesunden Sensibilisierung für die Thematik und angemessenen Sicherheitsmassnahmen kann die innovative Technologie der prädiktiven Analyse tatsächlich der Verbesserung von Arbeitsplätzen dienen.

Zusammenfassung

Prädiktive Analytik am Arbeitsplatz (People Analytics) hilft Arbeitgebern, den Betrieb zu optimieren. Die eingesetzten Algorithmen können jedoch Arbeitnehmer diskriminieren. Das Diskriminierungsschutzrecht vermag Diskriminierungen durch Algorithmen nicht zu erfassen, dies wegen Beweisschwierigkeiten, fehlender abschreckender Sanktionen und nur punktuell anwendbarer Diskriminierungs-Sondertatbestände. Das mit dem Diskriminierungsschutzrecht zusammenhängende Datenschutzrecht bietet nur limitierte Auskunfts- und Erklärungsrechte. Zudem orientiert es sich an den Bearbeitungsprozessen statt an den Auswirkungen auf die Persönlichkeit der Betroffenen. Aufgrund des mangelhaften Rechtsschutzes sind neue Formen der Regulierung von Algorithmen erforderlich. Unseres Erachtens kommt die gesetzliche Verankerung eines Diskriminierungsschutzes durch Technikgestaltung in Frage. Zu drängen ist allerdings auch auf Selbstregulierung und Eigenverantwortung. Mit einer gesunden Sensibilisierung für die Thematik und angemessenen Sicherheitsmassnahmen kann die innovative Technologie der prädiktiven Analyse der Verbesserung von Arbeitsplätzen dienen.

Résumé

L'analyse prédictive au le lieu de travail (People Analytics) aide les employeurs à optimiser leurs opérations. Cependant, les algorithmes utilisés peuvent être discriminatoires à l'égard des employés. La législation sur la protection contre la discrimination n'est pas en mesure d'empêcher la discrimination par algo-

rithmes en raison de difficultés de preuve, de l'absence de sanctions dissuasives et de l'application sélective de motifs spéciaux de discrimination. La loi sur la protection des données qui est liée à la législation sur la protection contre la discrimination n'offre que des droits limités en matière d'information et d'accès. En outre, elle est axée sur les processus de traitement des données plutôt que sur les effets sur la personnalité des personnes concernées. En raison de l'absence de protection juridique, de nouvelles formes de régulation des algorithmes sont nécessaires. A notre avis, l'ancrage juridique de la protection contre la discrimination dès la conception technologique pourrait être envisagé. Toutefois, l'autorégulation et la responsabilité personnelle sont également nécessaires. Avec une prise de conscience saine du problème et des mesures de sécurité appropriées, la technologie innovante d'analyse prédictive peut contribuer à l'amélioration des lieux de travail.