ECONSTOR
*Make Your Publications Visible.*

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Göritz, Lorena; Beinke, Jan Heinrich; Stattkus, Daniel; Thomas, Oliver

**Article — Published Version**

# Diversity Governance: Gestaltung eines Diversitätssensiblen Ansatzes zur Automatisierten Identifikation und Reduktion Impliziter Diskriminierung in der Unternehmenskommunikation

HMD Praxis der Wirtschaftsinformatik

*Suggested Citation:* Göritz, Lorena; Beinke, Jan Heinrich; Stattkus, Daniel; Thomas, Oliver (2024) : Diversity Governance: Gestaltung eines Diversitätssensiblen Ansatzes zur Automatisierten Identifikation und Reduktion Impliziter Diskriminierung in der Unternehmenskommunikation, HMD Praxis der Wirtschaftsinformatik, ISSN 2198-2775, Springer Fachmedien Wiesbaden, Wiesbaden, Vol. 61, Iss. 6, pp. 1567-1586, https://doi.org/10.1365/s40702-024-01090-3

This Version is available at:
https://hdl.handle.net/10419/316229

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

http://creativecommons.org/licenses/by/4.0/deed.de

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

HMD Praxis der Wirtschaftsinformatik (2024) 61:1567–1586
https://doi.org/10.1365/s40702-024-01090-3

SPEKTRUM

# Diversity Governance: Gestaltung eines Diversitätssensiblen Ansatzes zur Automatisierten Identifikation und Reduktion Impliziter Diskriminierung in der Unternehmenskommunikation

**Lorena Göritz · Jan Heinrich Beinke · Daniel Stattkus · Oliver Thomas**

Eingegangen: 30. Januar 2024 / Angenommen: 5. Juni 2024 / Online publiziert: 24. Juli 2024
© The Author(s) 2024

**Zusammenfassung** Diversity-Programme verfolgen nicht nur das Ziel, wirtschaftliche Potenziale zu erschließen, sondern auch, die Entwicklung von Unternehmen hin zu einer zeitgemäßen und ethisch vertretbaren Unternehmenskultur zu fördern. Trotz der positiven Absichten der Unternehmen, Diversity-Programme umzusetzen, zeigen aktuelle Daten nur langsame Fortschritte in diesem Bereich. Mit der vorliegenden Studie wird ein innovativer Ansatz zur Förderung von Diversity in Unternehmen vorgestellt, der darauf ausgerichtet ist, implizite Diskriminierung in der Unternehmenskommunikation abzubauen. Ähnlich der *Corporate Governance* zielt die im Folgenden vorgestellte *Diversity Governance* darauf ab, die Handlungsspielräume von Mitarbeitenden und Führungskräften so zu gestalten, dass sie dazu befähigt werden, im Sinne von Diversity-Prinzipien zu kommunizieren. Im Rahmen dieser Studie werden Ergebnisse zur Konzeption und zur technischen Umsetzung einer Diversity-Governance-Software präsentiert, die Daten der Unternehmenskommunikation z. B. in Form von Texten und Bildern automatisiert auf implizite Diskriminierung analysiert. Eine automatisierte Prüfung der Unternehmenskommunikation

Lorena Göritz · Jan Heinrich Beinke · Daniel Stattkus · ✉ Oliver Thomas
Smart Enterprise Engineering, Deutsches Forschungszentrum für Künstliche Intelligenz, Hamburger Straße 24, 49084 Osnabrück, Deutschland
E-Mail: oliver.thomas@dfki.de

Lorena Göritz
E-Mail: lorena.goeritz@dfki.de

Jan Heinrich Beinke
E-Mail: jan.beinke@dfki.de

Daniel Stattkus
E-Mail: daniel.stattkus@dfki.de

Oliver Thomas
Fachgebiet Informationsmanagement und Wirtschaftsinformatik, Universität Osnabrück, Hamburger Straße 24, 49084 Osnabrück, Deutschland

durch Verfahren der Künstlichen Intelligenz (KI) bietet einen entscheidenden Vorteil hinsichtlich der Geschwindigkeit und wird bei großen Analyseeinträgen auch nicht durch menschliche Phänomene wie Ermüdungseffekte in ihrer Konsistenz beeinflusst. Neben der Prüfung ist die Software auf die Weiterentwicklung der analysierten Texte und Bilder programmiert und unterstützt dabei durch KI-basierte Empfehlungen, bspw. für alternative Formulierungen. Um Feedback zur Konzeption und zur technischen Implementierung der Software zu erhalten, wurden zwölf Interviews mit Expert:innen aus dem Themenbereich *Diversity* durchgeführt, um aus den Ergebnissen praktische Handlungsempfehlungen für die erfolgreiche Umsetzung von *Diversity Governance* abzuleiten.

**Schlüsselwörter** Corporate Governance · Diversität · Diskriminierung · Unconscious Bias · Künstliche Intelligenz · Large Language Model

## Diversity Governance: Developing a Diversity-sensitive Approach to Automatically Identify and Reduce Implicit Discrimination in Corporate Communications

**Abstract** Diversity programs aim not only to unlock economic potential, but also to support companies' development toward a contemporary and ethical corporate culture. Despite companies' positive intentions to implement diversity programs, recent data indicate slow progress in this area. Our study presents a novel and innovative approach to promoting diversity in business that aims to reduce implicit discrimination within corporate communications. Similar to corporate governance, the diversity governance we propose seeks to shape the actions of employees and managers in ways that empower them to communicate in accordance with diversity principles. In this study, we present results related to the design and technical implementation of diversity governance software designed to automatically analyze corporate communication data, for example in the form of texts and images, for implicit discrimination. An automated assessment of corporate communications through artificial intelligence offers advantages in terms of speed and consistency, unaffected by human limitations such as fatigue in the case of extensive analysis input. Beyond assessment, the software focuses on enhancing the analyzed text and images by providing AI-based recommendations, such as alternative phrases. To gather feedback on the design and technical implementation of our software, we conducted twelve interviews with experts in the field of diversity. Finally, we used our findings to derive practical recommendations for the successful implementation of diversity governance.

**Keywords** Corporate governance · Diversity · Discrimination · Unconscious bias · Artificial intelligence · Large language model

## 1 Diversity als Erfolgsfaktor

Die zunehmende Beachtung von *Diversity* in Unternehmen hat nicht nur eine ethische Dimension, sondern schlägt sich auch positiv in gesellschaftlichen und wirtschaftlichen Kennzahlen nieder. Laut einer Statistik von Statista (2024) empfanden im Jahr 2020 80 % der Frauen und 62 % der Männer, dass ein Diversity-Management in Unternehmen einen erheblichen positiven Einfluss auf den wirtschaftlichen Erfolg hat. Dies wird auch bei der Betrachtung der ethnischen Vielfalt in Unternehmen in den USA bestätigt: Laut Zheng (2023) stieg der Anteil von Minderheiten in verschiedenen Jobpositionen von 26,4 % im Jahr 1996 auf 39 % im Jahr 2018 an. Zusätzlich zeigen Zhengs Ergebnisse, dass sowohl der Anteil von Minderheiten als auch der Frauenanteil positiv mit dem Wirtschaftswachstum korrelieren. Dies lässt sich durch eine Zunahme der Beschäftigung sowie die daraus resultierenden Fähigkeiten und die zusätzlichen Perspektiven erklären, die die Innovationsfähigkeit fördern.

Darüber hinaus besteht ein positiver Zusammenhang zwischen Wirtschaftskennzahlen wie dem Bruttoinlandsprodukt und dem Verbraucherpreisindex sowie der ethnischen Vielfalt und dem Frauenanteil in Unternehmen. So erzielen Unternehmen mit einer vielfältigen Führungsebene nach Andrevski et al. (2014) nicht nur einen höheren Marktanteil, sondern auch bessere Leistungen als homogene Unternehmen. Dies wird durch die höhere Wettbewerbsintensität von Unternehmen mit ethnischer Vielfalt erklärt, da häufiger neue Wettbewerbsaktionen eingeführt werden. Mit Vielfalt geht ein vermehrter Austausch von Wissen und Informationen innerhalb des Unternehmens. Vielfältige Perspektiven und Hintergründe innerhalb der Teams eröffnen wiederum die Möglichkeit, Probleme umfassender zu analysieren, was zu fundierten Entscheidungen führt und die Innovationskraft des Unternehmens stärkt (Østergaard et al. 2011). Gerade in der heutigen Wissensgesellschaft sind innovative Ideen unerlässlich, um die Resilienz und die Wettbewerbsfähigkeit von Unternehmen zu stärken.

Zusätzlich trägt die Förderung von Vielfalt in der Unternehmenskultur dazu bei, ein inklusives Arbeitsumfeld zu schaffen, in dem sich die Mitarbeitenden akzeptiert fühlen, was sich wiederum positiv auf ihre Arbeitsmotivation und ihre Produktivität auswirken kann (Ashcraft et al. 2016). In diesem Zusammenhang gewinnen Diversity-Programme zunehmend an Bedeutung. Konzeptionell werden Diversity-Trainingsprogramme als Programme definiert, die darauf abzielen, positive Interaktionen zwischen Gruppen zu erleichtern, Vorurteile und Diskriminierung abzubauen und die Fähigkeiten, das Wissen und die Motivation von Menschen zu verbessern, mit unterschiedlichen Menschen zu interagieren (Bezrukova et al. 2012).

Diversity-Programme zielen nicht nur auf die Erschließung wirtschaftlicher Potenziale ab, sondern auch auf die Entwicklung von Unternehmen hin zu einer zeitgemäßen, ethisch vertretbaren Unternehmenskultur, die den Zielen für nachhaltige Entwicklung der Vereinten Nationen (2024) entspricht. Trotz der positiven Absichten der Unternehmen, Ziele wie Geschlechtergleichheit, weniger Ungleichheiten, menschenwürdige Arbeit und Wirtschaftswachstum umzusetzen, werden laut aktuellen Daten nur langsame Fortschritte in diesem Bereich verzeichnet. Dies wird in „The Sustainable Development Goals Report 2023“ der Vereinten Nationen (2023)

deutlich, insbesondere beim Vergleich von Europa mit Ländern in Südostasien und Subsahara-Afrika, auch wenn diese Ungleichheiten teilweise durch die Auswirkungen der Covid-19-Pandemie erklärt werden können.

Die langsamen Fortschritte und Misserfolge bei der Umsetzung von Diversity-Programmen werden auch in der *Harvard Business Review* aufgegriffen (Dobbin und Kalev 2016). Laborstudien zufolge kann bei klassischen Programmen wie obligatorischen Diversity-Schulungen und Beschwerdesystemen ein Gefühl der Überwachung vermittelt, Widerstand bei den Mitarbeitenden hervorgerufen und Voreingenommenheit aktiviert werden (Dobbin und Kalev 2016). Zudem können strukturelle Zwänge Maßnahmen zur Förderung der Vielfalt entgegenstehen (Noon 2018). Diese Ergebnisse unterstreichen die Notwendigkeit neuer Ansätze, um eine diversitätsbewusste Unternehmensausrichtung zu begünstigen. Becker (2016) beschreibt eine solche Ausrichtung als *Managing Diversity*, womit Benachteiligungen von Personengruppen gegenüber anderen Gruppen abgebaut und beseitigt werden sollen.

Im Rahmen der vorliegenden Studie wird ein Ansatz fokussiert, der vor allem Benachteiligungen in der Unternehmenskommunikation in Form von benachteiligenden Bildern und Texten adressiert. Die Sprache, die Menschen jeden Tag verwenden, spiegelt ihre internen Denkprozesse wider und ist ein Abbild davon, wie sie die Welt sehen (Pennebaker 2011). Sprache kann Stereotypen durch geschlechtsspezifische, ethnische oder kulturelle Ausdrücke verstärken, was zu Vorurteilen und Diskriminierung führen kann, die wiederum die Innovation und den Geschäftserfolg von Unternehmen beeinträchtigen können (Correll 2017). Unternehmen, die auf eine inklusive Sprache achten, fördern somit ein respektvolles und vielfältiges Arbeitsumfeld, das sich langfristig positiv auf deren Leistung und deren Ruf auswirkt. Angesichts der wachsenden gesellschaftlichen Akzeptanz von *Diversity* bei Kund:innen und Mitarbeitenden müssen Unternehmen dieses Thema frühzeitig prominent platzieren und in ihre *Corporate Governance* integrieren, um langfristig wettbewerbsfähig zu bleiben. Insbesondere die Generation Z wird in den kommenden Jahren verstärkt in den Arbeitsmarkt eintreten und entsprechende ethische Anforderungen an Unternehmen stellen (Pandita 2021). Die Ansprache dieser potenziellen neuen Fachkräfte durch eine diversitätssensible Unternehmenskommunikation ist daher von besonderer Bedeutung.

Eine ausschließlich menschliche Prüfung der Unternehmenskommunikation ist jedoch nicht nur zeitaufwändig und damit kostenintensiv, sondern ebenfalls anfällig für Fehler und Schwankungen in der Beurteilung. Eine automatisierte Überprüfung der internen und externen Unternehmenskommunikation durch KI-Verfahren ist zum einen in Bezug auf die Geschwindigkeit von Vorteil und wird zum anderen auch bei großen Analyseeingaben nicht von menschlichen Erscheinungen wie Ermüdung in ihrer Konsistenz beeinflusst. In diesem Beitrag geht es daher um die Exploration der folgenden Forschungsfrage: Wie kann KI effektiv genutzt werden, um implizite Diskriminierung in der Unternehmenskommunikation automatisiert zu identifizieren und zu reduzieren?

Zur Beantwortung dieser Forschungsfrage wurde die Diversity-Governance-Software (DG-Software) entwickelt, mit der Unternehmenskommunikation automatisiert auf implizite Diskriminierung untersucht wird. So wie implizite Einstellungen als unbewusste mentale Assoziationen zwischen einem Objekt und einem bestimmten

Merkmal definiert werden, wird implizite Diskriminierung als Diskriminierung gefasst, die außerhalb des Bewusstseins stattfindet (Bertrand et al. 2005).

Hauptziel der DG-Software ist es, Mitarbeitenden und Führungskräften die Möglichkeit zu geben, interne und externe Unternehmenskommunikation in Form von Texten und Bildern mit wenigen Klicks daraufhin zu überprüfen, ob bewusst oder unbewusst unterrepräsentierte Gruppen diskriminiert werden. Damit werden auch marginalisierte Gruppen unterschiedlicher Diversitätsmerkmale berücksichtigt, die aufgrund unbewusster Erwartungen potenziell Unsicherheiten ausgesetzt sind. Neben der Prüfung ist die DG-Software auch auf die Weiterentwicklung der analysierten Texte und Bilder ausgerichtet. Hierbei werden KI-basierte Empfehlungen z. B. in Form von alternativen Formulierungen abgegeben.

Der Beitrag ist wie folgt gegliedert: Zunächst wird die Bedeutung von Diversity im Kontext der *Corporate Governance* geklärt. Anschließend werden Ansätze zur automatisierten Identifikation und Reduktion impliziter Diskriminierung in der Unternehmenskommunikation vorgestellt. Danach werden die Konzeption und die technische Umsetzung erläutert. Es folgen die Evaluation des Projektes, die Diskussion und die Ableitung von Handlungsempfehlungen für die *Diversity Governance*. Der Beitrag schließt mit einem Fazit und einem Ausblick.

## 2 Bedeutung von Diversity im Kontext der Corporate Governance

Inzwischen wird die Bekämpfung von Diskriminierung in mehreren Gesetzgebungen thematisiert. So werden im Allgemeinen Gleichbehandlungsgesetz (AGG) die europäischen Antidiskriminierungsrichtlinien umgesetzt. Mit seiner Einführung im Jahr 2006 wird das Ziel verfolgt, Diskriminierung aufgrund von Ethnie, Geschlecht, Religion, Behinderung, Alter und sexueller Identität zu verhindern, wie in § 1 AGG festgelegt ist. Die im AGG festgeschriebenen Pflichten und Rechte für Arbeitgeber:innen und Arbeitnehmer:innen beziehen sich sowohl auf den Bewerbungsprozess als auch auf bestehende Arbeitsverhältnisse. Informationen und Ressourcen zum AGG bietet die Antidiskriminierungsstelle des Bundes (2024), die auch die Öffentlichkeit über Diskriminierung aufklärt. Sie dient auch als Anlaufstelle für Beratung und führt wissenschaftliche Untersuchungen zur Diskriminierung durch (Antidiskriminierungsstelle des Bundes 2024).

Trotz gesetzlicher Bestimmungen werden jedoch immer noch bestimmte Personengruppen diskriminiert und aufgrund von Merkmalen wie Geschlecht, Hautfarbe, ethnischer Herkunft, sexueller Orientierung oder Religionszugehörigkeit benachteiligt. Dies hat nachhaltig negative direkte und indirekte Auswirkungen auf das Unternehmen, die sich anhand verschiedener Perspektiven exemplarisch aufzeigen lassen:

1. **Verlust an Talent und Fachkenntnissen**: Diskriminierung und erfahrene Benachteiligung führen dazu, dass qualifizierte Fachkräfte gar nicht erst für ein Unternehmen gewonnen werden können oder es deshalb verlassen. Durch den Verlust von Mitarbeitenden gehen Ressourcen verloren und es fehlt der entsprechende

Austausch von Fachwissen (Barmeyer et al. 2019), was sich direkt auf die Innovationsfähigkeit und die Produktqualität auswirkt (Gao et al. 2021).

2. **Geringere Arbeitsproduktivität**: Diskriminierung wirkt sich negativ auf die Arbeitsproduktivität aus. Mitarbeitende, die sich benachteiligt fühlen, neigen dazu, demotiviert zu sein, was zu nachlassender Arbeitsleistung und geringerer Effizienz führt (Gebert 2004). Eine positive und inklusive Arbeitsumgebung hingegen fördert die Motivation von Mitarbeitenden und steigert die Arbeitsproduktivität.
3. **Schlechtere Kund:innenbindung/-ansprache:** Einer Studie von Nazir et al. (2016) zufolge begünstigt ein positives Markenimage die Loyalität, das Markenbewusstsein, die Markenerfahrung, die Zufriedenheit und die Bindung. Dabei bezieht sich das Markenimage auf die Art, wie eine Marke von Verbraucher:innen wahrgenommen wird, und umfasst Eindrücke, Überzeugungen, Gefühle, Assoziationen und Erwartungen gegenüber der Marke (Nazir et al. 2016). Umgekehrt kann Diskriminierung zu einem Verlust von Kund:innen führen, wenn diese die Werte des Unternehmens als inakzeptabel empfinden. Die Kund:innenbindung kann abnehmen, was sich wiederum negativ auf den Umsatz und den langfristigen Geschäftserfolg auswirkt. Dazu zeigt sich, dass Werbung für Geschlechtervielfalt das Glaubenssystem über eine Organisation verbessert. So wurde in vier Studien festgestellt, dass weiße Männer ein toleranteres Klima in einem Unternehmen erwarten, wenn das Unternehmen auf Geschlechtervielfalt hinweist, die sich jedoch nur auf weiße Frauen beschränkt (Wilton et al. 2019).

Dies zeigt, dass sich *Diversity* zunehmend als Schlüsselelement einer effektiven *Corporate Governance* erweist und eine entsprechende Verankerung von Diversity-Prinzipien entscheidend für die Förderung einer integrativen und verantwortungsvollen Unternehmenskultur ist. So deuten Forschungsergebnisse darauf hin, dass die Integration von *Diversity* in die Führungsebenen einen deutlichen positiven Einfluss auf die Unternehmensperformance hat (Page 2008; Andrevski et al. 2014). Dazu kann laut aktuelleren Forschungsergebnissen von Díaz-Fernández et al. (2020) die Integration von Diversity-Merkmalen im Top-Management, z. B. in Bezug auf den Bildungshintergrund und internationale Erfahrungen, zu hohen Leistungen im Unternehmen führen. Auch demografische Merkmale sowie managementbezogene Kenntnisse sind in diesem Zusammenhang von wesentlicher Bedeutung. Dabei ist zu beachten, dass die Eigenschaften von Top-Management-Teams in komplexer Weise miteinander verbunden sind und unterschiedliche Kombinationen von Diversity-Merkmalen zum Erfolg führen können. Somit tragen Unternehmen durch eine aktive Förderung und Umsetzung von *Diversity* in ihren Führungsebenen potenziell zu ihrer Wettbewerbsfähigkeit bei. Daneben ist die Integration von *Diversity* in die *Corporate Governance* auch für das Risikomanagement relevant. Durch die Berücksichtigung eines breiten Spektrums von Stimmen und Perspektiven können Unternehmen potenzielle Risiken umfassender erkennen und effektiver darauf reagieren. Darüber hinaus hilft *Diversity*, Gruppendenken in Entscheidungsprozessen zu reduzieren, so dass fundiertere Unternehmensentscheidungen getroffen werden. Um vielfältige Talente anzuziehen und zu entwickeln, setzen erfolgreiche Unternehmen wie die Allianz, SAP und die Deutsche Telekom gezielt Maßnahmen ein (Kewes

**Tab. 1** Ergebnisse der Marktanalyse

| Fokus | | Kommerzielle Anbieter | Kurzbeschreibung |
|---|---|---|---|
| Verschiedene Diversitäts-merkmale | Gesamter HR-Rekrutie-rungsprozess | Applied | Applied bietet eine Software, die Vorurteile im Bewerbungsprozess minimiert. Sie optimiert Stellenausschreibungen, anonymisiert Bewerber:innen und bewertet deren Eignung anhand von Lebensläufen und Fragen, um die besten Kandidat:innen unabhängig von Vorurteilen zu finden |
| | | TalVista | TalVista stellt eine Software zur Verringerung von Vorurteilen im Bewerbungsprozess zur Verfügung. Sie optimiert Stellenausschreibungen, strukturiert den Einstellungsprozess und anonymisiert Lebensläufe, um voreingenommene Schlussfolgerungen zu vermeiden. Anschließend werden strukturierte Vorstellungsgesprächspläne erstellt |
| | | X0PA AI | X0PA AI ermöglicht einen vorurteilsfreien Einstellungsprozess, indem Bewerber:innen anonymisiert und Urteilsverzerrungen im Bewerbungsprozess minimiert werden. Das Tool gleicht Lebensläufe mit Stellenanforderungen ab, um unvoreingenommen die besten Bewerber:innen zu identifizieren |
| | | Textkernel | Textkernel ermöglicht einen vorurteilsfreien Einstellungsprozess, indem Bewerber:innen anonymisiert und Urteilsverzerrungen im Bewerbungsprozess minimiert werden. Das Tool analysiert und bewertet Lebensläufe und fasst Fähigkeiten in Arbeitszeugnissen zusammen |
| | | Textio | Textio deckt geschlechtsspezifische Ausdrücke, rassistische Sprache, Altersvorurteile und Behindertenfeindlichkeit in Texten auf. Es bietet Verbesserungsvorschläge für Wörter und Phrasen, um eine breitere Bewerber:innenansprache zu erreichen. Zusätzlich zum gesamten HR-Rekrutierungsprozess analysiert Textio auch Performance-Feedbacks |
| | Stellenaus-schreibungen | Ongig | Ongig analysiert Stellenanzeigen auf Formulierungen, die bestimmte Personengruppen ausschließen, und macht Verbesserungsvorschläge, um eine breitere Ansprache von Bewerber:innen zu erreichen. Ein Schwerpunkt liegt dabei auf der Aufdeckung geschlechtsspezifischer Vorurteile und allgemeinen Hinweisen zur Optimierung von Stellenanzeigen (z. B. hinsichtlich ihrer Lesbarkeit) |
| | | Get-Optimal | Get-Optimal nutzt KI, um Unternehmen im Bewerbungsprozess zu unterstützen. Es verfasst inklusive Stellenausschreibungen, verteilt diese automatisch und optimiert die Reichweite für eine größere Sichtbarkeit |
| Geschlecht | Allgemein | Grammarly | Grammarly überprüft Texte auf sprachliche Richtigkeit und vermeidet geschlechtsspezifische Verzerrungen in der Autokorrektur |
| | | Scribbr | Scribbr bietet ein Online-Lektorat und Plagiatsprüfungen an und hat eine kostenpflichtige Genderprüfung als zusätzlichen Service |
| | | Language-Tool | LanguageTool prüft Texte auf Sprachkorrektheit und bietet genderneutrale Alternativen für geschlechtsbezogene Begriffe an |

2022). Dazu gehören eine fortschrittliche Personalentwicklung und die Förderung von Vielfalt.

## 3 Ansätze zur automatisierten Identifikation und Reduktion impliziter sprachlicher Diskriminierung in Unternehmen

Zur Identifizierung von Unternehmen, die speziell Lösungen zur automatisierten Erkennung impliziter sprachlicher Diskriminierung in Unternehmen anbieten, wurde eine systematische Internetrecherche mit Hilfe der indexbasierten Suchmaschine Google durchgeführt. Die Recherche konzentrierte sich auf den deutschen Markt, schloss aber auch Technologien aus anderen Ländern ein. Neben diesen kommerziellen Anbietern existieren auch nicht-kommerzielle Websites, die im Rahmen von Forschungsvorhaben zu sogenannten *agentic* und *communal words* entstanden sind[1]. Anschließend wurden die identifizierten kommerziellen Anbieter von zwei unabhängigen Personen aus dem Forschungsteam analysiert, um ihren Fokus zu bestimmen. Dabei wurden sowohl die einbezogenen Diversity-Merkmale als auch die Anwendungsdomänen der Dienstleistungen untersucht. Methodisch war die Analyse an das Vorgehen von Vogel et al. (2020) angelehnt. Eventuelle Unterschiede in den Analysen wurden gemeinsam diskutiert, um zu einer konsistenten Bewertung zu gelangen. Die Ergebnisse dieses Prozesses sind in Tab. 1 zusammengefasst. Darin werden nicht nur die Anbieter und der Fokus aufgeführt, sondern auch ihre Dienstleistungen beschrieben.

Die Ergebnisse der Recherche zeigen, dass sich einige Firmen bereits mit dem Potenzial automatisierter Ansätze zur Identifizierung und zum Abbau impliziter sprachlicher Diskriminierung am Markt auseinandersetzen. Allerdings liegt der Fokus der meisten Lösungen derzeit primär auf spezifischen Diversity-Merkmalen wie dem Geschlecht oder auf bestimmten Anwendungsbereichen, insbesondere im Bereich der Personalauswahl. Im Gegensatz zu den bisherigen Ansätzen zielt die DG-Software jedoch auf eine umfassende Förderung von Vielfalt und berücksichtigt dabei zum einen unterschiedliche Diversitätsmerkmale und adressiert zum anderen die interne und externe Unternehmenskommunikation in verschiedenen Bereichen, z. B. Schulungsunterlagen für das Personalwesen, Webseiten und Social-Media-Beiträge für das Marketing, Kund:innenkommunikation für den Vertrieb sowie textbasierte Produkte für Forschung und Entwicklung.

## 4 Konzeption und technische Implementierung

Da eine implizite sprachliche Diskriminierung unbewusst erfolgt, ist sie oft nicht direkt sichtbar und daher schwer zu erkennen. Der Mangel an Sichtbarkeit und Bewusstsein spiegelt sich in dem Begriff *Unconscious Bias* wider, der in der Forschung häufig verwendet wird. Unter *Unconscious Bias* werden tief verwurzelte, unbewusste Stereotype verstanden, die das Verhalten auf natürliche, automatisierte

[1] z. B. genderdecoder.wi.tum.de.

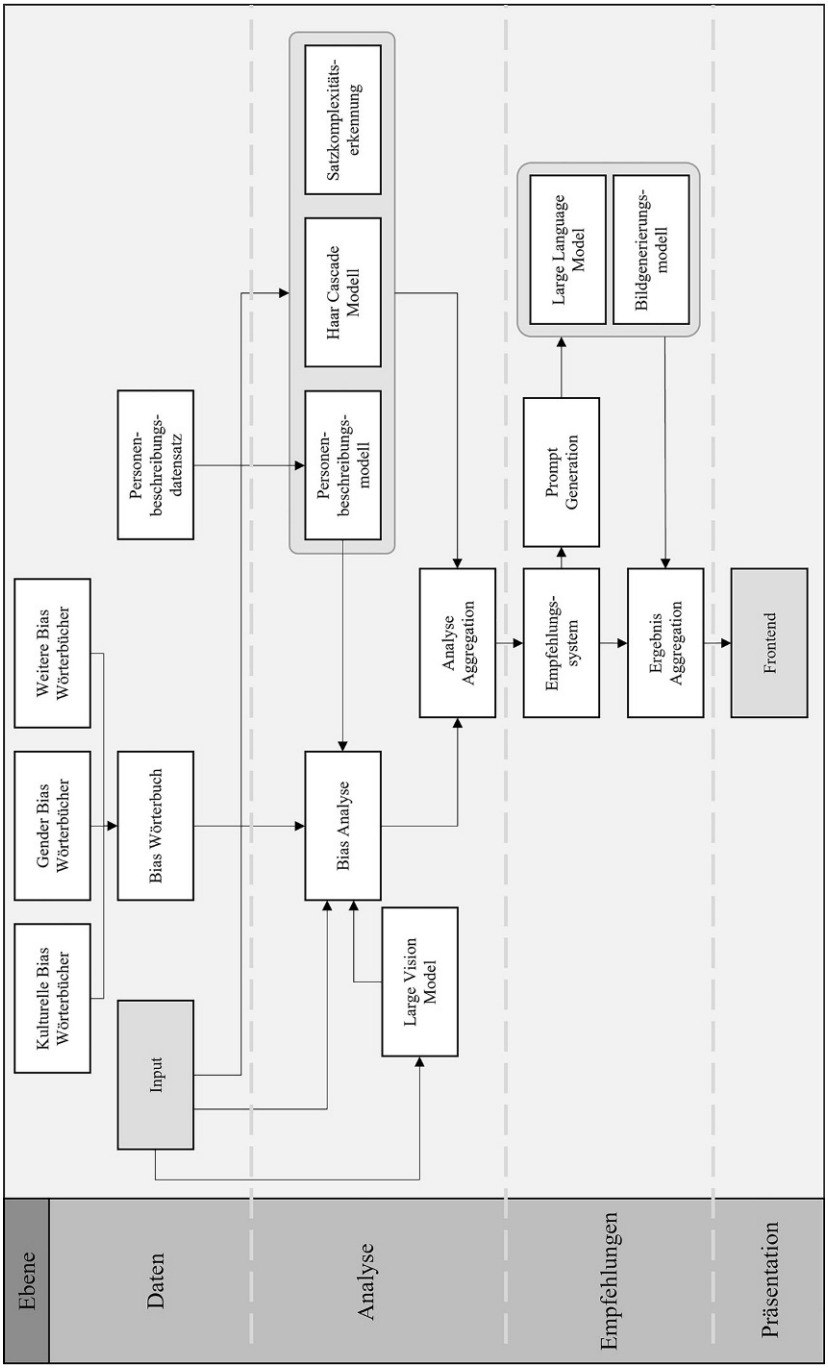


**Abb. 1** Architektur der Diversity-Governance-Software

und unbeabsichtigte Weise beeinflussen können (Noon 2018). Die Aufdeckung dieser impliziten Verzerrungen und Diskriminierungen erfordert ein tiefes Verständnis von sowohl linguistischen als auch gesellschaftlichen Aspekten. Technologien und KI können in diesem Zusammenhang Hilfestellung leisten. Insbesondere die KI-generierte Empfehlung diversitätssensibler Alternativvorschläge zur Überarbeitung von Unternehmenstexten wird erst durch die in den Jahren 2022 und 2023 weit fortgeschrittene Entwicklung von *Large Language Models* (LLM) wie OpenAIs GPT-Modellen, Googles Bard-Modell oder OpenSource-Modellen wie *Meta AIs LLaMA* ermöglicht. In diesem Zusammenhang sind auch Modelle zur Bildgenerierung wie OpenAIs *Dalle-E* und *Stable Diffusion* von Stability AI zu nennen. Die Kombination dieser Modelle mit herkömmlichen Methoden des NLP und der *Computer Vision* eröffnet neue Dimensionen für die automatisierte Erkennung von impliziter sprachlicher Diskriminierung in Unternehmen. Im Folgenden werden die Konzeption und die technische Implementierung der DG-Software beschrieben, bei der KI eingesetzt wird, um implizite sprachliche Diskriminierung automatisch zu identifizieren und zu reduzieren. In Abb. 1 werden die implementierten KI-Komponenten in einer Softwarearchitektur veranschaulicht.

Bei den inhaltlichen Ansätzen zur Identifikation impliziter sprachlicher Diskriminierung können unterschiedliche Aspekte berücksichtigt werden. Zum einen werden die eingegebenen Daten mittels bereits existierender Bias-Wörterbücher auf einen **stereotypen Sprachgebrauch** hin überprüft. Unter stereotypem Sprachgebrauch wird dabei die Verwendung von Sprache verstanden, die z. B. ausschließlich einer heteronormativen Lebenswelt, einer männlichen Sozialisation oder einer höheren

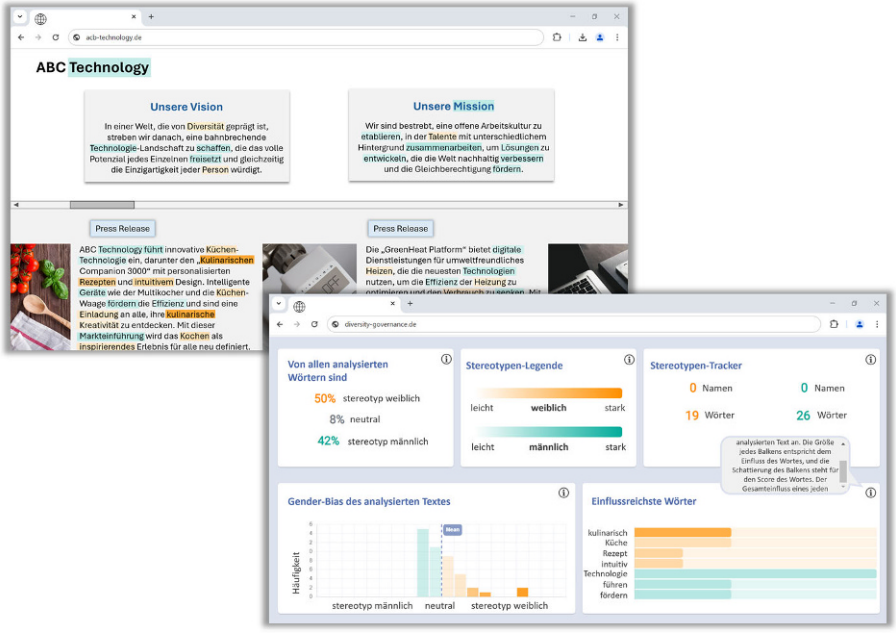


**Abb. 2** Identifikation geschlechterstereotyper Wörter am Beispiel einer fiktiven Unternehmenswebsite

Bildungsschicht entstammen. In Abb. 2 wird die Umsetzung dieses Analyseaspekts am Beispiel der fiktiven Unternehmenswebseite von deutlich.

Dabei werden die Wörter mit einem Gender-Bias auf einem Zwilling der Website farblich hervorgehoben: türkis für stereotyp männlich und orange für stereotyp weiblich. Zudem werden die Ergebnisse der Website-Analyse auf einem Dashboard zusammengefasst. Um die Verwendung von stereotyper Sprache automatisiert zu überprüfen, sind bestehende Bias-Wörterbücher implementiert. Verwendet werden umfangreiche Wörterbücher zu Geschlechterstereotypen (Göritz et al. 2022) und kulturellen Stereotypen (z. B. Jha et al. 2023). Durch die Modifikation einiger Trainingsdatensätze, insbesondere durch die im Folgenden beschriebene Erweiterung der Bias-Wörterbücher, ist eine Übertragung auf andere Bereiche möglich, z. B. die Überprüfung hinsichtlich alters- oder religionsbezogener Stereotype. Die Wörterbücher dienen als Grundlage, um Wörter und Phrasen mit stereotyper Konnotation im Text zu erkennen. Die DG-Software analysiert dazu die einzelnen Wörter und weist ihnen, ähnlich wie bei einer Sentiment-Analyse, einen spezifischen Wert zu, um ihre stereotype Konnotation zu bestimmen.

Des Weiteren wird die Bedeutung der diversen **Darstellung von Personen** in Texten und Bildern hervorgehoben, um stereotype Rollenerwartungen zu durchbrechen. Gemäß der *Social Role Theory* von Eagly und Wood (2012) können visuelle Stereotypen einen Einfluss darauf haben, ob sich Personen nach den erwarteten sozialen Rollen richten. Sie können sich negativ auf die Gesellschaft auswirken, indem sie Menschen in Rollen zwängen und sie daran hindern, ihr wahres Potenzial auszuschöpfen. Anstatt ihre eigenen Talente zu entwickeln, passen sie sich möglicherweise den Erwartungen der Gesellschaft an.

Ein Ansatz, stereotype Rollenerwartungen aufzubrechen, ist die ausgewogene Darstellung von Personen hinsichtlich aller Diversitätsmerkmale in Texten und Bil-

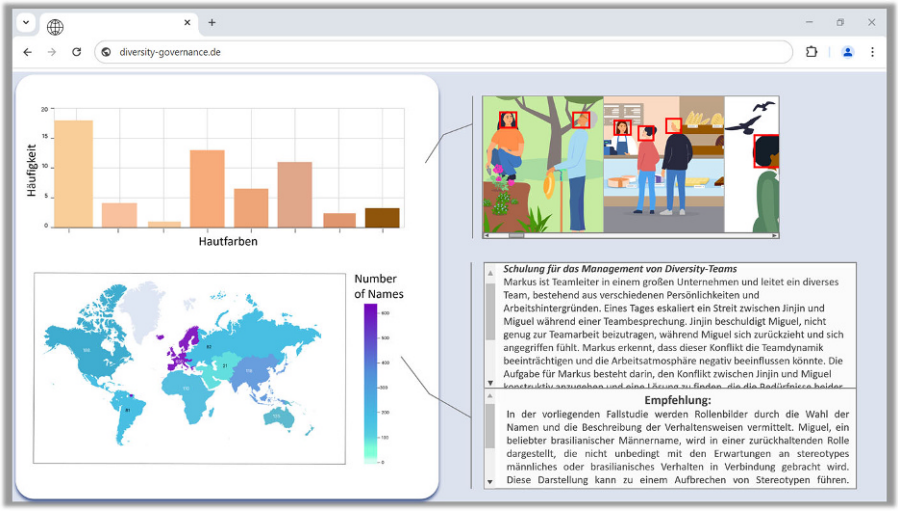


**Abb. 3** Analyse kulturell stereotyper Personen- und Rollendarstellungen am Beispiel fiktiver Schulungsunterlagen

dern, um jeder Person die Möglichkeit zu geben, sich mit Personen zu identifizieren, und so Vorbilder zu schaffen. In Abb. 3 werden die Ergebnisse der Überprüfung auf kulturell stereotype Personen- und Rollendarstellungen durch die DG-Software am Beispiel von fiktiven Schulungsunterlagen gezeigt. Dabei wird die Darstellung von Personen unterschiedlicher Hautfarbe in Bildern zusammengefasst und die Darstellung kultureller Rollen in Texten anhand der Namensetymologie analysiert.

Um geschlechterstereotype Rollen in Texten zeiteffizient zu erkennen, werden eingegebene Daten auf Personen und stereotype Wörter untersucht, die zur Beschreibung von Personen verwendet werden. Zu diesem Zweck wurde ein eigenes Modell trainiert. Der verwendete Datensatz bestand aus Sätzen, die Personen in bestimmten Kontexten beschreiben. Jeder Satz wurde mit Labels versehen, die mit der Person verbundene Wörter wie Namen, Personalpronomen, Berufe, Tätigkeiten und Adjektive enthielten. In Kombination mit Bias-Wörterbüchern und statistischen Daten, z. B. zur Geschlechterverteilung in Berufen, kann die Darstellung der im Text identifizierten Personen analysiert werden.

Um die Analysekapazitäten der DG-Software zu erweitern, wurde neben der Textverarbeitung auch die Verarbeitung von Bildmaterial integriert. Diese Erweiterung ermöglicht die Untersuchung visueller Inhalte auf stereotype Darstellungen über die reine Textanalyse hinaus. Dazu werden zum einen Methoden der *Computer Vision* eingesetzt. Ein zentrales Element ist dabei die Verwendung des Haar-Cascade-Modells, das Gesichter in Bildern identifiziert. Durch diese Technik können große Datensätze von Unternehmen nach abgebildeten Personen durchsucht werden, um anschließend die systematische Darstellung von Personen nach verschiedenen Merkmalen wie Hautfarbe und Kleidung zu untersuchen. Zum anderen werden *Large Vision Models* verwendet, um den gesamten Inhalt eines Bildes zu analysieren und in Textform zu übersetzen. Diese textuellen Bildbeschreibungen werden dann mit den implementierten Wörterbüchern und dem speziell entwickelten Modell zur Erkennung stereotyper Personenbeschreibungen kontrolliert. Mit diesem Ansatz können nicht nur einzelne stereotype Wörter oder Phrasen identifiziert, sondern auch komplexe visuelle Darstellungen und ihre Kontexte erfasst werden.

Ein weiterer relevanter Aspekt für die diversitätsbewusste Überarbeitung von Unternehmenssprache ist die ***Differenzierung nach Sprachniveaus***, um so einen niederschwelligen Zugang zu den vermittelten Inhalten für alle zu gewährleisten. In Branchen, in denen bestimmte Gruppen unterrepräsentiert sind, können stereotype Erwartungen zu Unsicherheiten führen und dadurch Fluktuationen erhöhen. Daher ist es von besonderer Bedeutung, eine Überforderung grundsätzlich zu vermeiden und einen Zugang zu relevanten Themen zu ermöglichen.

Eine automatisierte Analyse von Texten bezüglich ihrer Satzkomplexität ist bspw. mittels des Flesch-Index möglich und erlaubt die Bestimmung des erforderlichen Sprachniveaus der Texte (Flesch 1948). Der Flesch-Index ist ein bewährtes Instrument zur Beurteilung der Lesbarkeit und der Verständlichkeit von Texten. Dabei wird basierend auf Satzlänge und Silbenanzahl pro Wort gemessen, wie komplex ein Text ist. Eine weitere Möglichkeit, die Darstellung von Unternehmensinhalten zu vereinfachen, ist die Ergänzung von Texten durch Bilder. Ein ausgewogenes Text-Bild-Verhältnis kann dazu beitragen, die Gesamtkomplexität eines Dokuments zu

verringern, da Bilder den Textinhalt veranschaulichen und somit das Verständnis erleichtern.

Die Implementierung der automatisierten KI-gestützten Überarbeitungsfunktion stellt den letzten Schritt in der Entwicklung der DG-Software dar. Dazu werden die Ergebnisse der vorangegangenen Analysen aggregiert, um konkrete Empfehlungen für die Überarbeitung von Texten und Bildern zu generieren. Für die Textüberarbeitung verwendet die DG-Software ein LLM, das diversitätssensible Alternativen vorschlägt. Bei diesem Verfahren werden die Häufigkeit stereotyper Wörter und die systematische Darstellung von Personen analysiert. Auf Basis dieser Ergebnisse wird dann ein Prompt erstellt, mit dem die Überarbeitungsempfehlungen in natürliche Sprache übersetzt werden können. Dabei werden nicht nur Vorschläge zur Vermeidung stereotyper Sprache angeboten, sondern auch Begründungen für die empfohlenen Anpassungen bereitgestellt. Ein LLM dient auch dazu, die Satzkomplexität in Texten direkt unterschiedlichen Sprachniveaus anzupassen. Dabei können Texte in eine einfachere Sprache umformuliert werden, ohne den ursprünglichen Inhalt oder die intendierte Aussage zu verändern. Hierbei ist zu betonen, dass die DG-Software lediglich Überarbeitungsempfehlungen gibt; die endgültige Entscheidung darüber, was davon übernommen wird, jedoch dem Menschen überlassen ist.

Ähnlich wird bei der Überarbeitung von Bildmaterial vorgegangen. Hier werden die durch *Computer Vision* identifizierten Inhalte und die daraus resultierenden Analysen genutzt, um konkrete Vorschläge für die Bildgestaltung zu entwickeln. Die DG-Software generiert dazu zunächst einen Prompt, mit dem das LLM die Nutzer:innen auf mögliche Probleme im aktuellen Bildmaterial hinweist. Anschließend wird ein zweiter Prompt erstellt, der als Grundlage für ein Bildgenerierungsmodell dient. Dieses Modell erzeugt einen alternativen Bildvorschlag, der auf der Analyse und den Vorgaben des Prompts basiert. Diese Methode ermöglicht es, visuelle Stereotypen zu identifizieren und Alternativen für vielfältigere und inklusivere Darstellungen anzubieten. Bildgenerierungsmodelle werden auch verwendet, um die Unternehmenssprache verschiedenen Sprachniveaus entsprechend zu überarbeiten, indem visuelle Elemente zur Ergänzung des Textes erstellt werden. Diese lizenzfreien Bilder sind speziell darauf ausgerichtet, das Verständnis komplexer Textpassagen zu erleichtern.

## 5 Evaluation des Ansatzes sowie der prototypischen Implementierung

### 5.1 Evaluationsmethodik

Um Feedback zur Konzeption und zur technischen Implementierung der DG-Software zu erhalten, wurden zwölf Expert:innen (E) aus dem Bereich *Diversity* interviewt. Ausgewählt wurden Personen aus dem akademischen Bereich und der Industrie, die sich beruflich bspw. in der Beratung, als Gleichstellungsbeauftragte oder in der Unternehmenskommunikation mit dem Thema beschäftigen, um sowohl Forschungs- als auch Praxisperspektiven in die Ergebnisse einfließen zu lassen. Die Interviews dauerten zwischen 40 und 65 min und wurden aufgezeichnet. Sie folgten den Richtlinien von Myers und Newman (2007), so dass die halbstrukturierten In-

terviews anhand eines Interviewleitfadens durchgeführt wurden, der offene Fragen enthielt und eine flexible Anpassung an die Expert:innen ermöglichte. Die ersten Fragen zielten auf die Vorstellung der Befragten und die Beschreibung ihrer beruflichen und persönlichen Erfahrungen mit dem Forschungsthema ab. Danach lag der Schwerpunkt auf der Evaluation der DG-Software. In diesem Zusammenhang wurden die Fragen auf Grundlage der im Verlauf des Interviews gewonnenen Erkenntnisse kontinuierlich angepasst (Glaser et al. 1968). Infolgedessen wurden die Interviews im Laufe der Studie immer gezielter und spezifischer.

Nach der Durchführung der Interviews wurden diese transkribiert und zweimal induktiv codiert. Dabei wurden Ziele und Herausforderungen differenziert, die sich bei der Konzeption und der technischen Umsetzung einer Software zur automatisierten Erkennung impliziter sprachlicher Diskriminierung ergeben. Dem Ansatz der *Grounded Theory* folgend, wurden kontinuierlich Konzepte und Muster in den Interviewdaten verglichen (Strauss und Corbin 1990). In der ersten Kodierungsrunde sichteten zwei Forscher:innen die Interviewdaten und führten eine offene Kodierung durch. In der zweiten Kodierungsrunde wurde die Zuordnung der Codes mit zwei weiteren Forscher:innen diskutiert und angepasst.

### 5.2 Ergebnisse

Die Ergebnisse der Interviews wurden zu drei Zielen (Z) und vier Herausforderungen (H) zusammengefasst, die sich bei der Konzeption und der technischen Umsetzung ergeben. Zu den Zielen gehört die automatisierte Erkennung impliziter sprachlicher Diskriminierungen in der Unternehmenskommunikation (Z1). Der Fokus liegt dabei auf der nachvollziehbaren Erklärung der Wirkung von Unternehmenssprache. Dies ist vor allem deshalb von Bedeutung, weil implizite sprachliche Diskriminierung in der Regel nicht direkt auf der Satz- oder Bildoberfläche erkennbar ist, sondern bewusst sichtbar gemacht werden muss. Die DG-Software zielt jedoch nicht nur auf das bloße Erkennen impliziter sprachlicher Diskriminierung ab, sondern bietet auch konkrete Handlungshilfen in Form von intelligenten Empfehlungen für z. B. alternative Formulierungen (Z2). Durch das Anbieten konkreter Alternativen werden Nutzer:innen an die Hand genommen, so dass ihr Gefühl der Hilflosigkeit reduziert wird. Dieses Gefühl kann entstehen, wenn Menschen unsicher sind, wie sie im Sinne der *Diversity Governance* kommunizieren sollen. Konkrete Alternativen geben Orientierung und zeigen direkte Handlungswege auf. Zudem werden Nutzer:innen bei der Bedienung der DG-Software die kontinuierlich hinsichtlich impliziter sprachlicher Diskriminierung geschult. Durch das Sichtbarmachen impliziter Diskriminierungen in Texten und Bildern werden sie für unbewusste Verzerrungen sensibilisiert und dazu befähigt, in Zukunft im Sinne der *Diversity Governance* zu kommunizieren (Z3).

Zu den Herausforderungen, die bei der Konzeption und der technischen Implementierung einer Software zur automatisierten Erkennung impliziter sprachlicher Diskriminierung entstehen können, zählt unter anderem das Risiko, dass Nutzer:innen Reaktanz entwickeln (H1): „Ich erlebe es aber auch so, dass sehr viel Widerstand gegen den Gebrauch der Gendersprache oder der inklusiven Sprache vorhanden ist. Auch bei vielen meiner Kunden: ‚Kommen Sie mir bloß nicht da-

mit!‘ und ‚Das ist zu kompliziert!‘“ (E4). Wie bereits in der *Harvard Business Review* (2016) beleuchtet, besteht die Gefahr, dass Diversity-Programme bei Mitarbeiter:innen ein Gefühl der Überwachung vermitteln und Widerstand hervorrufen. Eine zentrale Anforderung an die DG-Software ist daher, dass insbesondere das User-Experience-Design unaufdringlich und nicht belehrend wirkt.

Des Weiteren wiesen die Expert:innen darauf hin, dass mit solcher Software im Zweifelsfall Stereotype weiter reproduziert werden könnten (H2): „Die Grundfrage bei mir ist immer bei solchen Tools: Nützen sie, um Bias aufzulösen oder Bias zu verstärken?“ (E9) So könnte bspw. die Markierung von stereotyp männlichen Wörtern in den analysierten Texten bisher inaktive Stereotype aktivieren und bei den Nutzer:innen den Fehlschluss hervorrufen, dass die Aussagen zutreffend sind. Dieser Herausforderung kann durch klischeefreie und aufklärende Darstellungen im Design begegnet werden.

Die Ergebnisse der Interviews weisen zudem auf die Herausforderung (H3) hin, dass unterschiedliche Anwendungsfälle der DG-Software individuelle Funktionen und damit ein hohes Maß an Flexibilität und Anpassbarkeit erfordern: „Stattdessen würde ich das über Fragen machen, die kritisch hinterfragen. Also so wirklich nach dem Motto: Was ist mein Ziel?“ (E11) Je nach Anwendungsfall werden unterschiedliche Analysen und Empfehlungen benötigt. Dieser Anforderung kann durch KI-gestützte Lösungen begegnet werden, die einen hohen Grad an Personalisierung ermöglichen.

Die vierte Herausforderung (H4) besteht in einer zeitlosen Funktionalität, so dass die DG-Software immer aktuell bleibt: „Sprache bewegt sich, entwickelt sich. […] [V]or drei Jahren hätte keiner ‚Studierende‘, ‚Prüfende‘, ‚Mitarbeitende‘ gesagt, ohne dass alle gucken.“ (E1). Auch diese Anforderung kann durch KI-basierte Lösungen realisiert werden, z. B. in Form eines selbstlernenden Systems. Insbesondere im Hinblick auf den sich ständig verändernden Umgang mit inklusiver Sprache ist es essenziell, dass die DG-Software kontinuierlich angepasst und aktualisiert wird.

## 6 Diskussion

In der bisherigen Forschung wurde gezeigt, dass herkömmliche Programme wie obligatorische Diversity-Schulungen ein Gefühl der Überwachung vermitteln, Widerstand bei den Mitarbeitenden hervorrufen und Vorurteile aktivieren können (Noon 2018). Mit der vorliegenden Studie wird ein innovativer Ansatz zur Förderung einer diversitätsbewussten Ausrichtung von Unternehmen präsentiert, um implizite sprachliche Diskriminierung in der Unternehmenskommunikation automatisiert zu reduzieren. Er wurden im Rahmen wissenschaftlicher Forschung entwickelt, prototypisch umgesetzt und evaluiert. Die daraus gewonnenen Erkenntnisse können in die Entwicklung eines marktreifen Produktes einfließen.

Es werden bereits automatisierte Ansätze kommerziell angeboten. Dabei werden jedoch primär spezifische Diversitätsmerkmale wie das Geschlecht fokussiert oder der Anwendungsbereich ist insbesondere auf Stellenausschreibungen beschränkt. Somit sind die Analysen auf wenige Wörter begrenzt und nur schwer übertragbar. Eine ganzheitliche Anwendbarkeit automatisierter Analysen auf die gesamte interne

und externe Unternehmenskommunikation ist jedoch essenziell für die erfolgreiche Implementierung einer Diversity-Governance-Strategie. Durch die Nutzung öffentlich verfügbarer Text-Corpora und bereits trainierter Sprachmodelle können die Analysen der DG-Software schnell auf vielfältige Anwendungsbereiche adaptiert werden. So können mithilfe des Ansatzes die interne und die externe Unternehmenskommunikation in verschiedenen Bereichen analysiert werden, bspw. Schulungsunterlagen für das Personalwesen, Webseiten und Social-Media-Beiträge für das Marketing, Kund:innenkommunikation für den Vertrieb sowie textbasierte Produkte für Forschung und Entwicklung.

Neben der Anpassungsfähigkeit an verschiedene Anwendungsdomänen bietet die DG-Software im Vergleich zu manuellen Analysen den Vorteil, dass weniger zeitliche und personelle Ressourcen beansprucht werden. Eine manuelle Überprüfung von Texten und Bildern ist zudem nicht nur zeit- und kostenaufwändig, sondern auch anfällig für menschliche Ermüdungseffekte und damit für Fehler und Inkonsistenzen. Eine automatisierte Analyse ist effizienter und konsistenter. Die entwickelte DG-Software soll Mitarbeitenden und Führungskräften ein effizientes Werkzeug bereitstellen, um sprachliche Diskriminierung mit minimalem Aufwand zu erkennen und zu eliminieren. Aufgrund des virtuellen und damit einfachen Zugangs kann bei digitalen Ansätzen wie der DG-Software die Zielgruppe auf alle Mitarbeiter:innen eines Unternehmens ausgeweitet werden, so dass sie auch zur Demokratisierung von Diversity-Programmen beitragen.

Bei der DG-Software als neuem Ansatz sollen allerdings bisherige Risiken traditioneller Diversity-Schulungen berücksichtigt werden. Anstatt Mitarbeitende mit allgemeinen Ratschlägen und theoretischen Konzepten zu überladen, bietet die Software konkrete und individuelle Beispiele für Alternativformulierungen, um so die Intentions-Verhaltens-Lücke zu reduzieren (Gill und Olson 2023). Diese praktische Herangehensweise ermöglicht es den Mitarbeitenden, direkt umsetzbare Lösungen zu verstehen und zu bedienen, was dem Gefühl der Hilflosigkeit entgegenwirkt. In Anlehnung an Gill und Olson (2023) soll die DG-Software eine eher autonomiefördernde als eine direktive Kommunikation über *Diversity* mit den Mitarbeitenden ermöglichen und so Reaktanz und Widerstand reduzieren. Mitarbeitende müssen nicht mehr zur Teilnahme an Schulungen verpflichtet werden, da mit der DG-Software ein selbstgesteuerter Ansatz vorliegt. Durch die automatische Überprüfung der eigenen sprachlichen Diskriminierungen können die Mitarbeitenden fundierte Entscheidungen im Sinne der *Diversity Governance* treffen, ohne sich überwacht zu fühlen.

Bei der Evaluation wurden vier bedeutende Herausforderungen bei der Konzeption und der technischen Implementierung der DG-Software herausgestellt. Bei Bewältigung dieser Herausforderungen kann die DG-Software im Gegensatz zu vielen herkömmlichen Diversitätsprogrammen erfolgreich zur Umsetzung einer Diversity-Governance-Strategie beitragen.

Zur Reduzierung des Risikos von Reaktanz gegenüber dem Thema *Diversity* während der Softwarenutzung ist insbesondere im Bereich des User-Experience-Designs eine umfassende und durchdachte Herangehensweise erforderlich (H1). Es ist zunächst von entscheidender Bedeutung, den Nutzer:innen die Ziele und den Nutzen der DG-Software klar zu vermitteln. Sie sollten verstehen, warum die DG-

Software relevant ist und welchen Vorteil sie sowohl für sie persönlich als auch für ihr Unternehmen bietet. Zusätzlich sollte die DG-Software für verschiedene Nutzer:innengruppen zugänglich und benutzer:innenfreundlich gestaltet sein, wobei diverse kulturelle Hintergründe, Sprachen und Fähigkeiten berücksichtigt werden müssen. Ferner sollte eine aufdringliche oder bevormundende Darstellung des Themas vermieden werden. Stattdessen sind die Informationen zu vermitteln, ohne belehrend zu wirken oder Nutzer:innen unter Druck zu setzen. Es können zudem positive Verstärkungsmechanismen implementiert werden, um Nutzer:innen zur Auseinandersetzung mit dem Thema Diversity zu ermutigen. Zusätzliche Sicherheits- und Datenschutzmaßnahmen können das Vertrauen der Nutzer:innen in die DG-Software weiter stärken. Dabei ist es essenziell, das Fachwissen verschiedener Disziplinen wie Psychologie, Soziologie, Wirtschaftswissenschaften und Informatik in der Konzeption der DG-Software zu einzuholen, um eine umfassende Perspektive auf die Bedürfnisse der Nutzer:innen zu erhalten und entsprechend zu agieren. Weiterhin sollten die Werte im Rahmen der DG-Software klar kommuniziert werden, da Authentizität und Ethik entscheidend sind, um das Vertrauen der Nutzer:innen zu gewinnen.

Um das Risiko der Reproduktion von Stereotypen durch eine DG-Software (H2) zu verringern, können verschiedene Maßnahmen ergriffen werden. Schon im frühen Stadium der Konzeption und der Umsetzung der DG-Software sollte das Team, das an der Entwicklung beteiligt ist, z. B. mittels Schulungen und Workshops für Stereotypen und Vorurteile sensibilisiert werden, um diese erkennen und vermeiden zu können. Es ist außerdem sicherzustellen, dass das Entwicklungsteam selbst vielfältig zusammengesetzt ist, um verschiedene Perspektiven und Erfahrungen einzubringen und somit blinde Flecken in Bezug auf Stereotypen zu vermeiden. Die Darstellung von Personen in Grafiken, Bildern oder Texten sollte ebenfalls frei von stereotypen Zuschreibungen sein, um deren Reproduktion bei der Nutzung der DG-Software zu verhindern. Dies kann erreicht werden, indem in der DG-Software Fallstudien oder Erfolgsgeschichten genutzt werden, die die Vielfalt und die unterschiedlichen Hintergründe von Menschen zeigen. Des Weiteren können Feedback-Mechanismen implementiert werden, damit Nutzer:innen problematische oder stereotype Inhalte direkt melden können.

Um verschiedene Nutzungsszenarien durch die DG-Software zu ermöglichen, die individuelle Funktionen und Flexibilität erfordern (H3), sind verschiedene Ansätze denkbar: So sollte die DG-Software modular entwickelt werden, um flexibel erweiterbar zu sein und neue Funktionen hinzufügen sowie Modifikationen vornehmen zu können. Darüber hinaus können manuelle Einstellungsmöglichkeiten und Filter integriert werden, die es den Nutzer:innen ermöglichen, die Funktionalität und die Inhalte der DG-Software ihren individuellen Bedürfnissen anzupassen. Zudem können teilautomatisierte Personalisierungstechnologien und adaptive KI-Algorithmen eingesetzt werden, um die DG-Software selbstlernend zu gestalten. Basierend auf dem Nutzer:innenverhalten und den Präferenzen kann die DG-Software kontextsensitive Inhalte liefern oder personalisierte Funktionen vorschlagen, um den individuellen Anforderungen gerecht zu werden. Ebenso gilt es, Mechanismen zur kontinuierlichen Rückmeldung der Nutzer:innen umzusetzen, um die DG-Software gemäß den Anforderungen zu modellieren und Anpassungen vorzuschla-

gen, die die individuellen Anwendungsfälle besser unterstützen. Klare Erklärungen und Anleitungen sollten bereitgestellt werden, um zu veranschaulichen, inwiefern die Funktionen und die Anpassungen der DG-Software den persönlichen Ansprüchen genügen.

Um sicherzustellen, dass die Software-Funktionalität insbesondere in Bezug auf den Umgang mit inklusiver Sprache und sich verändernden Normen der Vielfalt und Inklusion aktuell bleibt (H4), können zum einen selbstlernende Algorithmen implementiert werden, um aus dem Nutzer:innenverhalten kontinuierlich zu lernen und die Software anzupassen. So kann KI dazu beitragen, die DG-Software ständig zu verbessern und zu aktualisieren, um sich wandelnde Sprachmuster und inklusive Praktiken zu berücksichtigen. Zum anderen sollten Mechanismen integriert werden, die es ermöglichen, die DG-Software in Echtzeit zu aktualisieren oder regelmäßige Überprüfungen durchzuführen, um sicherzustellen, dass sie aktuellen inklusiven Sprachrichtlinien entspricht und an *Best Practices* ausgerichtet ist. Auch Feedbackmechanismen, die es den Nutzer:innen ermöglichen, Bedenken oder Vorschläge zur Verbesserung der DG-Software zu äußern, sind von großer Bedeutung. Diese Einbeziehung in den Entwicklungsprozess kann dazu beitragen, die DG-Software kontinuierlich zu optimieren. Des Weiteren ist es entscheidend, die Branche zu beobachten und Benchmarking-Studien durchzuführen, um die Konzeption ähnlicher Softwarelösungen zu verstehen und somit bewährte Methoden zu identifizieren und zu adaptieren.

## 7 Fazit, Limitationen und Ausblick

Im vorliegenden Artikel wurde der Ansatz der *Diversity Governance* erläutert. Im Vergleich zur *Corporate Governance* zielt dieser Ansatz darauf ab, die Handlungsspielräume von Mitarbeitenden und Führungskräften so zu gestalten, dass sie befähigt werden, gemäß den Prinzipien der *Diversity Governance* zu kommunizieren. Zur Förderung von Diversity-Maßnahmen kann der Einsatz technologiebasierter Strategien wertvolle Unterstützung bieten. Ein Beispiel hierfür ist die DG-Software, die darauf abzielt, implizite Diskriminierungen in der Unternehmenskommunikation automatisiert zu erkennen und über das reine Reporting von Kennzahlen hinaus konkrete Handlungsempfehlungen in Form von intelligenten Vorschlägen bereitzustellen.

Diese Studie ist wie jede andere Forschungsbemühung nicht frei von Einschränkungen. Als Methode wurden die systematische Internetrecherche und die Marktanalyse gewählt, um marktreife Ansätze zur automatisierten Erkennung impliziter sprachlicher Diskriminierung in Unternehmen zu ermitteln. Die Berücksichtigung des Marktreifegrads war aufgrund der hohen Komplexität der sozialen Dynamiken des Untersuchungsgegenstands zwar von Bedeutung, jedoch blieben die Ergebnisse aufgrund dieser Herangehensweise auf eine praktische Perspektive beschränkt. Für eine zukünftige theoretische Einbettung ist es erforderlich, dass die Ergebnisse enger mit den aktuellen wissenschaftlichen Methoden zur automatisierten Erkennung impliziter sprachlicher Diskriminierung in Unternehmen verknüpft werden. Zukünftige

Forschung sollte zudem darauf abzielen, weitere Möglichkeiten herauszuarbeiten, wie KI im Sinne der *Diversity Governance* genutzt werden kann.

Eine weitere Einschränkung der vorliegenden Studie besteht darin, dass bei der Evaluation der DG-Software der direkte quantitative Einfluss auf einstellungs- und verhaltensbezogene Faktoren (z. B. Reaktanz) potenzieller Nutzer:innen nicht gemessen wurde. Daher ist weitere Forschung notwendig, um die Auswirkungen von KI-gestützten Diversity-Governance-Ansätzen empirisch zu messen und die Identifizierung sowie die Reduzierung von Diskriminierung durch Automatisierung zu unterstützen.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf http://creativecommons.org/licenses/by/4.0/deed.de.

## Literatur

Andrevski G, Richard OC, Shaw JD, Ferrier WJ (2014) Racial diversity and firm performance: the mediating role of competitive intensity. J Manage 40(3):820–844. https://doi.org/10.1177/0149206311424318

Antidiskriminierungsstelle des Bundes (2024) Aufgaben der Antidiskriminierungsstelle des Bundes. https://www.antidiskriminierungsstelle.de/DE/ueber-uns/unsere-aufgaben/unsere_aufgaben_inhalt.html. Zugegriffen: 6. Mai 2024

Ashcraft C, McLain B, Eger E (2016) Women in tech: the facts. National Center for Women & Technology (NCWIT), Colorado

Barmeyer C, Bausch M, Moncayo D (2019) Cross-cultural management research: topics, paradigms, and methods—a journal-based longitudinal analysis between 2001 and 2018. Int J Cross Cult Manag 19(2):218–244. https://doi.org/10.1177/1470595819859603

Becker M (2016) Was ist Diversity Management? In: Fereidooni K, Zeoli AP (Hrsg) Managing Diversity: Die diversitätsbewusste Ausrichtung des Bildungs- und Kulturwesens, der Wirtschaft und Verwaltung, S 291–317

Bertrand M, Chugh D, Mullainathan S (2005) Implicit discrimination. Am Econ Rev 95(2):94–98

Bezrukova K, Jehn KA, Spell CS (2012) Reviewing diversity training: where we have been and where we should go. AMLE 11(2):207–227. https://doi.org/10.5465/amle.2008.0090

Correll SJ (2017) SWS 2016 feminist lecture: reducing gender biases in modern workplaces: a small wins approach to organizational change. Gend Soc 31(6):725–750

Díaz-Fernández MC, González-Rodríguez MR, Simonetti B (2020) Top management team diversity and high performance: an integrative approach based on upper echelons and complexity theory. Eur Manag J 38(1):157–168. https://doi.org/10.1016/j.emj.2019.06.006

Dobbin F, Kalev A (2016) Why diversity programs fail. Harvard Business Review. https://hbr.org/2016/07/why-diversity-programs-fail. Zugegriffen: 14. Dez. 2023

Eagly AH, Wood W (2012) Social role theory. Handbook of theories of social psychology. SAGE, Los Angeles
Flesch R (1948) A new readability yardstick. J Appl Psychol 32(3):221
Gao J, Wu J, You J, Smith M (2021) Migrant entrepreneurs and firm innovation. Account Finance 61(5):6069–6112. https://doi.org/10.1111/acfi.12834
Gebert D (2004) Durch diversity zu mehr Teaminnovativität? Betriebswirtschaft 64(4):412–430
Gill LJ, Olson MA (2023) A dual-process framework for diversity training to reduce discrimination in organizational settings. Social Issues Policy Review 17(1):79–103
Glaser BG, Strauss AL, Strutzel E (1968) The discovery of grounded theory; strategies for qualitative research. Nurs Res 17(4):364
Göritz L, Stattkus D, Beinke JH, Thomas O (2022) To reduce bias, you must identify it first! Towards automated gender bias detection. ICIS 2022 Proceedings. 10. https://aisel.aisnet.org/icis2022/data_analytics/data_analytics/10
Jha A, Davani A, Reddy CK, Dave S, Prabhakaran V, Dev S (2023) A stereotype benchmark with broad geo-cultural coverage leveraging generative models. Comput Lang. https://doi.org/10.48550/arXiv.2305.11840
Kewes T (2022) Diversitäts-Ranking: Allianz, SAP und Deutsche Telekom sind die Gewinner. Handelsblatt. https://www.handelsblatt.com/unternehmen/management/management-diversitaets-ranking-allianz-sap-und-deutsche-telekom-sind-die-gewinner/28374278.html. Zugegriffen: 14. Dez. 2023
Myers MD, Newman M (2007) The qualitative interview in IS research: examining the craft. Inf Organ 17(07):2–26. https://doi.org/10.1016/j.infoandorg.2006.11.001
Nazir B, Ali M, Jamil M (2016) The impact of brand image on the customer retention: a mediating role of customer satisfaction in Pakistan. Int J Bus Manag Invent 5(3):56–61
Noon M (2018) Pointless diversity training: unconscious bias, new racism and agency. Work Employ Soc 32(1):198–209. https://doi.org/10.1177/0950017017719841
Page S (2008) The difference: how the power of diversity creates better groups, firms, schools, and societies—new edition. Princeton University Press, Princeton https://doi.org/10.1515/9781400830282
Pandita D (2021) Innovation in talent management practices: creating an innovative employer branding strategy to attract generation Z. Int J Innov Sci 14(3/4):556–569
Pennebaker JW (2011) The secret life of pronouns. New Sci 211(2828):42–45
Statista Research Department (2024) Umfrage zu Diversity-Management und wirtschaftlichem Erfolg von Unternehmen 2020. https://de.statista.com/statistik/daten/studie/1226685/umfrage/relevanz-von-diversity-management-fuer-wirtschaftlichen-erfolg-von-unternehmen. Zugegriffen: 6. Mai 2024
Strauss A, Corbin J (1990) Basics of qualitative research. SAGE
Vereinte Nationen (2023) The sustainable development goals report 2023. https://unstats.un.org/sdgs/report/2023/The-Sustainable-Development-Goals-Report-2023.pdf
Vereinte Nationen (2024) Ziele für nachhaltige Entwicklung. UNRIC – Regionales Informationszentrum der Vereinten Nationen. https://unric.org/de/17ziele/. Zugegriffen: 6. Mai 2024
Vogel J, Schuir J, Thomas O, Teuteberg F (2020) Gestaltung und Erprobung einer Virtual-Reality-Anwendung zur Unterstützung des Prototypings in Design-Thinking-Prozessen. HMD 57(3):432–450
Wilton LS, Sanchez DT, Unzueta MM, Kaiser C, Caluori N (2019) In good company: when gender diversity boosts a company's reputation. Psychol Women Q 43(1):59–72
Zheng L (2023) How racial diversity and gender diversity in job positions affect the economy? CE 17(2):109–127
Østergaard CR, Timmermans B, Kristinsson K (2011) Does a different view create something new? The effect of employee diversity on innovation. Res Policy 40(3):500–509

**Hinweis des Verlags** Der Verlag bleibt in Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutsadressen neutral.