

Digital Future Challenge

#4

Verantwortungsvoller Umgang mit
künstlicher Intelligenz



Inhaltsverzeichnis

Editorial

Grußwort Dr. Volker Wissing	4
Vorwort Deloitte-Stiftung	5
Vorwort Initiative D21	6
DFC #4 auf einen Blick	7
Use Cases	8
Jury	9
Mentor*innen	10
Die Halbfinalteams und deren Ausarbeitungen im Überblick	11
Die Finalteams und deren Ausarbeitungen im Überblick	12

Beiträge der Studierenden

EduPin	13
AI Allies	19
GridHarmony	27
PowerBalance	33
BeyondTheBarrier	39

Dialogreihe Digital Responsibility

KI for Future?	46
Ein Jahr mit ChatGPT	47
Is this fAIr?	48

Anhang

ESG-Kriterien	50
Sustainable Development Goals	51
Danksagung an die Unterstützer*innen	52
Impressum und Informationen	53



Editorial

EduPin

AI Allies

GridHarmony

PowerBalance

Beyond
TheBarrier

Dialogreihe

Anhang

Editorial

Grußwort

Liebe Leserinnen und Leser,

digitaler Fortschritt lebt von innovativen Ideen – von Ideen, wie sie die Teilnehmerinnen und Teilnehmer der Digital Future Challenge entwickeln. Die Projekte und Anwendungen, die bei der jetzt schon vierten Auflage des Wettbewerbs ins Rennen gegangen sind, haben mich tief beeindruckt. Denn sie setzen nicht nur technologisch Akzente, sondern spiegeln auch eine enorme Begeisterung für Neues und für das Gestalten von Gegenwart und Zukunft wider. Genau das brauchen wir, um die vielfältigen Chancen der Digitalisierung noch besser nutzen zu können.

Zu den digitalen Gamechangern für Leben und Alltag, für Arbeiten und Produzieren zählt die künstliche Intelligenz. Der verantwortungsvolle Umgang mit ihr war der Schwerpunkt dieser Wettbewerbsrunde. Klar ist: KI ist gekommen, um zu bleiben. Deshalb ist es zum einen wichtig, dass wir uns mit ihr auseinandersetzen und Kompetenzen im Umgang mit KI vertiefen. Zum anderen wollen und müssen wir in Deutschland und Europa vorne mit dabei sein, wenn es um das Entwickeln vertrauenswürdiger KI geht. Wir dürfen das Spielfeld nicht anderen überlassen. Deshalb gestalten wir die nötigen Bedingungen dafür, dass die vielen klugen Köpfe, die sich mit KI beschäftigen, ihre Ideen bei uns verwirklichen können. Dazu gehört zum Beispiel, dass wir Freiräume schaffen, für praxistaugliche Regeln sorgen, Technologieoffenheit

fördern und bürokratische Hürden abbauen. Das sind anspruchsvolle, fordernde Aufgaben. Die Digital Future Challenge zeigt aber, dass es sich lohnt, dran zu bleiben. Denn wir verfügen in Deutschland über eine riesige Innovationskraft, die wir entfesseln wollen.

Vielen Dank an alle, die die Digital Future Challenge organisieren und möglich machen, vor allem an die Deloitte-Stiftung, die Initiative D21 und die Jury. Ein großes Dankeschön außerdem an die Studentinnen und Studenten, die mitgemacht haben und dabei waren. Sie haben sich Gedanken über die Digitalisierung und die Zukunft gemacht. Damit haben sie sich selbst und andere inspiriert. Das ist wertvoll für uns alle, die gesamte Gesellschaft profitiert davon.

Dr. Volker Wissing MdB

Bundesminister für Digitales und Verkehr
Schirmherr der DFC



Dr. Volker Wissing MdB
Bundesminister für Digitales
und Verkehr, Schirmherr

Vorwort Deloitte-Stiftung



Dorothea Schmidt
Vorstand Deloitte-Stiftung



Thomas Northoff
Vorstandsvorsitzender der
Deloitte-Stiftung

Sehr geehrte Leser*innen,

wir stehen an der Schwelle einer neuen Ära, in der künstliche Intelligenz (KI) nicht mehr nur ein futuristisches Konzept ist, sondern in nahezu alle Aspekte unserer Gesellschaft und unseres täglichen Lebens Einzug erhält. Mit großen Chancen kommt auch große Verantwortung. Diese Technologie sollte dabei so gestaltet und eingesetzt werden, dass sie dem Wohl unserer Gesellschaft dient. Angesichts der immer weiter fortschreitenden Entwicklung von KI ist es daher von großer Bedeutung, dass wir einen verantwortungsvollen Umgang mit dieser Technologie sicherstellen. KI bietet enorme Chancen, z. B. Verbesserungen von medizinischen Diagnosen, personalisierte Erfahrungen und effizientere Arbeitsabläufe, um nur einige Beispiele zu nennen, aber birgt auch Risiken, wie z. B. Bias, Diskriminierung und Fehleranfälligkeit. Diese Risiken und

ethischen Herausforderungen gilt es daher zu berücksichtigen. Wir sollten darauf achten, dass KI-Entscheidungen transparent, fair und nachvollziehbar sind. Menschen sollten verstehen können, wie diese Entscheidungen zustande kommen, welche Auswirkungen sie haben und wie ggf. ändernder Einfluss genommen werden kann.

Um einen verantwortungsvollen Umgang mit KI zu fördern, ist eine gemeinschaftliche Anstrengung erforderlich. Der Austausch von Wissen und Best Practices ist entscheidend, um bewährte Verfahren zu teilen und voneinander zu lernen. Genau da setzt die Digital Future Challenge an, wir bringen mit der DFC verschiedene Akteur*innen zusammen: die junge Generation, die Zivilgesellschaft, die Unternehmen und wirtschaftsnahen Organisationen, die Wissenschaft und die Politik. Dieses Projekt zeigt, dass es möglich ist, alle Interessengruppen an einem Tisch zu versammeln, um die digitale Zukunft gemeinsam zu gestalten. Wir lernen voneinander, lassen uns inspirieren und können gemeinsam etwas bewegen. Die Digital Future Challenge ist die perfekte Plattform, um die Diskussion über KI-Ethik voranzutreiben und innovative Lösungen zu entwickeln, die das Potenzial der KI voll ausschöpfen und gleichzeitig sicherstellen, dass sie zum Wohle aller eingesetzt wird.

Die vierte Edition der Digital Future Challenge unter der Schirmherrschaft von Bundesminister Dr. Volker Wissing und mit Förderung des Bundesministeriums

für Digitales und Verkehr setzt genau hier an. Sie öffnet den Raum, um disruptive Ideen zu diskutieren und Fragestellungen rund um die verantwortungsvolle Nutzung von KI zu adressieren.

Als Vorstand der Deloitte-Stiftung, mit der wir unter dem Leitmotiv „Gemeinsam Zukunft unternehmen“ unser Engagement für Bildung und Wissenschaft bündeln und damit einen Beitrag zur Verbesserung des Innovationstransfers von der Hochschule in die Wirtschaft leisten, ist es uns ein persönliches Anliegen, Sie mit dieser Publikation zum Mitdenken und Mitdiskutieren anzuregen: Denn nur gemeinsam können wir unsere digitale Zukunft verantwortungsvoll und zum Wohle aller gestalten. Wir möchten Sie daher einladen, anhand der fünf Ideen unserer Finalteams mit uns in den Dialog zu treten, wie ein verantwortungsvoller Umgang mit KI aussehen kann und in der Praxis umgesetzt werden sollte. Die Projekte zeigen auf tolle und kreative Weise, wie die Digitalisierung zum Vorteil für unsere Gesellschaft gestaltet werden kann.

Wir wünschen Ihnen viel Freude beim Lesen!

Dorothea Schmidt

Dorothea Schmidt,
Vorstand Deloitte-Stiftung

Thomas Northoff

Thomas Northoff,
Vorstandsvorsitzender der Deloitte-Stiftung

Vorwort Initiative D21



Marc Reinhardt
Präsident Initiative D21 e. V.

Liebe Leser*innen,

welche digitale Zukunft wünschen wir uns und wie können wir diese verantwortungsvoll gestalten? Mit diesen Leitfragen adressiert die Digital Future Challenge Studierende, Unternehmen und Politik, um das gemeinsame Engagement zu fördern und Nachwuchskräfte bereits heute zu ermutigen, sich mit den großen Fragen für die digitale Zukunft auseinanderzusetzen. In diesem Jahr hat der Wettbewerb das Schwerpunktthema „Verantwortungsvoller Umgang mit künstlicher Intelligenz (KI)“ und dies aus drei Perspektiven in die Ideenschmiede der Studierenden gegeben:

- Ökologisch: Welchen Beitrag kann KI zur ressourcen- und klimaschonenden Nachhaltigkeit leisten?

- Sozial: Wie können KI-Systeme so entwickelt werden, dass sie die sozialen Auswirkungen berücksichtigen?
- Governance: Welchen Beitrag kann KI zur resilienteren Governance leisten?

Die Ausgangsfragestellungen (Use Cases) für die Digital Future Challenge stellen Organisationen zur Verfügung, die die Studierenden bei der Ideenentwicklung auch begleiten – wir ziehen an einem Strang und dafür möchte ich allen Beteiligten herzlich danken!

Das vergangene Jahr hat uns gezeigt, wie schnell sich KI entwickelt. Mit der steigenden Geschwindigkeit müssen wir umso schneller sicherstellen, dass KI ethisch, transparent und fair entwickelt und eingesetzt wird. In Zeiten, in denen (generative) KI oft als riskant diskutiert wird, zeigen die Studierendenprojekte ermutigend, dass dies nicht unausweichlich, sondern gestaltbar ist. KI kann verantwortungsvoll genutzt werden und darüber hinaus sogar in vielen Anwendungsfällen Nachhaltigkeit und positive soziale Auswirkungen fördern. Ich bin zuversichtlich, dass die junge Generation als Entscheidungsträger*innen von morgen mithilfe von KI wichtige Impulse für die öko-digitale Transformation und die Nutzung von KI setzen wird.

In dieser Publikation erhalten Sie einen Einblick in die Ideen der fünf Finalteams: Die Teams haben

sich etwa überlegt, wie Energie auch in Zeiten von knappen Energieressourcen mithilfe von KI fair verteilt werden kann, wie die Verkehrsplanung auch vulnerable Gruppen integrieren kann, um mehr Sicherheit im Straßenverkehr für Kinder zu erzielen, oder wie Chatbots zur Inklusion und Barrierefreiheit beitragen können. Im Finale haben die Studierenden nicht nur ihre kreativen Lösungsvorschläge des Wintersemesters 2023/24 gepitcht, sie haben auch die gesellschaftliche Relevanz in ihren Präsentationen verdeutlicht und damit einen wichtigen Mehrwert der Digital Future Challenge herausgestellt: Es ist wichtig, sich zu engagieren, für die Gesellschaft und die digitale Zukunft!

Ich wünsche Ihnen eine inspirierende Lektüre.
Herzliche Grüße

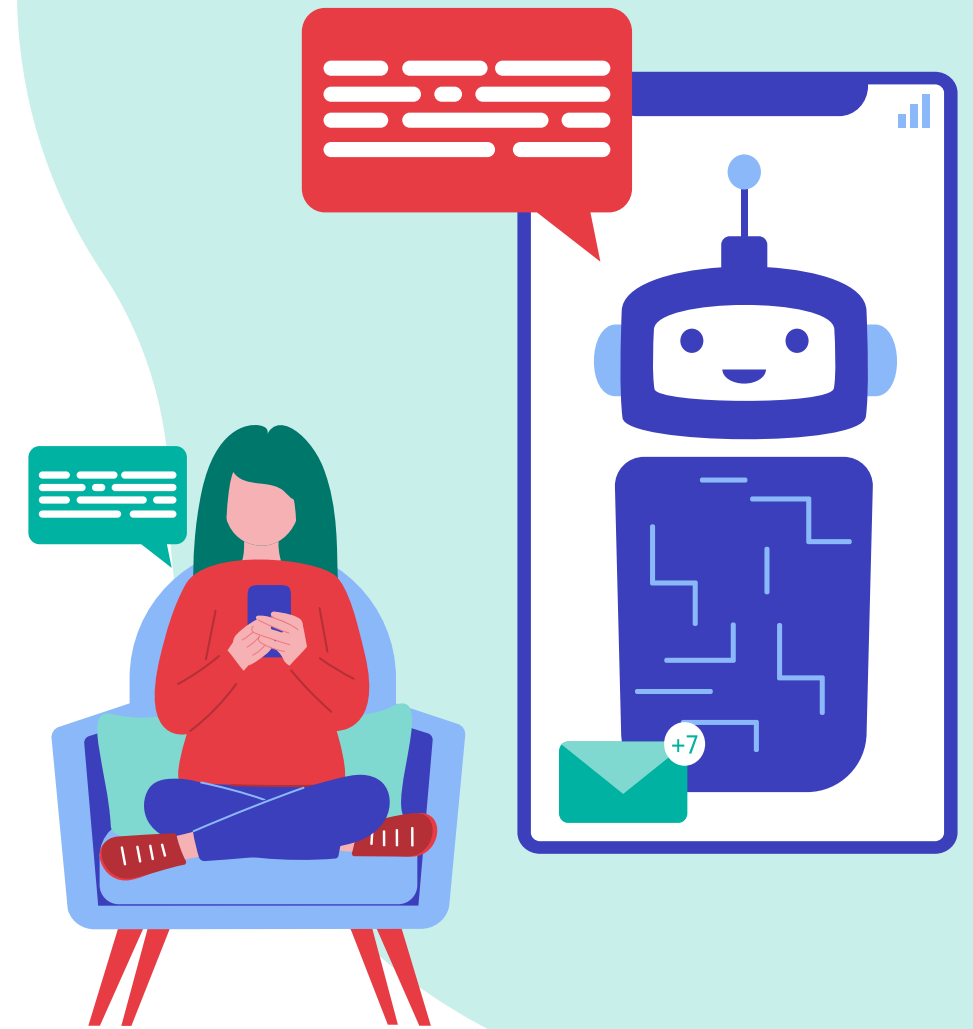
Marc Reinhardt,
Präsident Initiative D21 e. V.

DFC #4 auf einen Blick

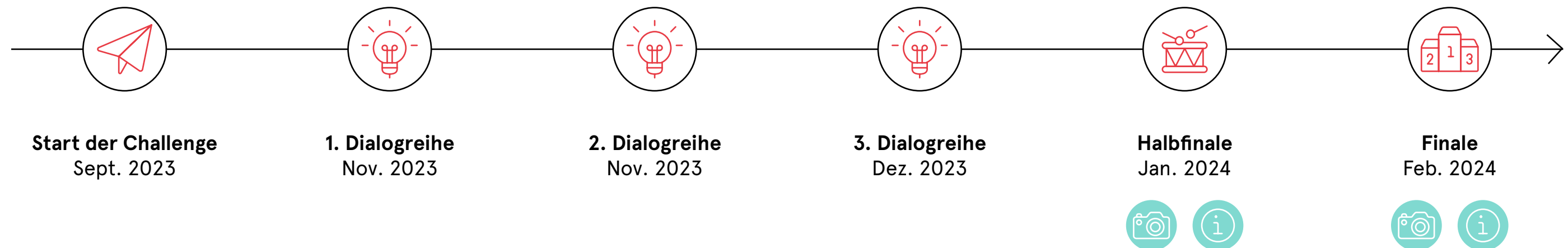
Die Integration von künstlicher Intelligenz ist mittlerweile in zahlreichen Bereichen unseres Lebens allgegenwärtig. Mit deren Anwendung gehen allerdings auch Herausforderungen einher, die eine sorgfältige Abwägung von Chancen und Risiken erfordern. Wie können wir die Anwendung von künstlicher Intelligenz in Übereinstimmung mit unseren ethischen Grundprinzipien gestalten? Diese Fragen des verantwortungsvollen Umgangs mit KI standen in der vierten Digital Future Challenge im Fokus. Die Grundlage des Wettbewerbs bilden neun ausgewählte Fallstudien, die einen verantwortungsvollen Umgang mit KI aus ökologischer, sozialer und Governance-Perspektive verdeutlichen. Interdisziplinäre und diverse Studierendenteams aus ganz Deutschland erarbeiten praxisnahe Lösungsvorschläge für die vorgestellten Use Cases.

Wie kann künstliche Intelligenz dazu beitragen, unsere Ressourcen effizienter zu nutzen und unser Klima zu schützen? Wie können KI-Systeme entwickelt werden, die soziale Auswirkungen berücksichtigen? Welche Rolle kann KI bei der Schaffung einer widerstandsfähigen Governance-Struktur spielen? Mit dem technologischen Fortschritt wächst auch die Pflicht von Unternehmen und Organisationen, sich dieser Fragen anzunehmen.

In dem Wettbewerb wird die gemeinsame Herangehensweise an Fragen und Gestaltung der digitalen Zukunft groß geschrieben: Politik, Unternehmen, Wissenschaft und Zivilgesellschaft engagieren sich.



Zeitlicher Ablauf der Digital Future Challenge



Use Cases



Ökologische Perspektive

| BertelsmannStiftung Zusammenreffen von Klimawandel und Digitalisierung: Welche Rolle spielen KI und New Work?



#AlmeetsNewWork
#OrgaTransformation #Alsustainability



Faires KI-Energiemanagement und ICH



#Alfairness #DemandSideManagement
#FairEnergyConsumption



Kann eine KI einen Mehrwert hinsichtlich der Ökobilanz für eine Logistiklösung schaffen oder nicht?



#AlcarbonFootprint #LogisticsAI
#AlforCommonGood



Soziale Perspektive



Human-friendly Automation – Analyse der Auswirkungen auf die Betroffenen von Automatisierung



#Alautomatization #SkillsShortage
#ChangeManagement



Zielgerichtete Datensammlung für den Einsatz von KI in der datenbasierten Verkehrsplanung



#UnbiasedData #AlforCommonGood
#SafeTrafficPlaning



Zugängliche und chancengleiche Conversational-AI-Systeme in der Interaktion zwischen Mensch und Technologie



#DigitalInclusion #ConversationalAI
#AldiscriminationPrevention



Governance-Perspektive



Kann und darf KI schon allein programmieren?



#LLMprogramming #AlasDeveloper
#AlqualityCheck



KI-basiertes Transaktionsmonitoring als Lösung für Geldwäsche-Prävention



#MoneyLaunderingPrevention
#TrustworthyAI #FraudDetection



Wie können Kund*innen eindeutig darüber informiert werden, dass sie es mit einer KI zu tun haben?



#Altransparency #ChatbotTransparency
#HumanMachineInteraction

Jury



Eva Achterhold
AI-Fairness-Expertin und
IT Consultant



Ben Brake
Abteilungsleiter Digital- und
Datenpolitik (BMDV)



Tobias Fausch
Chief Technology Officer
(BayWA AG)

TANDEM



Prof. Dr. Cordula Meckenstock
Honorarprofessorin für
Compliance und Internationales
Strafrecht



Maria Hinz
Digitalkoordinatorin
(BARMER)



Yara Hoffmann
Journalistin und
TV-Moderatorin



Simone Kaiser
Leiterin Center for Responsible
Research and Innovation
(CeRRI, Fraunhofer IAO)

TANDEM



Jessica Wulf
Research Associate
(CeRRI, Fraunhofer IAO)



Dr. Andreas Nauerz
Chief Technology Officer und
Executive Vice President
(Bosch Digital)



Prof. Dr. Frank Passing
Professor für Digitale
Transformation (IU)



Heidi Schultz
Development Manager
(SAP)



Dr. Christina Strobel
TU Hamburg,
Gründerin AlgoTrust

Mentor*innen



Charlotte Altmann
Expertin für verantwortungsvolles Unternehmertum und ethische Softwareentwicklung



Manuel Dietrich
Experte für AI Ethics und GDPR



Dr. Sergio Genovesi
Experte für KI und KI-Ethik



Jens-Rainer Jänig
Experte für CDR, Digitalisierung und Kommunikation



Laura Kaufmann
Expertin für digitales Marketing und digitale Transformation



Klara Krieg
Expertin für AI-Fairness



Alina Lorenz
Expertin für Machine Learning, KI und digitale Ethik



Bartosz Przybylek
Experte für CDR, Digitalisierung und Geschäftsmodellentwicklung



Dr. Tobias Rodemann
Experte für Energiemanagement



Felix von Roesgen
Experte für CDR, Sustainability und AI-Ethics



Isabelle Schlegel
Expertin für CDR, KI und digitale Ethik

Die Halbfinalteams und deren Ausarbeitungen im Überblick

DIGITAL FUTURE CHALLENGE #4 HALBFINALE
 15.01. 2024

KEYNOTES
 Lajla Fetic, KARIM MUSTAGHNA

AI-ALLIES
 AURA ALLY: KEINE CHANCE für HATESPEECH
 GEBÄRDEN-SPRACHE, SPRACH-SENSIBEL
 MATCHING-PLATTFORM CAREER-PILOT: neue AUFGABEN TINDERN, ONBOARDING KARRIERE- und ZUKUNFTS-PLANUNG!
 COMFORT-LEVEL: TAGESAKTUELL ANPASSEN!
 CHALLENGES und GAMIFICATION: STEUERUNG ÜBER BOX

EDU PIN
 STOP! KINDER SCHÜTZEN
 DATEN SAMMELN, ABER SICHER! ANONYM, BEREINIGT

IBMATCH
 LERNEN MIT INTERSEKTIONALEN DATEN!
 VORSICHT VOR BIAS REPRODUKTION! GEGEN DEN BIAS ANARBEITEN!

GRID HARMONY
 NUDGING FÜR NETZ-STABILITÄT
 PER APP! ANREIZE, TRANSPARENT

MY COMFORT
 SMARTER STROM-ZÄHLER, COINS SAMMELN
 FAIR FÖRDERFAKTOR, live REAGIEREN KÖNNEN

INTERSECTIONAL AI
 PERSONA-AI KANN DIVERSE KUND*INNEN IMITIEREN!
 BIAS BARRIERE INDEX, STEUERUNG DURCH TASK FORCE

TEAM NORDERSTEDT
 INTER-SKILLS, AUFGABEN-PROFILE
 FACH-KRÄFTE, HR FÜHRT DEN DIALOG

BEYOND THE BARRIER
 PERSONA-AI KANN DIVERSE KUND*INNEN IMITIEREN!
 ANDERE CAI TESTEN!

WONDA
 VERTEILUNGS-SCHLÜSSEL, TRANSPARENTE PREISANPASSUNG

TECHNOLOGIE als WERKZEUG
 KREATIVITÄT ist TRAINIERBAR
 KRITERIEN in HETEROGENEN TEAMS ENTWICKELN!
 FRAGEN Lieben GEGENFRAGEN
 Die Ziele müssen vom MENSCHEN KOMMEN!
 KI OFFEN, FAIR UND VERANTWORTUNGS-VOLL GESTALTEN!

Neugierde
 wie MESSEN wir ARMUT? GESTALTEN!
 BREIT, TIEF

Team Nordstern
 PERFEKT FIT finden

DeloitteStiftung Initiative D21
 Gefördert durch: Bundesministerium für Digitales und Verkehr

Jonas Heidebrecht



Die Finalteams und deren Ausarbeitungen im Überblick

An der vierten Edition nahmen über 50 Studierendenteams aus der ganzen Bundesrepublik teil und hatten bis Ende November 2023 ihre Ideen eingereicht. Eine Jury wählte fünf Finalist*innenteams, die ihre Projekte stetig weiterentwickelten und beim Finale am 8. Februar 2024 präsentierten:

- 
Platz 1: EduPin (Technische Universität München) mit einem digitalen Ansteck-Pin für Kinder, der Bewegungsdaten anonym aufzeichnet und Kommunen bei der datenbasierten, kindgerechten Verkehrsplanung hilft.
 - 
Platz 2: AI Allies (Ludwig-Maximilians-Universität München) mit einer inklusiven Gestaltung von Chatbots für den Kund*innenservice.
 - 
Platz 3: GridHarmony (Technische Universität Hamburg) mit einer App, die KI-gestützt Vorschläge für faire Stromverteilung macht und dabei nachhaltige Entscheidungen durch ein Stromguthaben belohnt.
- PowerBalance** (Technische Universität München) mit einer Energieplattform, die Haushalte unterstützt, ihren Verbrauch anzupassen, Kosten zu sparen und als Gemeinschaft die Energiewende voranzubringen.
- BeyondTheBarrier** (Hochschule für Technik und Wirtschaft Berlin) mit einer KI für biasfreie Interaktionen, die Kundenbedürfnisse der Conversational Als zusammenführt, die beispielsweise Chatbots zugrunde liegen.



Finale im Bundesministerium für Digitales und Verkehr am 8.2.2024, Fotograf: Phil Dera

„Ich bin beeindruckt von der Vielfalt und Tiefe der Ideen, die die Studierenden präsentierten. Die DFC bietet die perfekte Bühne, um die frischen Ideen diverser Studierendenteams mit den aktuellen Entwicklungen in Unternehmen zusammenzubringen und so digitale Ethik und unternehmerische Verantwortung mit Leben zu füllen. Es ist inspirierend zu sehen, wie sie die Herausforderungen der digitalen Transformation angehen und eine Zukunft gestalten wollen, die unseren ethischen Grundsätzen entspricht und das Wohl von Gesellschaft und Umwelt im Blick hat.“

Maria Hinz, Digital Koordinatorin (BARMER), Jurymitglied



Editorial

EduPin

AI Allies

GridHarmony

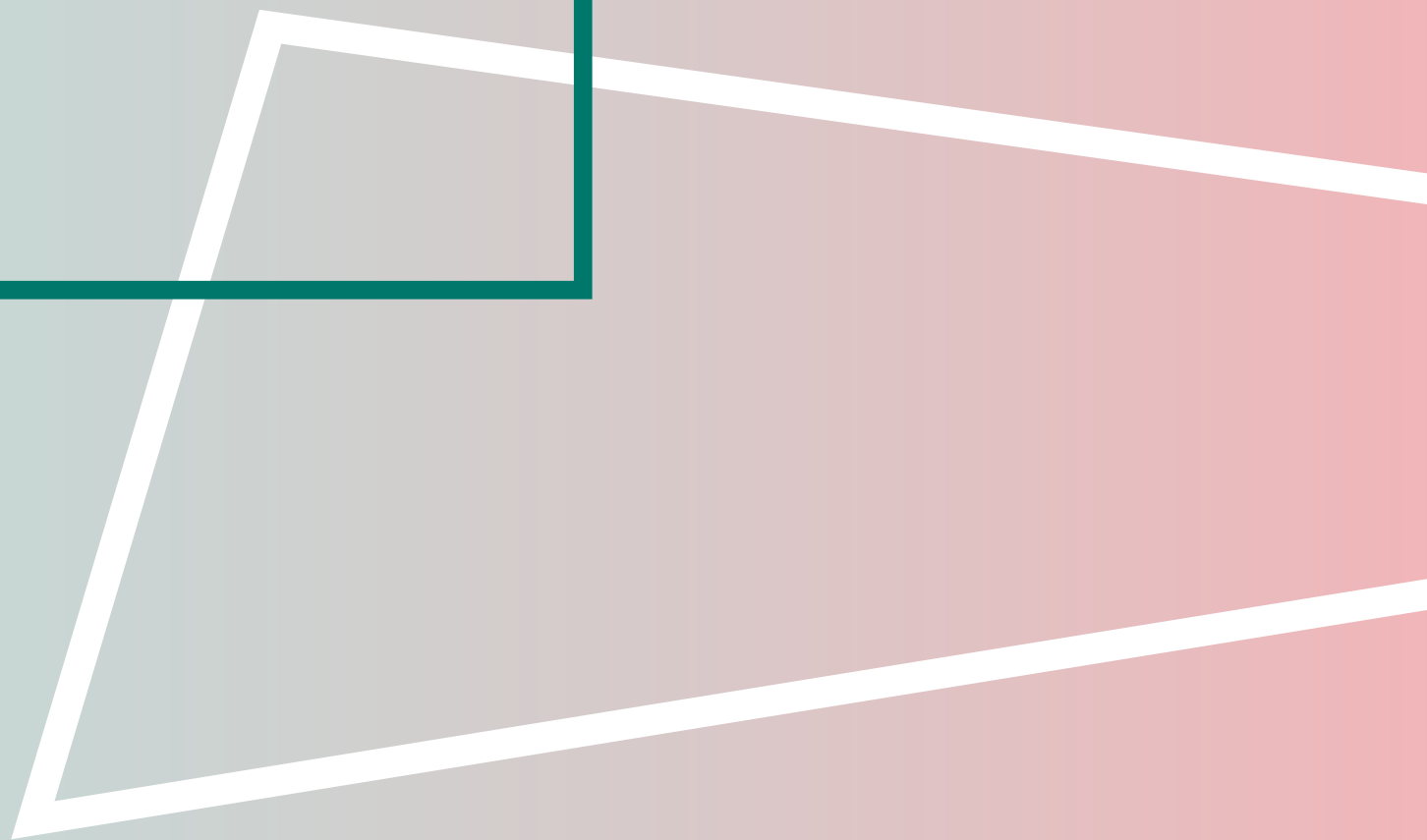
PowerBalance

Beyond
TheBarrier

Dialogreihe

Anhang

EduPin



EduPin – der smarte Verkehrsbegleiter für inklusive Kommunen



Team EduPin



David Javani
TU München



Tammy Schmidt
TU München



Alexander Wiener
TU München



Use Case: PD

Kernbotschaften/Abstract



Datenbasierte Verkehrsplanung verspricht effiziente und sichere Kommunen, insbesondere für nicht-motorisierte Verkehrsteilnehmer*innen. Sie basiert auf Bewegungsdaten, um die Nutzung der Infrastruktur zu verstehen und entsprechende Maßnahmen zu ergreifen. Die praktische Umsetzung datenbasierter Verkehrsplanung wird durch die unzureichende Verfügbarkeit repräsentativer Daten behindert. Insbesondere die Bewegungsdaten von Kindern wurden bisher kaum bis gar nicht erfasst, was zu einem Bias in den Daten führt und die Effektivität der Planung beeinträchtigt.



Der EduPin ist eine innovative Lösung, um Bewegungsdaten von Kindern zu erfassen und gleichzeitig eine edukative Komponente für die Verkehrssicherheit zu bieten. Durch den Einsatz von künstlicher Intelligenz ermöglicht er es, Kinder spielerisch Verkehrsregeln zu lehren und potenzielle Gefahren im Straßenverkehr zu erkennen. Dadurch wird nicht nur die unmittelbare Verkehrssicherheit erhöht, sondern auch ein nachhaltiger Lerneffekt erreicht.



Bewährte Technologie, ein strenger Katalog an Sicherheitskriterien, spielerische Interaktion, Sprachenauswahl und individualisierbares Design sorgen dafür, dass sowohl Eltern als auch Kinder vom EduPin überzeugt werden. Darüber hinaus beherbergt der EduPin das Potenzial, auf weitere Zielgruppen ausgeweitet zu werden. Neben dem sofortigen Mehrwert im Straßenverkehr für die tragende Person ermöglichen die durch den EduPin gesammelten Daten, die Verkehrsplanung in der Kommune auch mittel- und langfristig inklusiver und unter Berücksichtigung aller Bevölkerungsgruppen zu gestalten.

Herausforderung

In der Theorie besitzt datenbasierte Verkehrsplanung das Potenzial, effizient und für alle Bürger*innen sichere Kommunen zu bauen. Dabei sollte besonderes Augenmerk auf den nicht-motorisierten Umweltverbund gelegt werden, da Fußgänger*innen und Fahrradfahrer*innen im Straßenverkehr am meisten gefährdet sind. Sie nehmen außerdem oft Abkürzungen, Um- und Schleichwege, die nicht Teil des offiziellen Straßennetzes sind und daher bei Planung, Wartung und Gefahrenprävention nicht berücksichtigt werden. Bewegungsdaten zeigen die tatsächlichen Wege auf, die Menschen zurücklegen, und geben Aufschluss darüber, wie die gegebene Infrastruktur genutzt wird.

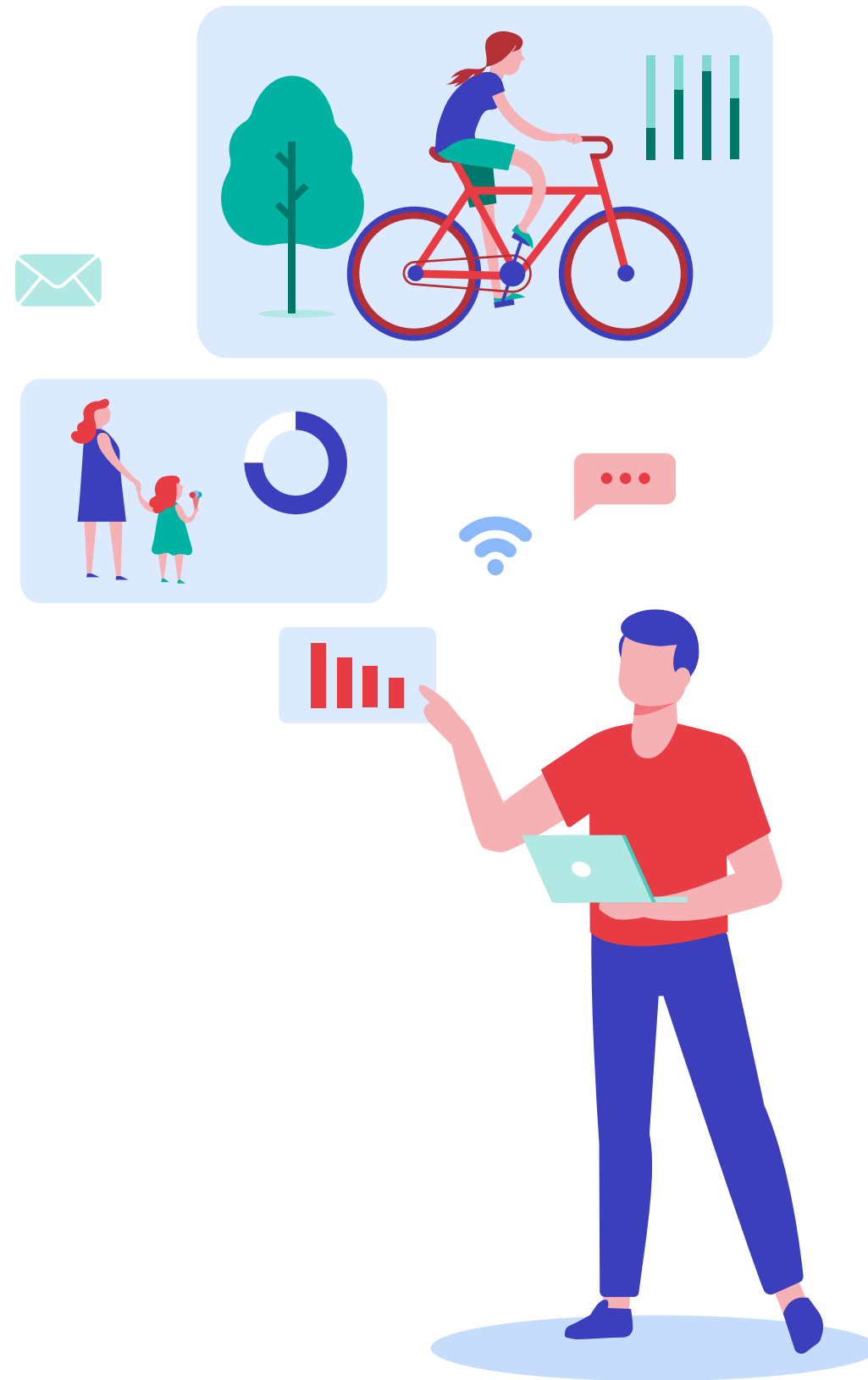
In der Praxis scheitert datenbasierte Verkehrsplanung vorwiegend an der Verfügbarkeit repräsentativer Daten. Die technischen Möglichkeiten sind gegeben: Moderne Smartphones können zurückgelegte Strecken immer und überall aufzeichnen. Es gibt erste Pilotprojekte, zum Beispiel in der Stadt Wiesloch (Wiesloch 2023), bei denen Teilnehmende ihre gesammelten Handydaten an die Kommune spenden.

Doch die Bereitschaft und/oder Mittel, solche Daten zu teilen, hängen stark von demografischen Faktoren ab. Laut Anita Klingel, Expertin für AI Governance von PD, nehmen an freiwilligen Projekten zur handybasierten Datensammlung überproportional viele junge Erwachsene teil, während Kinder und ältere Menschen aufgrund fehlender

technischer Geräte oder erhöhter Sicherheitsbedenken unterrepräsentiert sind. Dabei sollte genau auf diese vulnerablen Gruppen bei der Verkehrsplanung besonders Rücksicht genommen werden. Vor allem Kinder sind im Straßenverkehr weniger sichtbar, nehmen besonders häufig Abkürzungen und kennen nicht alle Straßenschilder und Verkehrsregeln. Im Jahr 2021 verunglückte im deutschen Straßenverkehr durchschnittlich alle 24 Minuten ein Kind unter 15 Jahren, wovon ein Großteil (59,3 %) zu Fuß oder mit dem Fahrrad unterwegs war (Statistisches Bundesamt 2020).

Selten nehmen Kinder an herkömmlichen Aktionen zur Datensammlung teil, da viele kein eigenes Smartphone besitzen (Statistisches Bundesamt 2023), und wenn doch, zögern aus Sicherheitsbedenken die meisten Eltern, Daten ihrer Kinder zu teilen. Noch schwieriger wird es, Eltern zur Teilnahme an einem solchen Projekt zu überzeugen, wenn sprachliche und kulturelle Barrieren dazukommen. Das Ergebnis: Bisherige gesammelte Daten stammen von einer sehr homogenen Bevölkerungsgruppe und führen dadurch zu einem Bias in den Daten, der sich bis in die Verkehrsplanung hinein erstrecken könnte.





Analyse

Angesichts der erkannten Problematik besteht dringender Bedarf, die Erfassung von Verkehrsdaten so attraktiv und inklusiv wie möglich zu gestalten. Dies dient nicht nur der allgemeinen Verbesserung der Verkehrsplanung, sondern auch dem Schutz und der Berücksichtigung von unterrepräsentierten Gruppen (wie Kindern) im Straßenverkehr. Eine datenbasierte Verkehrsplanung, die alle Verkehrsteilnehmer*innen einbezieht, kann dazu beitragen, Unfälle zu verhindern und die Sicherheit im Straßenverkehr zu erhöhen. Es ist daher von entscheidender Bedeutung, innovative Lösungen zu entwickeln, die nicht nur Daten erfassen, sondern auch einen präventiven und edukativen Ansatz verfolgen, um insbesondere vulnerable Gruppen zu schützen.

Daraus ergibt sich die Notwendigkeit eines Instruments wie des EduPins, der nicht nur Bewegungsdaten von Kindern anonymisiert aufzeichnet, sondern auch eine edukative Komponente bietet. Der EduPin ermöglicht es, Kinder spielerisch über potenzielle Gefahrenstellen im Straßenverkehr zu informieren und sie in Echtzeit Verkehrsregeln zu lehren. Diese Lösung trägt nicht nur mittel- und langfristig zur Verbesserung der Verkehrsplanung bei, sondern bietet auch einen unmittelbaren Nutzen für die Sicherheit und das Wohlergehen von Kindern im Straßenverkehr.

Zusätzlich zu den direkten Vorteilen für Kinder eröffnet der EduPin auch Möglichkeiten für eine

breitere gesellschaftliche Einbindung und den Ausbau von Bildungsmaßnahmen im Verkehrssektor. Eine inklusive Verkehrsplanung kann zu einer gesteigerten Attraktivität von Fuß- und Radverkehr führen, was wiederum zu einer Reduzierung des Autoverkehrs, weniger Staus, Verkehrsunfällen und Treibhausgasemissionen beitragen kann. Darüber hinaus fördert die Verkehrsbildung von Kindern einen nachhaltigen Lerneffekt und trägt langfristig zur Sicherheit im Straßenverkehr bei.

Die Grundlage des EduPins beruht auf bewährten Technologien, die bereits für Autofahrer*innen existieren. Allerdings wurde das Konzept für nicht-motorisierte Verkehrsteilnehmer*innen angepasst, insbesondere für Kinder. Der EduPin ist leicht an Jacken, Schulranzen oder Fahrradlenkern anzubringen und bietet durch die Nutzung eines eigenen Geräts eine größere Anonymität und Unabhängigkeit von großen Konzernen. Die Integration von akustischen Warnsignalen sorgt dafür, dass das Kind nicht visuell abgelenkt wird.

In Bezug auf Erweiterungen bietet der EduPin auch Potenzial für andere Bevölkerungsgruppen, darunter erwachsene Neubürger*innen, die mit dem deutschen Verkehrssystem nicht vertraut sind, sowie Menschen mit Sehbehinderung, einschließlich älterer Personen. Eine mögliche Kooperation mit Fahrrad- und Rucksackherstellern könnte die Distribution der Geräte erleichtern und die Reichweite der Lösung erweitern.

Handlungsempfehlungen/Lösung

Unser EduPin ist ein ansteckbarer digitaler Button, der Bewegungsdaten der Kinder anonymisiert aufzeichnet, ohne dass ein Smartphone erforderlich ist. Über eine integrierte Kamera erkennt der EduPin KI-basiert Verkehrsschilder und gibt über einen Lautsprecher dem Kind in einfacher Sprache Erklärungen. Der erste Prototyp kann neben Deutsch auch Farsi, Türkisch und Ukrainisch, die laut Anita Klingel meistgesprochenen Sprachen im deutschen Klassenzimmer, um so bereits von Anfang an Inklusion zu fördern. In der gewünschten Sprache erhält das Kind in Echtzeit Hinweise zu potenziellen Gefahrenstellen und lernt Verkehrsregeln kennen. Es wurde sich bewusst gegen einen Bildschirm oder Touchscreen entschieden, damit der EduPin ein Begleiter und keine Ablenkung ist. Abbildung 1 bildet den ersten Prototypen des EduPins ab und demonstriert den technischen Aufbau.

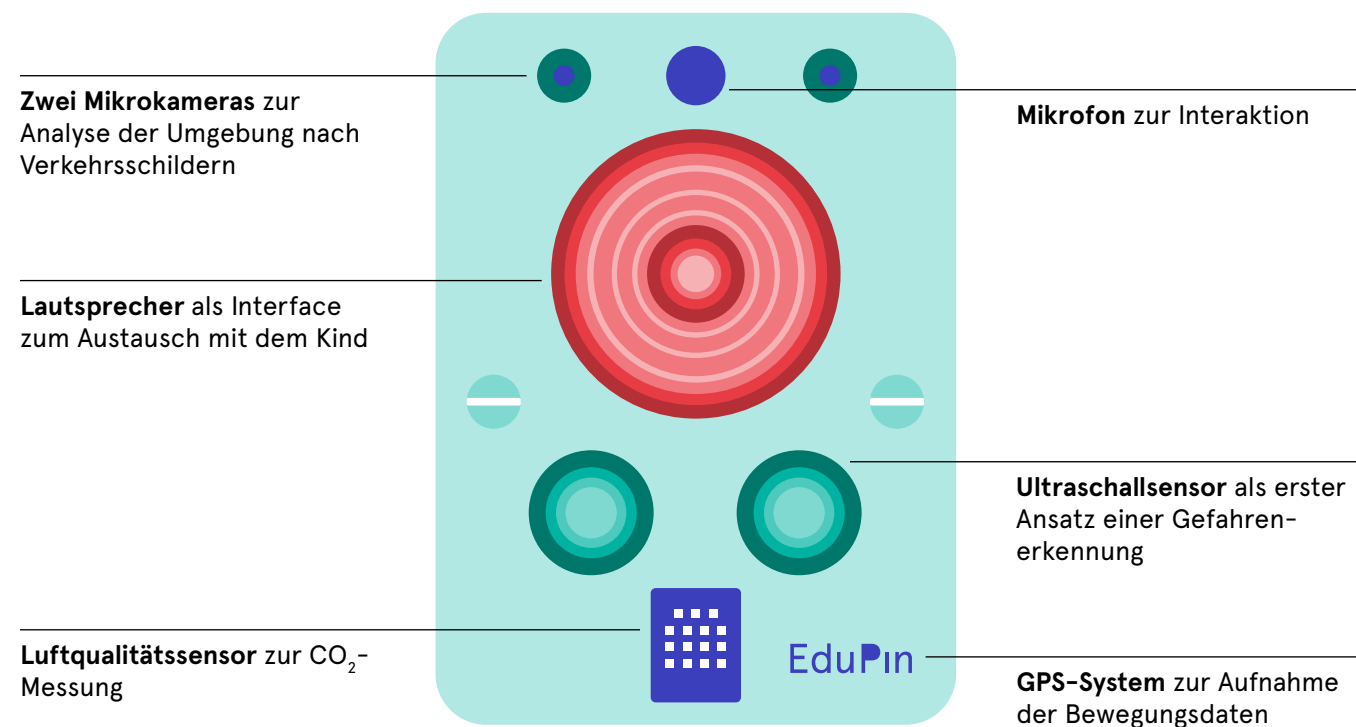


Abb. 1 – Der technische Aufbau des ersten Prototypen zum EduPin

Diese Lösung bietet nicht nur Aussicht auf eine mittel- bis langfristig erhöhte Verkehrssicherheit als Folge von datenbasierter Straßenplanung, sondern ermöglicht Kindern und ihren Eltern auch einen unmittelbaren Nutzen. Das Give-and-Take-Prinzip sieht vor, dass Kinder einen digitalen, edukativen Verkehrsbegleiter auf Leihbasis erhalten und im Gegenzug ihre Bewegungsdaten anonym und verschlüsselt an die Kommune spenden.

Neben den Eltern, die darüber entscheiden, ob ihr Kind mit dem EduPin ausgestattet wird oder nicht, sind die zweite Zielgruppe die Kinder selbst. Eine kleine Marktrecherche hat ergeben, dass zwei Aspekte ausschlaggebend sind, um Kinder vom EduPin zu überzeugen. Zum einen spielt das Design eine große Rolle. Über verschiedene Hüllen kann der EduPin individualisiert und optisch an den Geschmack des tragenden Kindes angepasst werden, um den erforderlichen Coolness-Faktor zu erhöhen.

Zum anderen steigert Gamification die Attraktivität des EduPins, vermeidet einen Automation Bias (also das blinde Verlassen auf Technik) und verstärkt den Lerneffekt. Dazu ist ein interaktives Frage-Antwort-Spiel geplant, bei dem der EduPin das Kind regelmäßig nach dem korrekten Verhalten bei Verkehrsschildern fragt. Über ein eingebautes Mikrofon kann das Kind spielerisch mit dem EduPin interagieren. Lob und Gamificationelemente helfen dabei, das Kind nachhaltig zu motivieren. Es ist wichtig zu erwähnen, dass das Mikrofon nur aktiv ist, nachdem der EduPin eine Frage gestellt hat oder mit „Hey EduPin“ direkt angesprochen wurde, so wie man es von Smartphone-Assistenten bereits kennt.

Da der EduPin hauptsächlich von Kindern genutzt wird, einer besonders vulnerablen Zielgruppe, ist es entscheidend, eine sichere Nutzung zu gewährleisten. Dazu gehören rechtliche Rahmenbedingungen wie die DSGVO-konforme Speicherung der Daten und die Unterzeichnung einer Haftungserklärung durch die Erziehungsberechtigten. Auch zukünftige rechtliche Regulierungen wie der EU AI Act müssen berücksichtigt werden (European Parliament 2023).

Technische Sicherheitskriterien umfassen keine Echtzeitübertragung, die Übertragung der Daten nach 24 Stunden mit Ende-zu-Ende-Verschlüsselung, die Speicherung auf einem lokalen Server, Edge Computing, einen On/Off-Knopf mit Lichtanzeige sowie eine automatische Löschung der ersten und letzten hundert Meter des aufgezeichneten Weges.

„Den EduPin soll es auch als Leihmodell geben. Damit schaffen wir Chancengleichheit, wenn Kinder die Geräte über ihre Schulen erhalten. Kommunen sparen Geld, weil sie nicht laufend neue Geräte anschaffen müssen, und wir vermeiden unnötigen Elektroschrott im Sinne der Circular Economy.“

Tammy Schmidt, EduPin

Fazit

Das Konzept der datenbasierten Verkehrsplanung verspricht eine effiziente und sichere Gestaltung unserer Städte und Gemeinden, insbesondere im Hinblick auf nicht-motorisierte Verkehrsteilnehmer*innen, die zu Fuß oder mit dem Fahrrad unterwegs sind. Es basiert auf der Nutzung von Bewegungsdaten, um die tatsächliche Nutzung der Infrastruktur zu verstehen und entsprechende Maßnahmen zu ergreifen. Trotz des Potenzials dieser Methode stoßen Projekte zur datenbasierten Verkehrsplanung jedoch auf Herausforderungen in der Praxis.

Die Hauptprobleme liegen in der Verfügbarkeit repräsentativer Daten, die für eine ganzheitliche Planung erforderlich sind. Besonders Kinder und ältere Menschen sind unterrepräsentiert, was zu einem Bias in den Daten führt und die Effektivität der Planung beeinträchtigen kann. Die Lösung ist der EduPin, ein innovatives Gerät, das Bewegungsdaten anonymisiert aufzeichnet und gleichzeitig eine edukative Komponente für Kinder bietet.

Der EduPin ermöglicht es Kindern, spielerisch Verkehrsregeln zu lernen und potenzielle Gefahren im Straßenverkehr zu erkennen, ohne auf persönliche Smartphones angewiesen zu sein. Durch gamifizierte Elemente und ein ansprechendes Design wird die Akzeptanz bei Kindern gefördert. Gleichzeitig werden weitreichende Datenschutz- und Sicherheitsbedenken berücksichtigt.

Die Implementierung des EduPins könnte nicht nur kurzfristig die Verkehrssicherheit für Kinder verbessern, sondern auch langfristig zu einer inklusiven Verkehrsplanung beitragen, die alle Bevölkerungsgruppen einbezieht. Eine breitere gesellschaftliche Einbindung des EduPins sowie potenzielle Erweiterungen auf andere Zielgruppen tragen dazu bei, Kommunen für alle Verkehrsteilnehmer*innen sicher und lebenswert zu gestalten.



Quellen:

Statistisches Bundesamt (2020): Verkehrsunfälle: Kinderunfälle im Straßenverkehr. o.O.
 Meyer, Helmut (2023): Elterntaxi: Besser nicht mit dem Auto zur Schule. Online verfügbar unter: <https://is.gd/GGeCib> (Abfrage: 24.11.2023).
 European Parliament (2023): EU AI Act: first regulation on artificial intelligence. Online verfügbar unter: <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence> (Abfrage: 24.11.2023).
 Statistisches Bundesamt (2023): Smartphone-Besitz von Kindern in Deutschland im Jahr 2022 nach Altersgruppen. Online verfügbar unter: <https://is.gd/AilUyM> (Abfrage: 29.11.2023).
 Wiesloch (2023): Radverkehrsdaten-App. Online verfügbar unter: <https://www.wiesloch.de/pb/radverkehrsdatenapp.html> (Abfrage: 02.02.2024).



Editorial

EduPin

AI Allies

GridHarmony

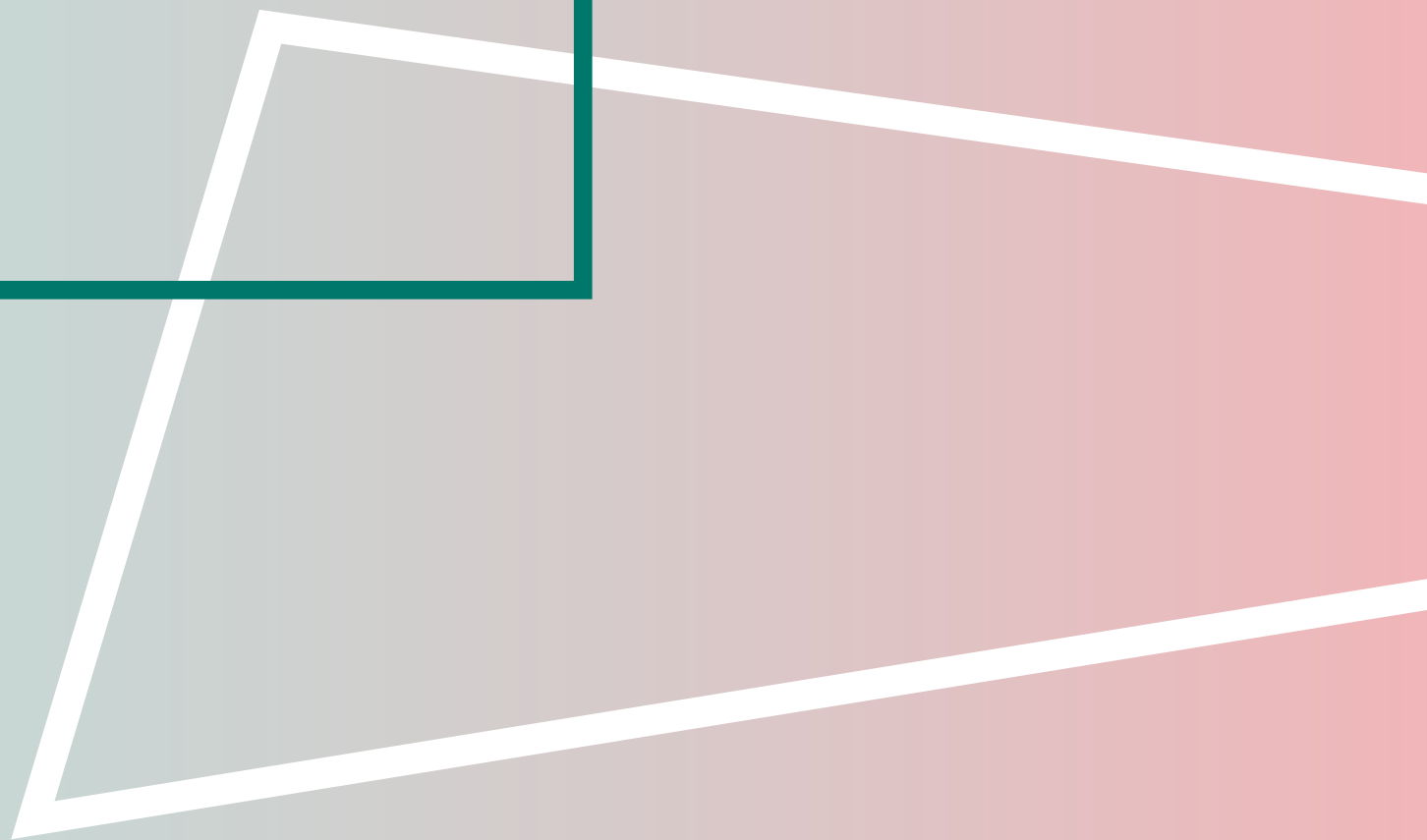
PowerBalance

Beyond
TheBarrier

Dialogreihe

Anhang

AI Allies



Conversational AI für alle: Barrierefreiheit und Gleichberechtigung in O₂ Telefónicas „Aura“



Team AI Allies



Nina Mandl
LMU München



Leon Oskui
LMU München



Mert Türkekul
LMU München



Kernbotschaften/Abstract



„Aura“, die Conversational AI von O₂ Telefónica, wird für umfassende Barrierefreiheit und Gleichberechtigung optimiert. Dies beinhaltet die Integration von Technologien zur Unterstützung von Menschen mit unterschiedlichen kognitiven und physischen Fähigkeiten. Ein besonderer Fokus liegt in der Unterstützung von hörbehinderten Menschen durch den Einsatz von Technologien für die Einbindung von Gebärdensprache in die Benutzer*innenschnittstelle.



Die Weiterentwicklung von „Aura“ zielt auf eine verbesserte Mensch-Computer-Interaktion ab. Dies umfasst eine intuitivere Benutzer*innenoberfläche, personalisierte Ansätze und die Einbeziehung von Audio-Optionen und Spracheingabefunktionen, um die Interaktion zu vereinfachen und zu personalisieren.



Ein Schwerpunkt liegt auf dem respektvollen Umgang mit allen Nutzer*innen. Durch fortgeschrittene Algorithmen zur Erkennung und Handhabung von Hate Speech und unangemessenen Inhalten strebt „Aura“ danach, eine sichere und positive Kommunikationsumgebung zu schaffen.

Use Case: O₂ Telefónica

Mitwirkende: Sarah Rachut, Helene von Schwichow, Markus Siewert

Herausforderung

Die Entwicklung von „Aura“, der Conversational AI von O₂ Telefónica, steht vor komplexen Herausforderungen, die technologische Innovation, Barrierefreiheit, ethische Verantwortung und soziale Inklusion umfassen. Einerseits erfordert die Integration fortschrittlicher Technologien, dass „Aura“ schnell, effizient und in der Lage ist, komplexe Anfragen zu bearbeiten. Andererseits muss die Barrierefreiheit gewährleistet werden. BIK für Alle (o. J.), der Deutsche Blinden- und Sehbehindertenverband e. V. (o. J.) und Holzer (2023) weisen darauf hin, dass digitale Technologien oft nicht für Menschen mit Seh- oder Hörbehinderungen optimiert sind, und führen weiterhin aus, wie man hier Fortschritte erzielen könnte. Insbesondere zeigen sie auf, dass die Einbeziehung von Gebärdensprache in digitale Interaktionsszenarien, wie sie von Pardasani et al. (2018) oder Apuzzo und Burrelli (2022) untersucht wurden, sowie auch Alternativtexte im Code zur Erkennung und Verbalisierung von Grafiken entscheidend sind.

Die Optimierung der Mensch-Computer-Interaktion erfordert eine Benutzer*innenoberfläche, die nicht nur intuitiv und benutzer*innenfreundlich ist, sondern auch die Bedürfnisse von Menschen mit

unterschiedlichen kognitiven Fähigkeiten berücksichtigt. Expert*innen wie Torres et al. (2019) heben die Wichtigkeit einer zugänglichen und ansprechenden Schnittstelle hervor. Die Herausforderung liegt darin, eine Schnittstelle zu schaffen, die sowohl einfache als auch komplexe Anfragen effizient handhabt, während sie gleichzeitig für alle Nutzer*innengruppen zugänglich ist.

Ein wesentlicher Aspekt der ethischen Verantwortung ist der respektvolle Umgang mit Nutzer*innen- und -interaktionen. Angesichts der Zunahme von Online-Hassrede und Diskriminierung ist es entscheidend, dass „Aura“ Mechanismen zur Erkennung und Vermeidung solcher Inhalte integriert. Die aktuelle Forschung zeigt hierzu erste Möglichkeiten zur Hate Speech Detection auf (Mandl 2020; Schmidt und Wiegand 2017; Sontheimer et al. 2011; Jahan et al. 2023; Mohiyaddeen und Siddiqi 2021). Es ist wichtig, ein System zu entwickeln, das nicht nur technisch leistungsfähig, sondern auch in der Lage ist, eine sichere und respektvolle Kommunikationsumgebung zu gewährleisten. Ein wichtiger Aspekt einer AI-gestützten Reduzierung von Hate Speech ist aber auch, nicht angemessene Einschränkungen der Meinungsfreiheit zu vermeiden.

Die Herausforderung besteht darin, diese verschiedenen Aspekte in einem harmonischen System zu vereinen, das technologisch fortschrittlich, zugänglich, fair und ethisch verantwortungsbewusst ist. Dies erfordert nicht nur fortlaufende technologische Entwicklungen, sondern auch eine kontinuierliche Auseinandersetzung mit ethischen Fragen und eine Anpassung an sich ändernde gesellschaftliche Normen und Nutzer*innenbedürfnisse. Die Entwicklung von „Aura“ ist daher ein fortlaufender Prozess, der eine enge Zusammenarbeit zwischen Technolog*innen, Ethiker*innen und Nutzer*innen erfordert, um eine inklusive, gerechte und verantwortungsvolle Conversational AI zu schaffen.



Analyse

Die Transformation von O₂ Telefónicas Conversational AI „Aura“ zu „Aura Ally“ (siehe Abb. 2) verkörpert ein tiefes Engagement für die Schaffung einer zugänglichen, inklusiven und fairen digitalen Welt. Unsere Vision innerhalb der Digital Future Challenge strebt danach, die Interaktion zwischen Mensch und Maschine durch die Schwerpunkte Barrierefreiheit, Chancengleichheit und respektvollen Umgang neu zu definieren. Durch die Implementierung von Technologien wie einem Gebärdensprach-Avatar und die Bereitstellung personalisierter Interaktionsmöglichkeiten adressieren wir proaktiv die Bedürfnisse von Menschen mit diversen kognitiven und physischen Fähigkeiten.

Die Notwendigkeit für solche Innovationen wird durch die offensichtlichen Lücken in der Zugänglichkeit digitaler Dienste unterstrichen. Es ist erkennbar, dass viele bestehende Systeme nicht ausreichend auf die Bedürfnisse aller Nutzer*innengruppen ausgerichtet sind. Dies unterstreicht die Dringlichkeit, unsere Ansätze zur Gestaltung von Conversational AI kontinuierlich zu überdenken und anzupassen, um eine inklusive und barrierefreie digitale Umgebung zu schaffen. Zusätzlich zu diesen technischen Anpassungen hat uns das Feedback der Jury im Halbfinale dazu motiviert, noch tiefer in die Themen Fairness und Inklusion einzutauchen und unsere Lösung entsprechend zu verfeinern.

„Gerade für gehörlose Menschen sind Mimik, Mundbilder, Emotion und Körpersprache sehr wichtig. Genau das fehlt aktuellen Avataren aber bisher – sie wirken in der visuellen Wahrnehmung nicht menschlich genug. Daran arbeiten wir.“

Nina Mandl, AI Allies

Abb. 2 – Avatar: „Aura Ally“





Analyse der Problematik und Herleitung der Lösung

Der Kern unserer Herausforderung liegt in der Schaffung einer Conversational AI, die nicht nur technisch ausgereift, sondern auch für alle Nutzer*innengruppen zugänglich und verständlich ist. Dies beinhaltet die Entwicklung einer Benutzer*innenoberfläche, die intuitiv und benutzer*innenfreundlich ist und gleichzeitig Menschen mit diversen körperlichen Einschränkungen berücksichtigt. Die Integration eines Gebärdensprach-Avatars und von Alternativtexten sind direkte Antworten auf diese Bedürfnisse, die zugleich das digitale Erlebnis für Menschen mit Hör- oder Sehbehinderungen revolutionieren.



Lösungsansätze und Konzepte

Um die Herausforderungen rund um digitale Barrierefreiheit, ethische Verantwortung und soziale Inklusion zu adressieren, setzen wir auf einen integrativen Ansatz bei der Entwicklung von „Aura Ally“. Unser Ziel ist es, eine Conversational AI zu realisieren, die technologisch fortschrittlich und zugleich für alle Nutzer*innengruppen inklusiv ist. Dazu gehört die Schaffung einer intuitiven Benutzer*innenoberfläche, die personalisierte Interaktionsmöglichkeiten bietet und somit eine individuell angepasste Nutzungserfahrung ermöglicht.

Im Zuge dessen konzentrieren wir uns auf die Optimierung der Mensch-Maschine-Interaktion, um eine natürliche und effiziente Kommunikation zu gewährleisten. Darüber hinaus streben wir an, durch den Einsatz von Technologien und spezifischen Anpassungen die digitale Zugänglichkeit für alle zu revolutionieren. Ein weiterer Schwerpunkt liegt auf dem respektvollen Umgang innerhalb der digitalen Kommunikation, um eine positive und sichere Austauschumgebung zu fördern.



Gesamtgesellschaftliche Relevanz

Die Weiterentwicklung zu „Aura Ally“ trägt entscheidend zur gesellschaftlichen Digitalisierung bei, indem sie nicht nur die Nutzer*innenerfahrung für eine diverse Basis verbessert, sondern auch als Modell für ethische Verantwortung und soziale Inklusion dient. Durch die Berücksichtigung der älteren Personengruppe durch Hilfestellungen wie Erklärvideos und anpassbare Schriftgrößen adressieren wir zudem eine oft vernachlässigte Nutzer*innengruppe.

Insgesamt setzt „Aura Ally“ neue Standards in der Conversational-AI-Branche und demonstriert das transformative Potenzial von Technologie, eine inklusive und gerechte Gesellschaft zu fördern. Unsere umfassende Herangehensweise, die sowohl technische Innovation als auch ein tiefes Verständnis für Nutzer*innenbedürfnisse vereint, dient als Leitfaden für zukünftige Entwicklungen in der digitalen Welt.

Diese umfassende Herangehensweise soll „Aura Ally“ als maßgebliche Lösung etablieren, die sowohl technische Exzellenz als auch ein tiefgreifendes Verständnis für die Diversität der Nutzer*innenbedürfnisse vereint, um eine inklusivere und gerechtere digitale Zukunft voranzutreiben.



Handlungsempfehlungen/Lösung

Die Neugestaltung von O₂ Telefónicas Conversational AI „Aura“ signalisiert unser Engagement für den Aufbau einer digitalen Welt, die für jeden zugänglich, inklusiv und gerecht ist. Indem wir uns auf die Verbesserung der Mensch-Maschine-Interaktion, die Förderung von Barrierefreiheit und Chancengleichheit sowie den respektvollen Umgang konzentrieren, tragen wir zu einer signifikanten Verbesserung der digitalen Erfahrung bei. Unser Ansatz integriert Gebärdensprache und Alternativtexte, um ein umfassend inklusives Chatbot-Erlebnis zu bieten, und berücksichtigt dabei die speziellen Bedürfnisse von Menschen mit unterschiedlichen Fähigkeiten, einschließlich älterer Nutzer*innengruppen.

Im Rahmen unseres Projekts haben wir „Aura Ally“ als Avatar konzipiert, der als eine erweiterte Assistentenfunktion für den bestehenden Chatbot „Aura“ von O₂ Telefónica fungiert. „Aura Ally“ bildet ein innovatives Bindeglied zwischen künstlicher Intelligenz und menschlichen Nutzer*innen. Der Avatar hebt sich durch seine Fähigkeit hervor, nicht nur akustische Kommunikation zu unterstützen, sondern auch visuelle Kommunikationsmodi, insbesondere die Gebärdensprache, anzubieten. Dadurch wird eine inklusive und umfassende Interaktionsmöglichkeit geschaffen, die sowohl hörende als auch gehörlose Benutzer*innengruppen effektiv anspricht und einbezieht. Diese Dualität ermöglicht es, sowohl hörende als auch gehörlose Nutzer*innengruppen anzusprechen und zu bedienen.

Darüber hinaus wurden von uns innovative Designvorschläge für die Nutzer*innenschnittstellen in mobilen (Abb. 3) und Web-Anwendungen (Abb. 4 und 5) entwickelt, um die Interaktion zwischen Benutzer*innen und Technologie zu optimieren und eine umfassende Benutzer*innenerfahrung zu gewährleisten.

Verbesserung der Mensch-Maschine-Interaktion

Unsere Lösung umfasst eine Audio-Option und eine Spracheingabe-Funktion, die eine intuitive und benutzer*innenfreundliche Schnittstelle für „Aura“ schaffen. Durch das Einholen von Nutzer*innenfeedback und die Anpassung an sich ändernde gesellschaftliche Normen stellen wir sicher, dass „Aura“ stets den Anforderungen und Wünschen der Nutzer*innen entspricht. Die Einbindung eines Erklärvideos und die Möglichkeit zur Anpassung von Schriftgröße und -dicke erleichtern insbesondere älteren Menschen die Nutzung, wodurch die Interaktion mit der Technologie intuitiver und zugänglicher wird.

Barrierefreiheit und Chancengleichheit

Durch die Implementierung von Gebärdensprache-Avataren wie „Aura Ally“ und die Bereitstellung von Alternativtexten im Code ermöglichen wir einen gleichberechtigten Zugang zu Informationen und Diensten. Die Berücksichtigung verschiedener Sprachen und Dialekte erweitert die Reichweite von

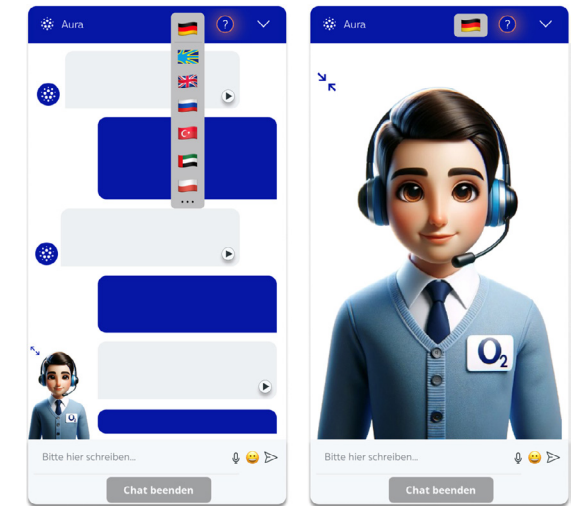


Abb. 3 – Sprachmodus (Deutsch) – Mobile Ansicht

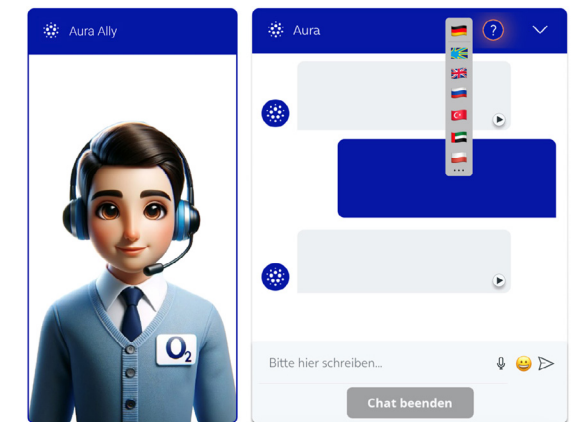


Abb. 4 – Sprachmodus (Deutsch) – Web-Ansicht

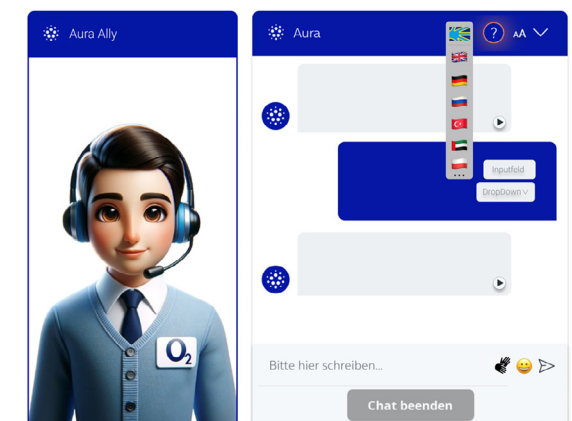
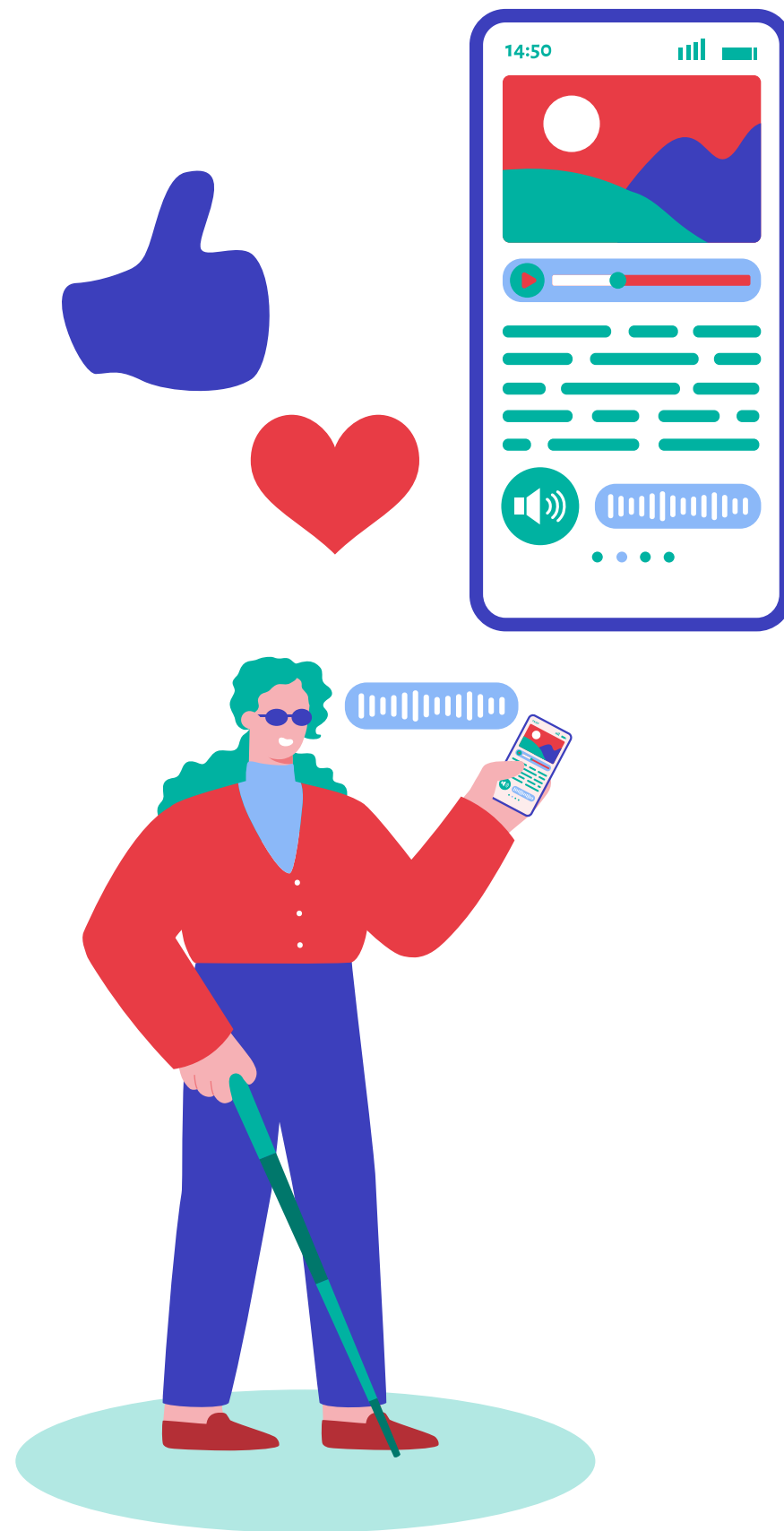


Abb. 5 – Gehörlosenmodus – Web-Ansicht



„Aura“ und ermöglicht eine individuellere Interaktion mit der AI. Diese Maßnahmen spiegeln unser Bestreben wider, digitale Hürden abzubauen und ein inklusives Erlebnis für alle Nutzer*innen zu schaffen.

Respektvoller Umgang

„Aura Ally“ wurde entwickelt, um bei jeder Interaktion einen persönlichen und höflichen Umgangston zu wahren. Durch die Abfrage des Namens und die Wahlmöglichkeit zwischen Du- und Sie-Form ermöglichen wir eine personalisierte Kommunikation. Die Entscheidung gegen eine spezifische Pronomenabfrage und für eine allgemeine Ansprache vermeidet Diskriminierung und Exklusion. Zudem greift „Aura Ally“ aktiv bei Beleidigungen oder sonstigen toxischen Inhalten ein, indem es mittels Hate Speech Detection unangemessene Inhalte identifiziert und mit einer höflichen, deeskalierenden Rückmeldung reagiert.

Zusammenfassend ist „Aura Ally“ nicht nur eine technologische Innovation, sondern ein Schritt hin zu einer inklusiveren, gerechteren und respektvolleren digitalen Gesellschaft. Durch die Kombination aus fortschrittlicher Mensch-Maschine-

Interaktion, barrierefreiem Zugang und einem Klima des Respekts stellen wir uns an die Spitze der Bemühungen, eine digitale Zukunft zu schaffen, die niemanden ausschließt. Unsere Entwicklungen basieren auf den Prinzipien der Barrierefreiheit, Chancengleichheit und des respektvollen Miteinanders, wodurch „Aura Ally“ neue Standards in der Industrie für Conversational AI setzt und das transformative Potenzial der Technologie demonstriert, eine inklusive und faire Gesellschaft zu unterstützen.

Unsere Handlungsempfehlungen richten sich an verschiedene Stakeholder*innen darunter Entwickler*innen, Nutzer*innen und Entscheidungstragende, und betonen die Wichtigkeit einer fortwährenden Weiterentwicklung und Anpassung von Conversational-AI-Systemen, um eine gerechte und inklusive digitale Umgebung zu fördern. Wichtig ist es hierbei auch, Nutzer*innenfeedback einzuholen, um deren konkrete Bedürfnisse besser verstehen und berücksichtigen zu können. Durch die Implementierung von „Aura Ally“ demonstrieren wir, wie technologische Innovationen genutzt werden können, um Barrieren zu überwinden und eine digitale Zukunft zu gestalten, die allen Nutzer*innen gerecht wird.

Fazit

Die Entwicklung und Implementierung von „Aura Ally“ markiert einen wesentlichen Wendepunkt in der Bemühung, eine digital zugänglichere, inklusivere und gerechtere Welt der Conversational AI zu gestalten. Durch die Konzentration auf Schlüsselemente wie die Mensch-Maschine-Interaktion, Barrierefreiheit, Chancengleichheit und den respektvollen Umgang miteinander haben wir einen umfassenden Ansatz vorgestellt, der darauf abzielt, die digitale Erfahrung für alle Nutzer*innengruppen zu verbessern. Die Integration von Gebärdensprache und Alternativtexten, zusammen mit einer intuitiven Benutzer*innenoberfläche und personalisierbaren Interaktionsmöglichkeiten, stellt sicher, dass „Aura Ally“ den unterschiedlichen Bedürfnissen und Fähigkeiten der Nutzer*innen gerecht wird und dabei eine effiziente und benutzer*innenfreundliche Plattform ist.

Unsere Arbeit unterstreicht das transformative Potenzial von Technologie, um Barrieren abzubauen und eine inklusive digitale Zukunft zu fördern. Die Berücksichtigung aller Nutzer*innengruppen, mit besonderem Augenmerk auf Menschen mit Beeinträchtigungen, durch die Bereitstellung von Hilfestellungen, anpassbaren Optionen und diversen Kommunikationswegen verdeutlicht unser Engagement für eine umfassende Zugänglichkeit und Gleichberechtigung. Die proaktive Handhabung von Hate Speech und unangemessenen Inhalten durch fortgeschrittene Algorithmen trägt zur Schaffung einer sicheren und positiven Kommunikationsumgebung bei.

Schließlich setzt „Aura Ally“ neue Standards in der Conversational-AI-Industrie, indem es nicht nur technische Innovation vorantreibt, sondern auch ein tiefes Verständnis für ethische Verant-

wortung und soziale Inklusion zeigt. Unsere Handlungsempfehlungen rufen zu einer fortgesetzten Entwicklung und Anpassung auf, um sicherzustellen, dass die digitale Welt für jeden zugänglich und gerecht ist. „Aura Ally“ dient als Leitfaden für zukünftige Entwicklungen in der digitalen Welt, demonstriert die Bedeutung einer integrativen Technologie und unterstreicht unsere Verpflichtung, eine inklusivere, gerechtere und respektvolle digitale Gesellschaft zu unterstützen.



Quellen:

- Apuzzo, C. & Burrelli, G. (2022). Designing Accessible Chatbots for Deaf People. 11th Mediterranean Conference on Embedded Computing (MECO), IEEE, S. 1-4.
- BIK für Alle. (o. J.). Alternativtexte für Grafiken. Verfügbar unter: <https://bik-fuer-alle.de/alternativtexte-fuer-grafiken.html> (Abfrage: 06.02.2024).
- Deutscher Blinden- und Sehbehindertenverband e.V. (DBSV). (o. J.). Zahlen und Fakten. Verfügbar unter: <https://www.dbsv.org/zahlen-fakten.html> (Abfrage: 06.02.2024).
- Holzer, D. (2023). Barrierefreiheit: Bayern setzt auf KI-Gebärdensprachdolmetscher. Verfügbar unter: <https://www.br.de/nachrichten/netzwelt/barrierefrei-digital-bayern-setzt-auf-ki-gebaerdendolmetscher,TePBur0> (Abfrage: 06.02.2024).
- Jahan, M. S. & Oussalah, M. (2023). A systematic review of Hate Speech automatic detection using Natural Language Processing. *Neurocomputing*, 126232.
- Mandl, T. (2020). Die Erkennung unangemessener Inhalte im Internet: KI-Verfahren, Evaluierung und Herausforderungen. *Bibliotheksdienst*, 54(3-4), 214-226.
- Mohiyaddeen, M. & Siddiqi, S. (2021). Automatic hate speech detection: A literature review. Verfügbar unter SSRN 3887383.
- Pardasani, A., Sharma, A. K., Banerjee, S., Garg, V. & Singha Roy, D. (2018). Enhancing the ability to communicate by synthesizing American Sign Language using Image Recognition in a chatbot for differently Aabled. In: 2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO). IEEE, 529-532.
- Schmidt, A. & Wiegand, M. (2017). "A survey on hate speech detection using natural language processing." In Proceedings of the fifth international workshop on natural language processing for social media. 1-10.
- Sontheimer, L., Schäfer, J. & Mandl, T. (2022). Enabling Informational Autonomy through Explanation of Content Moderation: UI Design for Hate Speech Detection. *Mensch und Computer 2022-Workshopband*.
- Torres, C., Franklin, W. & Martins, L. (2019). Accessibility in chatbots: the state of the art in favor of users with visual impairment. In: *Advances in Usability, User Experience and Assistive Technology: Proceedings of the AHFE 2018 International Conferences on Usability & User Experience and Human Factors and Assistive Technology, Held on July 21-25, 2018, in Loews Sapphire Falls Resort at Universal Studios, Orlando, Florida, USA 9*. Springer, 623-635.



Editorial

EduPin

AI Allies

Grid
Harmony

PowerBalance

Beyond
TheBarrier

Dialogreihe

Anhang

Grid Harmony

Für eine faire Energiezukunft: Die smarte Lösung mit GridHarmony!



Team GridHarmony



Felix Buchholz
TU Hamburg



Jakob Frerichs
TU Hamburg



Jan Loschen
TU Hamburg



Jonas Zeiske
TU Hamburg



Use Case: Honda Research Institute EU

Kernbotschaften/Abstract



Der Übergang zu nachhaltigen Energiesystemen, insbesondere mit Fokus auf erneuerbare Energien wie Wind und Sonne, präsentiert vielschichtige Herausforderungen, einschließlich potenzieller Versorgungslücken.



Unsere App zur Förderung von spielerischem Stromsparen für Privathaushalte reagiert auf diese Herausforderungen, indem sie eine Vielzahl von Lösungsansätzen kombiniert, darunter spielerisches Lernen, personalisierte Empfehlungen und Belohnungen. Durch die Integration von KI und Nudging fördern wir langfristige Verhaltensänderungen und tragen zur Stabilisierung des Stromnetzes sowie zur Sicherheit der Stromversorgung bei.



Das eigens entwickelte Fairnesskonzept, das auf Transparenz, Bewusstseinsbildung, Solidarität und Motivation basiert, gewährleistet eine gerechte und inklusive Nutzung der App. Insgesamt bietet unsere App eine umfassende Lösung für die Herausforderungen der Energiewende und leistet einen Beitrag zur Erreichung der energiepolitischen Ziele in Deutschland.

Mitwirkende: Prof. Dr. Maximilian Kiener und Dr. Jonas Bozenhard

Herausforderung

Die Transition zu nachhaltigen Energiesystemen, insbesondere mit Fokus auf Wind und Sonne, stellt uns vor vielschichtige Herausforderungen. Während die verfügbare Leistung bis 2030 von 99 GW auf etwa 90 GW sinken wird, kann die Spitzenlast auf bis zu 120 GW ansteigen, was zu einer potenziellen Versorgungslücke von 30 GW führt (vgl. hier und im Folgenden McKinsey 2023). Die Lösungsansätze auf der Angebotsseite sind vielseitig und zielen beispielsweise auf erhöhte Stromimporte oder auf das Bauen neuer Gaskraftwerke ab. Doch auch Privathaushalte können beispielsweise durch Smart Meters aktiv dazu beitragen, die Spitzenlasten zu drücken. Dies erfordert ein adaptives Verbrauchsmangement (Honda Research Institute Europe 2023).

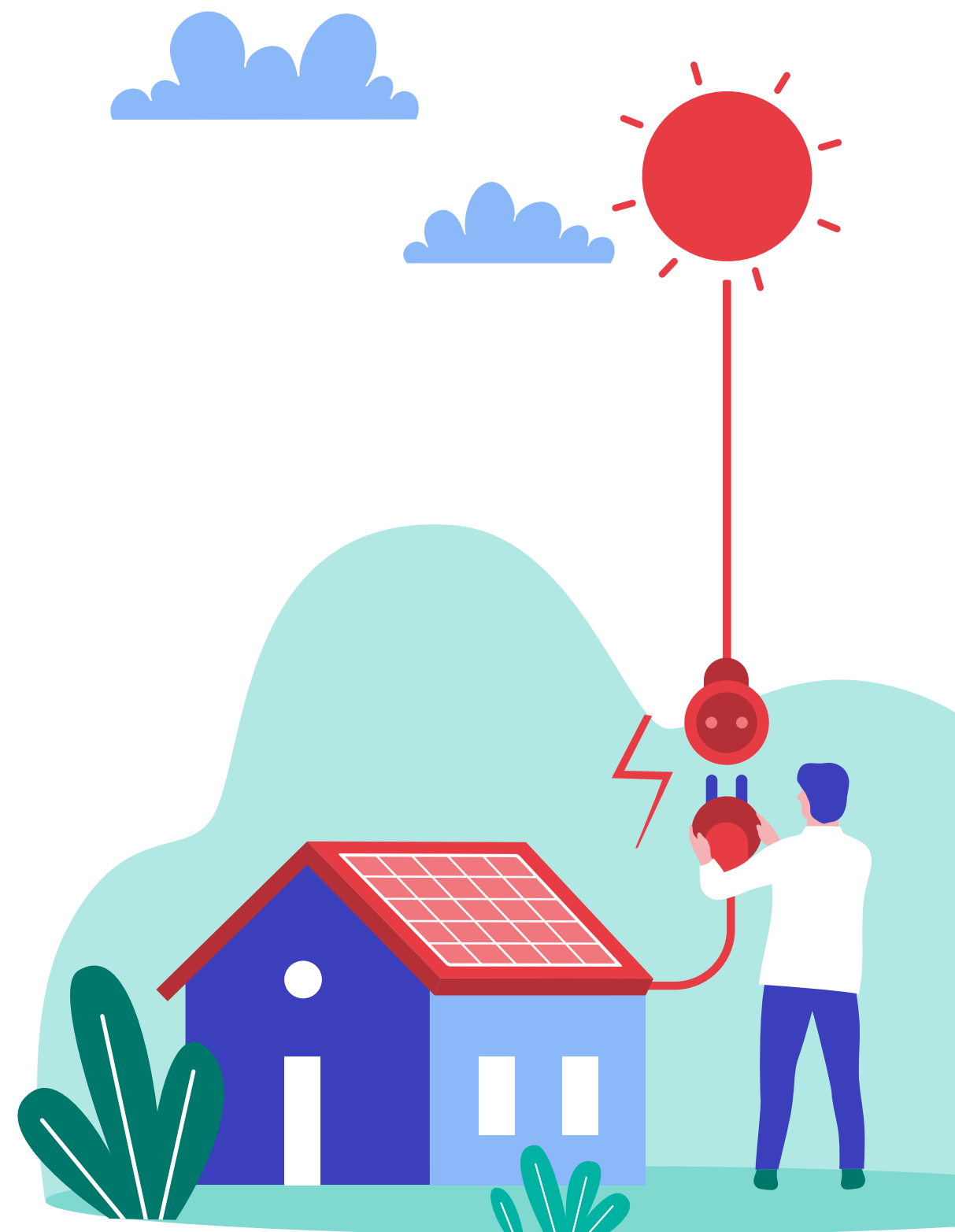
Konzentriert man sich weiter auf das Verbrauchsmangement, steht das Zusammenspiel aus künstlicher Intelligenz und Stromnutzer*innen besonders im Fokus. Wie bringt man Nutzer*innen dazu, einen Teil des Luxus nahezu unbegrenzter Energie aufzugeben? Was könnten attraktive Anreize für die Nutzer*innen sein, sich darauf einzulassen? Und welche Rolle spielen diesbezüglich Fairness, Inklusion und Datenschutz? Diese Fragen bergen weitere Herausforderungen, für die es einer Lösung bedarf.

Gerade die Integration von KI eröffnet neue digital-ethische Fragestellungen, wobei ein zentraler Aspekt die Möglichkeit für Nutzer*innen ist, ihre Entscheidungen weiterhin selbst zu treffen.

Dies geht über die rein technologische Ebene hinaus und offenbart eine psychologische Dimension. Die Akzeptanz des Systems ist maßgeblich von der wahrgenommenen Fairness abhängig, wodurch nicht nur die technische Effizienz, sondern auch die psychologische Dimension der Nutzer*innen-erfahrung entscheidend ist.

Die Integration von Smart Meters und verschiedenen Smart-Home-Technologien hingegen wirft die Frage auf, ob es gerecht ist, dass Nutzer*innen mit einem höheren Technologiestandard möglicherweise einen erweiterten Nutzen aus der KI ziehen könnten. Diese Herausforderung wurde während der Problemlösung erkannt und entsprechende Maßnahmen wurden ergriffen, um eine möglichst faire Lösung zu gewährleisten. Die Frage nach Chancengleichheit und einem inklusiven Zugang zu den Funktionen der App unterstreicht das Bestreben, Fairness in der Anwendung und Nutzung unabhängig von technologischen Unterschieden sicherzustellen.

Eine ganzheitliche Herangehensweise, die neben technischen Effizienz-Überlegungen und ethischen Erwägungen auch die psychologische Dimension und die gezielte (positive) Beeinflussung von Verhaltensweisen umfasst, ist entscheidend für den Erfolg nachhaltiger Energiesysteme (Honda Research Institute 2023).



Analyse und Handlungsempfehlungen/Lösung

Für uns ist die Entwicklung einer App zur Förderung spielerischen Stromsparverhaltens für private Haushalte eine direkte Reaktion auf die zunehmende Herausforderung der Gewährleistung von Stromversorgungssicherheit in Deutschland. Angesichts der Prognosen einer drohenden Stromlücke und der verschlechterten Indikatoren für die Energiewende ist es dringend erforderlich, innovative Lösungen zu implementieren, die sowohl an der Angebots- als auch der Nachfrageseite des Stromnetzes ansetzen (McKinsey 2023).

Unsere App adressiert dieses Problem, indem sie eine Vielzahl von Lösungsansätzen und Konzepten kombiniert. Zunächst einmal bietet sie einen spielerischen Ansatz, um das Bewusstsein für den Stromverbrauch zu schärfen und die Nutzer*innen zu motivieren, ihren Energieverbrauch zu optimieren. Die Integration einer KI ermöglicht es der App, personalisierte Handlungsempfehlungen basierend auf Wetterdaten, Erfahrungswerten und dynamischen Strompreisen zu geben (vgl. Abb. 6). Dies unterstützt nicht nur die individuelle Energieoptimierung, sondern trägt auch zur Stabilisierung des Gesamtsystems bei. Zudem werden Nutzer*innen durch subtiles Nudging innerhalb der App in die „richtige“ Richtung gelenkt, was jedoch keinesfalls mit Manipulation zu vergleichen ist.

Nudging hat sich als wichtiges Werkzeug für die App erwiesen, um langfristige Verhaltensänderungen in Bezug auf Stromsparen zu fördern. Durch

gezielte Interventionen, die auf Erkenntnissen der Verhaltenswissenschaft basieren, motivieren wir die Nutzer*innen dazu, energieeffizientes Verhalten zu praktizieren. Studien wie die von Allcott und Rogers (2014) sowie Ferraro und Price (2013) belegen genau das. Sie zeigen, dass Nudges vergleichbare oder gar stärkere Effekte haben können als finanzielle Anreize. Diese Erkenntnisse führten zur Gründung des Behavioural Insights Team durch die britische Regierung. Diese Gruppe implementiert Nudging in verschiedene öffentliche Dienstleistungen, wie zum Beispiel in Schreiben der britischen Finanzämter. Anknüpfend daran ist hier wichtig das eigens entwickelte Fairnesskonzept von GridHarmony, welches kurz „GRID-Modell“ genannt wird, zu erwähnen:

Mit „**Guarentee transparency**“ garantieren wir vollständige Transparenz gegenüber unserer Nutzer*innen. Diese erreichen wir durch viele kleine Infokacheln innerhalb der App, welche es ihnen ermöglichen, alles über die Funktions- und Arbeitsweise der App zu erfahren. Besonders ist uns wichtig, dass Informationen über die Arbeitsweise der KI, das Vorhandensein und die Sinnhaftigkeit des Nudging geteilt werden.

„**Raise awareness**“ geht auf den edukativen Mehrwert unserer App ein. Denn GridHarmony soll nicht nur dabei helfen, Strom zu sparen, sondern auch mithilfe des „Education-Tab“ die Möglichkeit bieten, anhand von Mini-Games mehr



Abb. 6 – Handlungsempfehlung



Abb. 7 – Education-Tab

über den Strommarkt, erneuerbare Energien und ein nachhaltiges Leben zu lernen (vgl. Abb. 7). Belohnt werden soll dies mit kleinen Gutscheinen für beispielsweise ÖPNV-Dienstleister und andere nachhaltige Unternehmen.

Des Weiteren steht der „**Community-Tab**“ von GridHarmony für den Punkt „**Incentivise solidarity**“ im „GRID-Modell“. Dort können Freunde hinzugefügt und sich mit ihnen in Rankings verglichen werden. Es kann sich aber auch ganz anonym mit Menschen aus der jeweiligen Region verglichen werden.

Zu guter Letzt steht „**Democratise motivation**“ dafür, dass GridHarmony für jeden eine Motivation bieten soll, Strom zu sparen – egal ob gesellschaftliche Unter-, Mittel- oder Oberschicht. Die Grundmechaniken der App sollen für jede*n einen Anreiz bieten, sich mit dem Thema Stromsparen und erneuerbaren Energien auseinanderzusetzen.

Zusammen ergeben diese Punkte das „**GRID-Modell**“ von GridHarmony.

In Bezug auf die gesamtgesellschaftliche Relevanz unserer Lösung ist anzumerken, dass die App nicht nur die individuelle Verbrauchssituation der Nutzer*innen verbessert, sondern auch einen Beitrag zur Reduzierung der Spitzenlast und zur Stabilisierung des Stromnetzes leistet. Dies ist entscheidend für die Sicherheit der Stromversorgung und die Vermeidung von Engpässen.

Die Analyse von McKinsey 2023 verdeutlicht, dass neben dem Ausbau erneuerbarer Energien auch Maßnahmen auf der Nachfrageseite erforderlich sind, um die Versorgungslücken zu schließen (vgl. Abb. 8 (McKinsey 2023)). Unsere App bietet genau das: Durch die Nutzung von Smart-Meter-Technologie und die Steuerung von Elektrogeräten wie Waschmaschinen in Zeiten günstiger Stromtarife ermöglicht sie eine flexible Anpassung des Stromverbrauchs an die verfügbare Energie.

Insgesamt bietet unsere App eine umfassende Lösung für die Herausforderungen der Energiewende und der Stromversorgungssicherheit in Deutschland. Durch die Kombination von spielerischem Lernen, personalisierten Empfehlungen und Belohnungen schaffen wir Anreize für eine nachhaltige Energiepraxis und leisten einen Beitrag zur Erreichung der energiepolitischen Ziele.

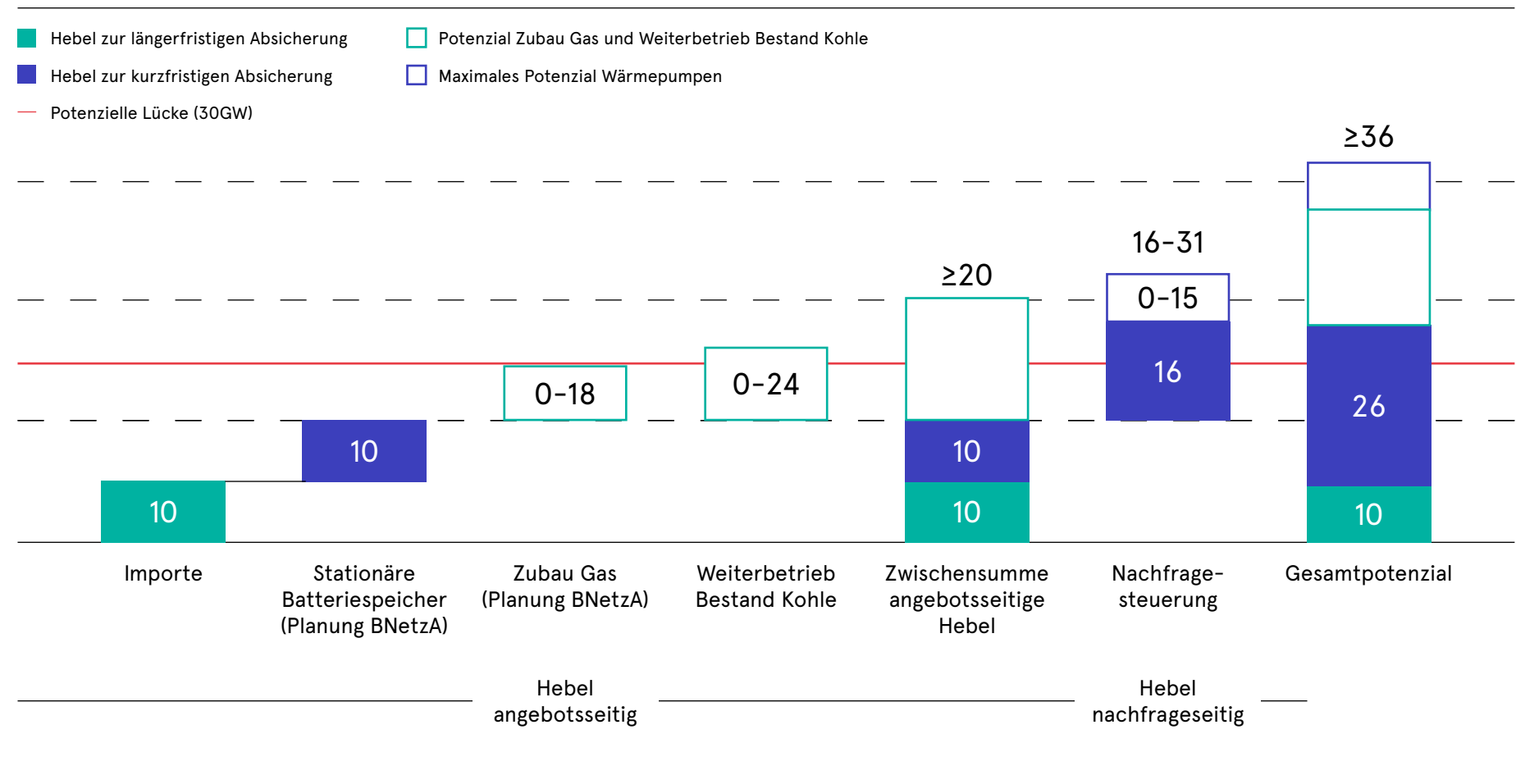


Abb. 8 – Angebots- und nachfrageseitige Hebel (McKinsey 2023)

Fazit

Die Umstellung auf nachhaltige Energiesysteme stellt uns vor eine komplexe Herausforderung, insbesondere im Hinblick auf den Rückgang der verfügbaren Leistung und das potenzielle Auftreten von Versorgungslücken. Lösungsansätze wie erhöhte Stromimporte und der Ausbau von Gaskraftwerken werden diskutiert, doch auch Verbraucher*innen können durch Smart Meters aktiv dazu beitragen, die Spitzenlasten zu drücken. Hier setzt unsere App an, indem sie eine spielerische Herangehensweise mit KI-Unterstützung kombiniert, um das Bewusstsein für den Stromverbrauch zu schärfen und Energieeinsparungen zu fördern.

Die Integration von KI und Smart-Meter-Technologie wirft jedoch ethische Fragen auf, insbesondere im Hinblick auf Fairness, Inklusion und Datenschutz. Unsere Lösung adressiert diese Bedenken, indem sie Transparenz, Bildung und soziale Anreize fördert. Das „GRID-Modell“ von GridHarmony garantiert Transparenz, erhöht das Bewusstsein durch Bildungsinhalte und schafft eine Gemeinschaft von Nutzer*innen, die sich

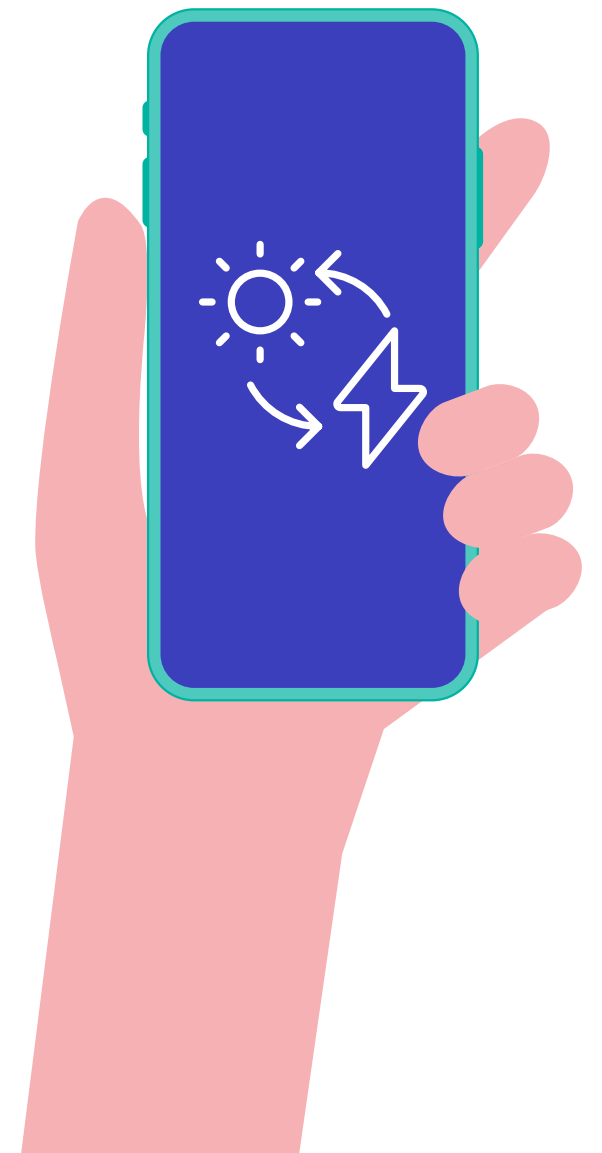
gegenseitig motivieren. Durch diese ganzheitliche Herangehensweise soll eine demokratische Motivation für Stromsparverhalten geschaffen werden.

Die Effektivität von Nudging als Verhaltensänderungsinstrument wird durch Studien unterstützt, die zeigen, dass Nudges vergleichbare oder sogar stärkere Effekte haben können als finanzielle Anreize. Unsere App nutzt gezieltes Nudging, um langfristige Verhaltensänderungen in Bezug auf Stromsparen zu fördern. Indem sie auf Erkenntnissen der Verhaltenswissenschaft basiert und personalisierte Empfehlungen bietet, trägt sie zur individuellen Energieoptimierung und zur Stabilisierung des Gesamtsystems bei.

Insgesamt bietet unsere App eine umfassende Lösung für die Herausforderungen der Energiewende und der Stromversorgungssicherheit. Durch die Kombination von spielerischem Lernen, personalisierten Empfehlungen und sozialen Anreizen leisten wir einen Beitrag zur Erreichung der energiepolitischen Ziele.

„Es geht bei sowas auch immer ums Schaffen von Anreizen: Wer die App herunterlädt, um Geld zu sparen, profitiert bei uns von Rabatten und anderen Boni. Wer sich zum Beispiel gerne mit Freund*innen misst, kann das dank unseres Gamification-Ansatzes über ein Ranking tun.“

Felix Buchholz, GridHarmony



Quellen:

Allcott, H., & Rogers, T. (2014). How Long Do Treatment Effects Last? Persistence of Energy Saving Behaviors Over Time. *Review of Economics and Statistics*, 96(5), 917–934.
 Energiewendeindex von McKinsey: Versorgungssicherheit unter Spannung (2023): McKinsey & Company (auch online unter: <https://www.mckinsey.de/news/presse/2023-03-06-energiewende-index> [Abfrage: 06.02.2024]).
 Ferraro, P. J., & Price, M. K. (2013). Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment. *Review of Economics and Statistics*, 95(1), 64–73.
 Honda Research Institute: Case Study 2023 (auch online unter: https://assets.foleon.com/eu-central-1/de-uploads-7e3kk3/19783/dfc2023_hri.1a57457a1252.pdf [Abfrage: 06.02.2024]).



Editorial

EduPin

AI Allies

GridHarmony

Power
Balance

Beyond
TheBarrier

Dialogreihe

Anhang

Power Balance

PowerBalance – DER Gamechanger für die Energiewende!



Team PowerBalance



Mursal Mohammadi
TU München



Felix Schnippering
TU München



Charlotte Vorbauer
TU München



Kernbotschaften/Abstract



Die Herausforderung einer nachhaltigen Energieversorgung, die Unvorhersehbarkeit erneuerbarer Energiequellen und die Notwendigkeit einer gerechten Energieverteilung bilden den Ausgangspunkt für die Entwicklung von PowerBalance. Dieses innovative System, basierend auf künstlicher Intelligenz und smarten Stromzählern, zielt darauf ab, die Energieverteilung effizienter, gerechter und nachhaltiger zu gestalten.



Durch die Echtzeit-Datenanalyse, die Berechnung eines fairen Förderfaktors und ein Bonus-Malus-System werden individuelle Verbrauchsmuster optimiert und an die Verfügbarkeit von Energie angepasst. PowerBalance steht nicht nur für ein fortschrittliches Energiemanagement, sondern verkörpert auch die Prinzipien der Verteilungsgerechtigkeit, indem es jedem Haushalt ermöglicht, seinen idealen Verbrauchsbereich zu finden, und Anreize setzt, diesen auch nutzen.



Durch die PowerBalance profitieren Verbraucher*innen und Stromversorger primär finanziell. Es entsteht ein neuer Stellhebel zur Stabilisierung des Strommarkts, was zu signifikanten Kostenreduktionen führt. Der Anteil benötigter fossiler Energieträger sinkt, was geringere CO₂-Emissionen und ein früheres Erreichen der Klimaziele zur Folge hat.

Use Case: Honda Research Institute EU

Herausforderung

„Wir können der Sonne nicht sagen, wann sie scheinen soll, und dem Wind nicht, wann er zu wehen hat.“

Wir stehen vor dem grundlegenden Dilemma, dass unsere aktuellen Methoden zur Stabilisierung der Energieversorgung – wie der Rückgriff auf Gas- und Kohlestrom, der Ausbau kostspieliger Batteriespeicher und Stromimporte – weder nachhaltig sind, noch unseren Zielen einer umweltfreundlichen Energiepolitik entsprechen. Zudem sind bei der Umstellung auf erneuerbare Energien Schwankungen unvermeidlich. Vor diesem Hintergrund der natürlichen Unbeständigkeit von Wind- und Solarenergie gilt es, eine gerechte und effiziente Verteilung von Energie zu gewährleisten.

Aktuell müssen Stromversorger Engpässe marktseitig kompensieren. Die hohen Kosten für diese Versorgungssicherheit werden von Stromversorgern durch teure Pauschalbeträge an die Endverbraucher weitergegeben. Smart Meters spielen eine Schlüsselrolle für flexiblere Strompreise und bieten einen Lösungsansatz, um auf diese Herausforderungen zu reagieren.

Derzeit existiert ein überwiegend marktbasierendes Energiemanagementsystem, das wenig Raum für individuelle oder kollektive Entscheidungsmöglichkeiten lässt. Verbraucher*innen haben kaum Einfluss darauf, ihre Energieverbrauchsmuster anzupassen und davon zu profitieren, Energie zu nutzen,

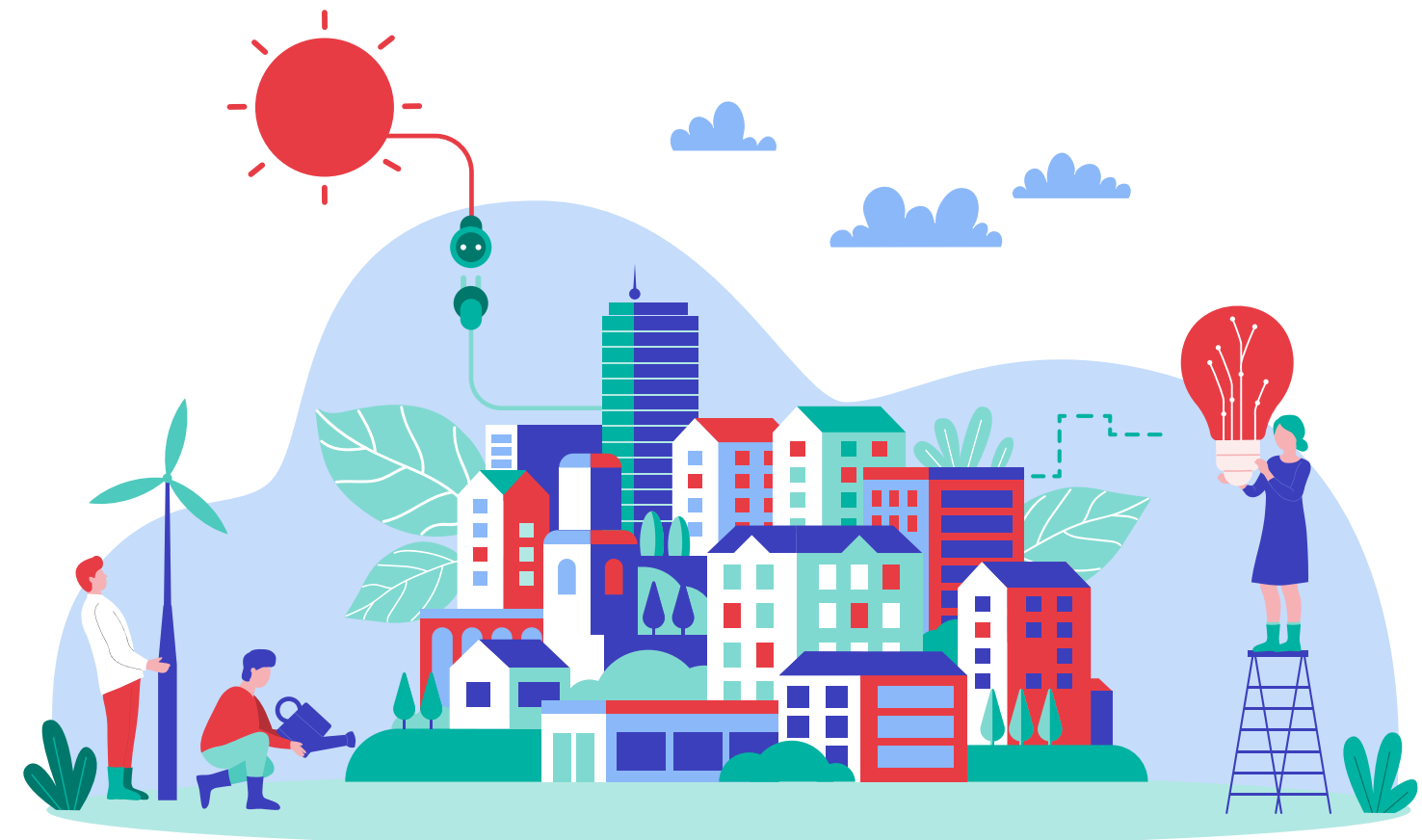
wenn die aktuelle Verfügbarkeit hoch ist. Stattdessen sind sie mit einem statischen Preisgefüge konfrontiert, das keine Anreize für Energieeinsparungen oder eine Anpassung des Verbrauchsverhaltens bietet. Diese Situation führt zu einem ineffizienten Energieverbrauch und einer Verteilung der Energiekosten ohne Rücksicht auf individuelle oder soziale Umstände.

Stellen Sie sich nun vor, Sie wüssten nicht, wie Ihre persönlichen Lebensumstände aussehen – wie viele Kinder Sie haben, wie viel Sie verdienen, ob Sie in der Stadt oder auf dem Land leben. Sie befänden sich hinter John Rawls' Schleier des Nichtwissens: Wie würden Sie unter diesen Umständen die Verteilung von Energie entscheiden? Gemäß Rawls' The-

orie der Gerechtigkeit (Rawls 1999) würden Sie wahrscheinlich ein System bevorzugen, das soziale Faktoren berücksichtigt und benachteiligten Gruppen einen erleichterten Zugang zur Energie ermöglicht, ohne dass dieser allein vom Preis abhängig ist.

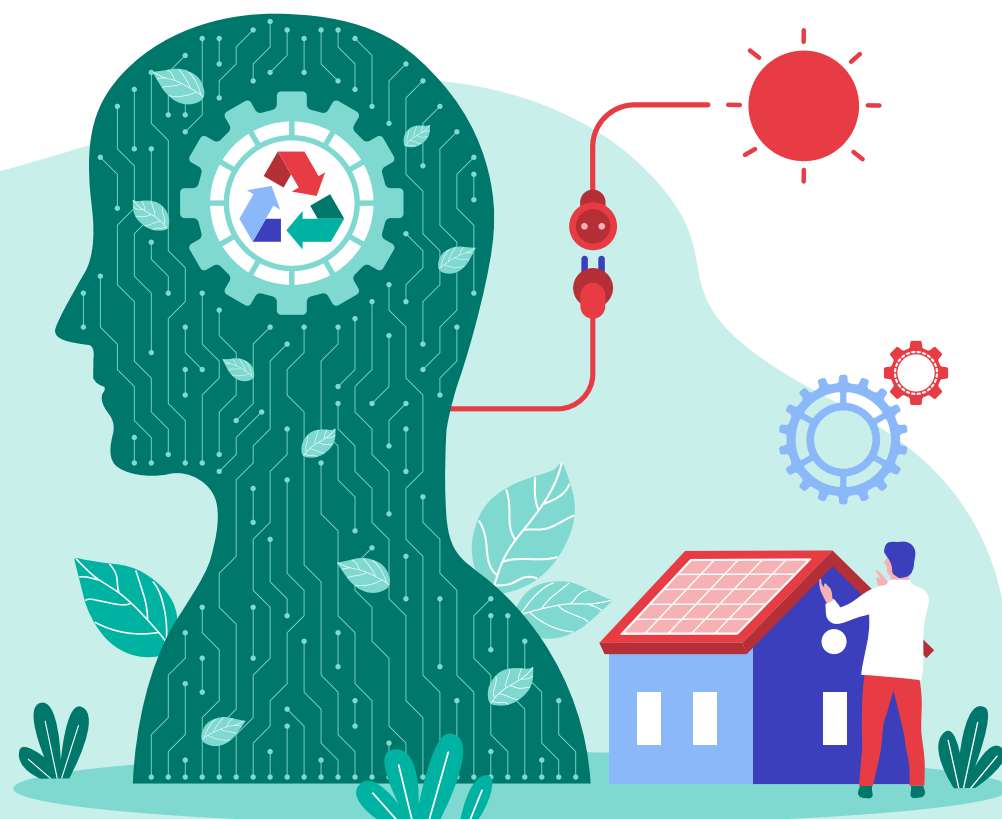
Diese Überlegungen führen uns zu unserer Vision von PowerBalance: einem System, das nicht nur effizient ist, sondern auch durch seine Fairness überzeugt.

PowerBalance zielt darauf ab, eine Energieverteilung zu schaffen, die über bloße Marktdynamiken hinausgeht und sicherstellt, dass jede*r – unabhängig von der eigenen sozioökonomischen Situation – gerechten Zugang zu Energie hat.



Analyse

Die zentrale Herausforderung besteht darin, ein System zu entwickeln, das gerecht ist und die Bedürfnisse aller Nutzer*innen berücksichtigt. Die Frage der Fairness ist besonders heikel, da es sich bei der Energieversorgung um die Verteilung eines Grundrechts handelt (BverfGE 125, 175). Hier setzt unsere Idee an: PowerBalance bietet eine innovative Lösung für das Problem der schwankenden Energieversorgung, indem sie als KI-basiertes Steuerungssystem Anreize für den individuellen idealen Stromverbrauch setzt. Durch die Anwendung von PowerBalance kann eine effektive Angleichung der Nachfrage an das Stromangebot erreicht werden.



Echtzeit-Datenanalyse mithilfe von smarten Stromzählern und KI

Der erste Schritt umfasst die Sammlung und Auswertung von Stromverbrauchsdaten in Echtzeit mittels Smart Meters. Das bereits verabschiedete Smart-Meter-Gesetz sieht einen flächendeckenden Rollout von Smart Meters in Deutschland bis 2032 vor (BMWK 2023). Diese Geräte liefern präzise Verbrauchsinformationen, die durch die fortschreitende Einführung gemäß dem Smart-Meter-Gesetz zunehmend verfügbar werden. PowerBalance nutzt KI-Technologien, um aus diesen Daten Verbrauchsmuster zu identifizieren und Prognosen zum zukünftigen Energiebedarf zu erstellen. Bestehende Datenquellen zur deutschlandweiten Prognose von Verfügbarkeit und Bedarf (vgl. www.smarde.de; Bundesnetzagentur) werden ebenfalls eingebunden.

Berechnung des fairen Förderfaktors (FFF) für ein gerechtes System

Die Implementierung des fairen Förderfaktors (FFF) stellt einen zentralen Pfeiler des PowerBalance-Systems dar, um ein ausgewogenes und gerechtes Energiemanagement zu ermöglichen. Dieser Faktor wird spezifisch berechnet, um eine gleichförmige und gerechte Energieverteilung sicherzustellen. Ziel ist es, den optimalen Verbrauchsbereich – die sogenannte PowerBalance-Sweetzone – für jeden Haushalt individuell fest-

zulegen. Der FFF zieht dabei eine Vielzahl von sozialen Faktoren wie Haushaltsgröße, Einkommen und andere relevante Aspekte in Betracht. Diese präzise und maßgeschneiderte Berechnungsweise erlaubt es, jedem Haushalt einen angepassten idealen Verbrauchsbereich zuzuordnen, der gerecht und zugleich realitätsnah ist.

Unser benutzer*innenfreundlicher FFF-Rechner, der analog zu einem Brutto-Netto-Rechner funktioniert, ermöglicht es allen Nutzer*innen, auf transparente Weise mit den verschiedenen Faktoren zu experimentieren. Dadurch wird nicht nur Transparenz gefördert, sondern es werden auch individuelle Anpassungsmöglichkeiten und ein tieferes Verständnis für den Einfluss sozialer Faktoren auf den Energieverbrauch ermöglicht.

Diese Praxis spiegelt die Prinzipien der Verteilungsgerechtigkeit nach John Rawls wider, indem sie darauf abzielt, soziale und ökonomische Ungleichheiten zum Vorteil der am wenigsten Begünstigten zu gestalten (Rawls 1999), und stellt somit eine ethisch fundierte, sozial verantwortliche Berechnungsmethode dar.

Prognostizierung der Sweetzone

PowerBalance nutzt KI, um nach der Berechnung des fairen Förderfaktors (FFF) für jede*n Nutzer*in eine individuelle Sweetzone zu definieren. Diese Vorgehensweise ermöglicht es, den optimalen

Stromverbrauchsbereich unter Berücksichtigung des FFF und der verfügbaren Energiekapazitäten festzulegen. Anhand dieser Analyse generiert das System präzise Handlungsempfehlungen – beispielsweise den optimalen Zeitpunkt für das Waschen der Wäsche –, um den Energieverbrauch zu optimieren und das Netzwerk effizient zu nutzen (vgl. Abb. 9). Solche gezielten Vorschläge unterstützen Nutzer*innen dabei, Energie sparsamer einzusetzen, und tragen zur Stabilität des Stromnetzes bei. Darüber hinaus sollen Smart Devices angebunden werden können. So kann z. B. die Wallbox das Laden des Elektroautos zeitlich optimal automatisieren.

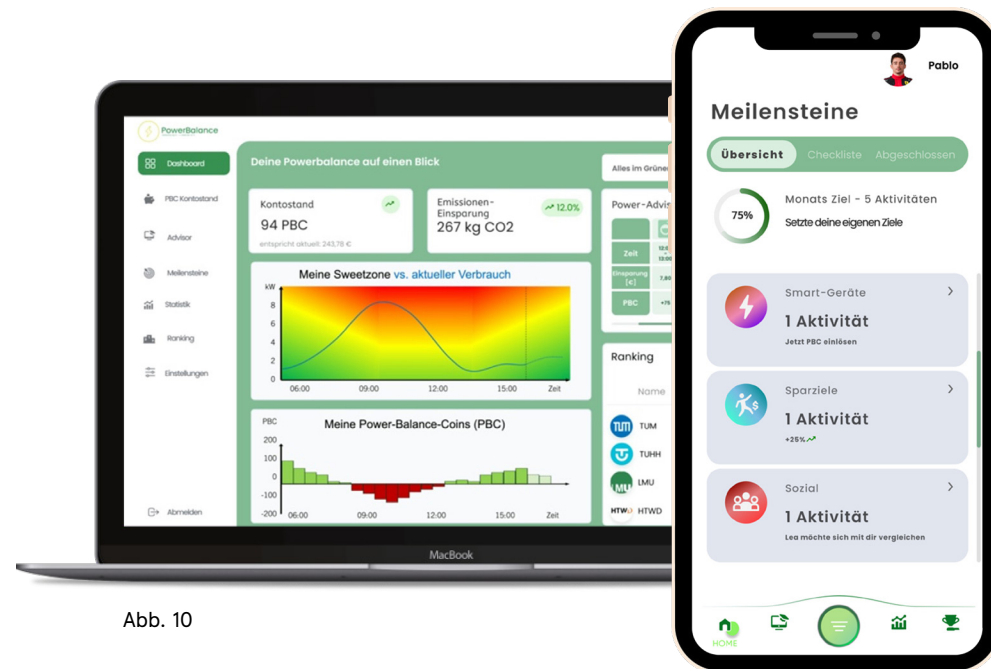


Abb. 10

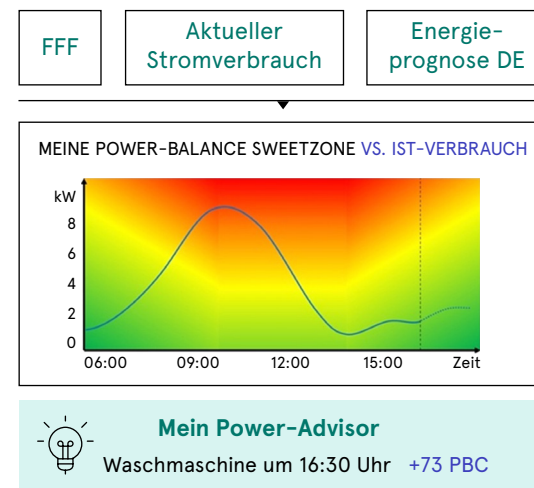


Abb. 9

PowerBalance Coins

Das Bonus-Malus-System der PowerBalance wird durch PowerBalance Coins realisiert, die als Anreize dienen, den Stromverbrauch an die Sweetzone anzupassen. Nutzer*innen erhalten Coins für energieeffizientes Verhalten oder geben welche ab bei übermäßigem Verbrauch. Der Wert der PBC wird dynamisch analog zum Strommarkt berechnet. Diese Coins sind vielseitig einsetzbar: Sie können gespendet werden oder zur Verrechnung mit der Stromrechnung genutzt werden. Ferner ermöglichen sie Investitionen in intelligente Geräte zur weiteren Verbrauchsoptimierung. Eine spielerische Komponente erlaubt den Vergleich des Verbrauchs mit anderen, fördert die Teilnahme und motiviert zu bewusstem Energieverbrauch. In unserer Desktop-Applikation sehen die Nutzer*innen auf einen Blick ihre Coins, die prognostizierte Sweetzone, Handlungsempfehlungen, aber auch weitere Anreize wie Rankings oder Einsparungstracker (vgl. Abb. 10).

Handlungsempfehlungen/Lösung

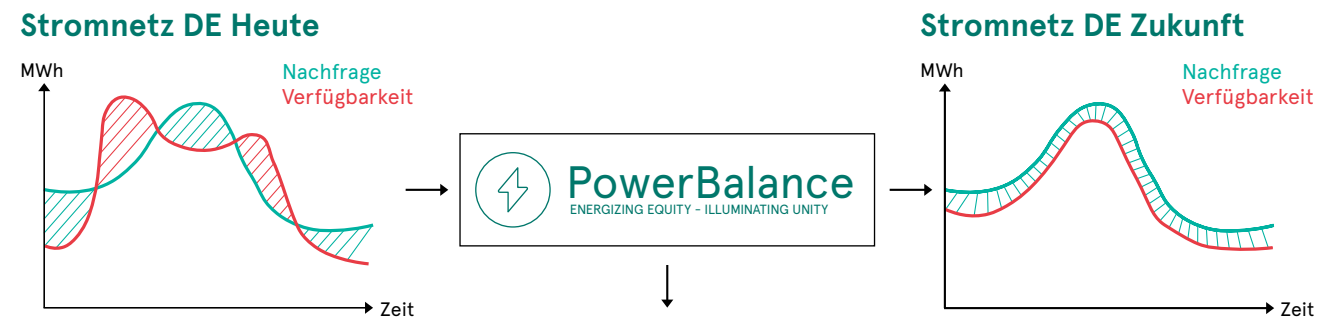
Die PowerBalance-Lösung bietet auf verschiedenen Ebenen Mehrwerte, die sowohl einzelne Akteur*innen als auch die Gesamtgesellschaft begünstigen. Energieversorger profitieren erheblich, da in Zeiten geringer Stromproduktion der Kauf teuren Zusatzstroms entfällt, was eine effizientere und kostengünstigere Betriebsführung ermöglicht. Da Energieversorger besonders in Knappheitsszenarien Vorteile sehen, ist der Plan, PowerBalance über diese auszurollen, wobei sie auch direkten Zugang zu den Endverbraucher*innen haben.

Verbraucher*innen ziehen Nutzen aus der Möglichkeit, Strom zu günstigeren Echtzeitpreisen zu beziehen und dabei PowerBalance Coins (PBC) zu sparen. Dies fördert einen bewussteren Energieverbrauch. Die Umwelt profitiert ebenfalls signifikant: Durch die optimierte Nutzung von Energiequellen werden CO_2 -Emissionen reduziert, da der Bedarf an fossilen Energieträgern sinkt. Dies trägt zur Verringerung des ökologischen Fußabdrucks bei.

Gesamtgesellschaftlich fördert PowerBalance Fairness und Gerechtigkeit, indem es allen Bevölkerungsschichten den Zugang zu einer stabilen, kosteneffizienten und vor allem auf ihre jeweilige soziale Situation abgestimmten Energieversorgung ermöglicht. Zudem stärkt das System das Gemeinschaftsgefühl durch Rankings und die Möglichkeit, PBC zu spenden, was den sozialen Zusammenhalt und die kollektive Verantwortung für eine nachhaltigere Energiezukunft fördert.

Darüber hinaus unterstützt PowerBalance den Staat dabei, Klimaziele schneller zu erreichen, indem es zu einer nachhaltigen Energiepolitik beiträgt. Das höhere Gemeinschaftsgefühl und die Förderung nachhaltiger Energieverbrauchspraktiken wirken sich positiv auf den gesamten Staat und die Gesellschaft aus.

Insgesamt verbindet PowerBalance wirtschaftliche, ökologische und soziale Nachhaltigkeit und schafft so einen umfassenden gesellschaftlichen Mehrwert (vgl. Abb. 11).



Durch die PowerBalance entsteht ein Mehrwert für ...



Abb. 11

Fazit

PowerBalance markiert einen innovativen Ansatz in der Energieversorgung, der technische Fortschritte mit den Prinzipien sozialer Gerechtigkeit vereint. Durch die Nutzung von KI und intelligenten Stromzählern passt es den Energieverbrauch dynamisch an die Verfügbarkeit erneuerbarer Energien an, wodurch Energieversorger erheblich von reduzierten Betriebskosten profitieren und dies an Verbraucher*innen aktiv weitergeben. Dies führt nicht nur zu einer effizienteren Nutzung von Ressourcen, sondern unterstützt auch eine Reduktion der CO₂-Emissionen und fördert die Erreichung von Klimazielen.

Auf gesellschaftlicher Ebene stärkt PowerBalance das Gemeinschaftsgefühl und fördert Fairness durch den Zugang zu gerechter und kosteneffizienter Energie für alle Bevölkerungsschichten. Das Bonus-Malus-System mit PowerBalance Coins motiviert zu einem verantwortungsvollen Energieverbrauch und ermöglicht eine spielerische Teilnahme am Energiemanagement, was den sozialen Zusammenhalt und das Bewusstsein für Energieeffizienz verstärkt.

Zusammenfassend verbindet PowerBalance wirtschaftliche, ökologische und soziale Nachhaltigkeit zu einem ganzheitlichen System, das weit über traditionelle Energielösungen hinausgeht. Es bietet eine zukunftsweisende Antwort auf die drängenden Energie- und Umweltfragen unserer Zeit und legt den Grundstein für eine gerechtere, nachhaltigere Energiezukunft.

Quellen:

Bundesministerium Für Wirtschaft Und Klimaschutz/BMWK (2023): Smart Meter-Gesetz final beschlossen: Flächendeckender Einsatz intelligenter Stromzähler kommt, [online] <https://www.bmwk.de/Redaktion/DE/Pressemitteilungen/2023/05/20230512-smart-meter-gesetz-final-beschlossen.html>. [abgerufen am 06.02.2024].
 Bundesregierung (2023): Intelligente Strommessung für die Energiewende, Die Bundesregierung informiert, [online] <https://www.bundesregierung.de/breg-de/schwerpunkte/klimaschutz/digitale-energiewende-2157184#:~:text=Ab%202025%20ist%20der%20Einbau,mit%20Smart%20Metern%20ausgestattet%20sein>. [abgerufen am 01.02.2024].
 BverfGE 125, 175, Beschluss des Ersten Senats vom 23. Juli 2014 – 1 BvL 10/12 –, Rn. 1–149, [online] https://www.bverfg.de/e/1s20140723_1bvl001012.html [abgerufen am 06.02.2024].
 John Rawls (1999): A Theory of Justice: Revised Edition, Cambridge, MA: Harvard University Press.



Editorial

EduPin

AI Allies

GridHarmony

PowerBalance

Beyond
TheBarrier

Dialogreihe

Anhang

Beyond TheBarrier

Wie ermöglicht man eine vorurteils- und barrierefreie Interaktion mit Conversational AI?



Team BeyondTheBarrier



Sidar Dalli
HTW Berlin



Fritz Arno Heitmann
HTW Berlin



Vasileios Perchanidis
HTW Berlin



Use Case: O₂ Telefónica

Kernbotschaften/Abstract



Immer mehr Menschen werden in ihrem alltäglichen Leben mit künstlicher Intelligenz konfrontiert. Der technologische Fortschritt macht vor keinem Lebensbereich halt – ganz im Gegenteil: Um kompetitiv zu bleiben und nicht abgehängt zu werden, sind Unternehmen dazu gezwungen, diese Technologie so schnell wie möglich in irgendeiner Form in ihre Dienstleistungen zu integrieren. So wird auch im Kund*innendienst „Conversational AI“ (CAI) immer verbreiteter.



Der schnelle technologische Fortschritt von künstlicher Intelligenz darf unter keinen Umständen gehemmt werden. Dies könnte die Position einer Nation auf der Weltbühne in vielerlei Hinsicht in Gefahr bringen, insbesondere wirtschaftlich, aber auch auch sicherheitstechnisch.



Gleichzeitig muss man erreichen, dass mit Fortschritt dieser essenziellen Technologie die sozialen Belange der Menschen nicht auf der Strecke bleiben. Es muss von Anfang an ein Handlungsrahmen gebildet werden, in dem sich die Technologie im Einklang mit den sozialen Bedürfnissen aller Menschen entwickeln kann.

Herausforderung

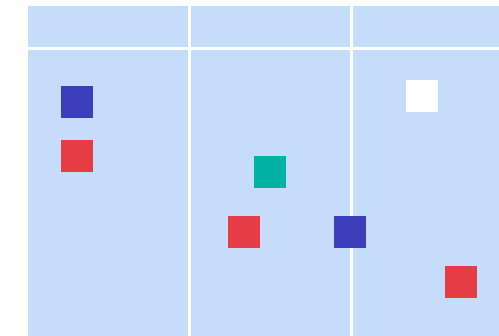
Technologien, die darauf abzielen, menschliche Interaktionen zu simulieren und zu unterstützen, stehen vor der Herausforderung, gerecht, zugänglich und frei von Vorurteilen zu sein.

Trotz des Potenzials, Dienstleistungen und Informationen breiter verfügbar zu machen, können digitale Barrieren Menschen mit eingeschränkten technischen, sprachlichen oder kognitiven Fähigkeiten und solche ohne Zugang zu den neuesten Technologien ausschließen. Ältere Menschen, Personen mit Behinderungen und Menschen in ländlichen oder wirtschaftlich und sozial benachteiligten Regionen sind besonders anfällig für solche Ausschlüsse. Die Entwicklung von Conversational AI, die universell zugänglich ist und adaptive Schnittstellen für verschiedene Benutzer*innenbedürfnisse bietet, ist eine fortlaufende ethische Herausforderung.

Ein weiterer kritischer Bereich sind die in Daten und Algorithmen eingebetteten Voreingenommenheiten. Conversational-AI-Systeme lernen aus riesigen Datenmengen, die historische und kulturelle Vorurteile widerspiegeln können. Wenn diese Systeme beispielsweise Trainingsdaten verwenden, die nicht vielfältig genug sind oder systematische Verzerrungen aufweisen, können ihre Antworten und Entscheidungen diese Voreingenommenheiten widerspiegeln. Dies kann zu diskriminierenden Praktiken führen, etwa wenn be-

stimmte Dialekte, Sprachmuster oder demografische Merkmale benachteiligt werden.

Darüber hinaus stellt sich die Frage der Verantwortlichkeit und Transparenz. Nutzer*innen von Conversational AI sollten verstehen können, wie Entscheidungen getroffen werden, insbesondere in kritischen Anwendungen wie im Gesundheitswesen, in der Rechtsprechung oder im Finanzwesen. Die Schaffung transparenter Systeme, die ihre Entscheidungsfindungsprozesse offenlegen, ist essenziell, um Vertrauen aufzubauen und Nutzer*innen zu ermöglichen, die Empfehlungen und Entscheidungen der AI nachzuvollziehen.





Analyse

Um gegen Voreingenommenheiten und Barrieren vorzugehen, muss man sie zunächst fest definieren und quantifizieren. Dieser Schritt ist unumgänglich, da man so ein kollektives Verständnis für die Natur und Ursachen dieser Probleme erhält. Somit wird es ermöglicht, Maßnahmen und Strategien auf einer breiten, gemeinsamen, aber insbesondere transparenten Grundlage zu entwickeln.

Das Definieren und Quantifizieren von Barrieren ist im Vergleich zu Voreingenommenheiten oft einfacher, was in der Natur dieser beiden Phänomene liegt. Barrieren sind in der Regel externe Hindernisse oder Einschränkungen, die relativ direkt beobachtet, gemessen und somit auch angegangen werden können. Sie manifestieren sich oft in Form von physischen Hürden, gesetzlichen oder institutionellen Regelungen und Zugangsbeschränkungen, die greifbar sind und objektiv identifiziert werden. Diskussionen über Barrieren sind häufig weniger emotional aufgeladen und könnten zumeist durch direkte Maßnahmen wie den Einsatz neuer Technologie oder Gesetzesänderungen relativ unkompliziert beseitigt werden.

Hingegen sind Voreingenommenheiten subtiler und komplexer, da sie sich aus persönlichen Überzeugungen, Einstellungen und unbewussten Vorurteilen zusammensetzen, die das Denken und Handeln von Individuen beeinflussen. Diese inneren Zustände sind nicht direkt sichtbar oder messbar und variieren stark zwischen Individuen und Kontexten.

Voreingenommenheiten zu definieren und zu quantifizieren, erfordert daher ein tiefes Verständnis psychologischer Prozesse sowie die Entwicklung und Anwendung spezifischer Methoden, wie Umfragen, Interviews und Experimente, um diese indirekten und oft unbewussten Einflüsse zu erfassen.

Hinzu kommt, dass Voreingenommenheiten sich dynamisch verändern und in verschiedenen Situationen unterschiedlich ausgedrückt werden können, was ihre Messung zusätzlich erschwert. Die Quantifizierung erfordert oft die Interpretation von Verhaltensmustern und die Anwendung komplexer statistischer Methoden, um indirekte Beweise für Voreingenommenheiten zu identifizieren, davon abgesehen ist die Anerkennung von persönlichen Voreingenommenheiten oft mit Widerstand verbunden, da dies das Selbstbild und die Identität der Menschen herausfordern kann. Dies erschwert das Führen offener und ehrlicher Diskussionen über Voreingenommenheiten.

Auch wenn das Definieren und Quantifizieren von Voreingenommenheiten und Barrieren ein äußerst komplexes Unterfangen ist, ist es eine unverzichtbare Voraussetzung für effektive Maßnahmen gegen Ungleichheit und Diskriminierung.



Handlungsempfehlungen/Lösung

Unsere Lösung ist ein Konzept der Testautomation, welches nicht nur bei Conversational AI, sondern auch bei Menschen im telefonischen Kundendienst angewandt werden kann, um Bias und Barrieren zu identifizieren.

Das Konzept besteht aus drei Komponenten, die sich gegenseitig ergänzen und somit eine nahezu perfekte Lösung für diesen Anwendungsfall bieten:

- einer eigenen Conversational AI namens PersonAI, deren Hauptmerkmal das Simulieren verschiedener „Personas“ ist. Hierbei handelt es sich um künstliche Repräsentationen verschiedener Kund*innentypen, die eine breite Palette an Persönlichkeiten und Bedürfnissen abdecken. Durch diese Vielfalt ist PersonAI in der Lage, echte, individuelle Kund*innendienstgespräche zeitgleich über mehrere Instanzen zu führen. Das bedeutet im Umkehrschluss, dass beim Testen von CAI jede Personengruppe und deren Subgruppen, seien sie noch so klein, angemessen vertreten werden können. Nach jeder Sitzung werden Berichte erstellt, die beschreiben, welche spezifischen Interaktionsattribute und Verhaltensmuster registriert wurden, ob sie für die angewandte Persona als Barriere oder Voreingenommenheit zählen und wenn ja, inwiefern diese mit den jeweiligen Eigenschaften der simulierten Personengruppe korrelieren. Diese Berichte werden an die zuständigen Stellen geleitet, zum Beispiel eine Qualitätsprüfung eines Kund*innendienstes oder der eines CAI-Entwicklers, damit sie die nötigen Anpassungen vornehmen können, um den beschriebenen Interaktionsmerkmalen in Zukunft vorzubeugen.
- einer Sammlung an definierten und quantifizierten Voreingenommenheiten und Barrieren, die als Trainingsdatensatz für PersonAI gelten und unserer CAI ermöglichen, solche zu erkennen, namens Bias-Barriere-Index. Dieser wird von einer Ethikkommission erstellt und gewartet. Die Implementierung der Ergebnisse erfolgt über Expert*innen aus der Datenanalyse.

- der Taskforce, einer Ethik-Kommission, welche nicht nur für den Bias-Barriere-Index zuständig ist, sondern auch die Ergebnisse von PersonAI stichpunktartig überprüft, um die Integrität und Effektivität der CAI zu gewährleisten. Sie setzt sich aus unabhängigen Expert*innen verschiedener Disziplinen zusammen, darunter auch Vertreter*innen von Minderheiten, um eine umfassende Perspektive zu gewährleisten.

In Kombination bieten diese drei Säulen eine sichere und effektive Möglichkeit, Kund*innendienste auf Voreingenommenheiten und Barrieren zu testen, was zu einem inklusiven Kund*innendienstelerlebnis für alle Personengruppen führt. Dabei darf auch nicht außer Acht gelassen werden, dass diese Form der Automatisierung des Weiteren dazu beiträgt, den Ressourcenaufwand in Form von monetären Mitteln und Arbeitskraft zu reduzieren. Damit wird auch dem Fachkräftemangel und dem demografischen Wandel begegnet.



Fazit

Wir befinden uns in einer sehr schnelllebigen Welt, nicht zuletzt wegen der rasant fortschreitenden Entwicklung der Technologie, die uns ohne Zweifel viele Vorteile beschert, aber auch Gefahren birgt. Es ist von größter Wichtigkeit, dass jeder Mensch auf der Welt die gleichen Voraussetzungen hat, um ein erfülltes Leben frei von Hass und Ungleichberechtigung zu führen. Um dies zu gewährleisten, müssen wir jedoch darauf aufpassen, dass unser soziales Verständnis und dessen Implementierung in allen Lebensaspekten mit derselben Geschwindigkeit fortschreiten. Ähnlich wie in einem Wettrennen kann selbst ein vermeintlich kleiner Geschwindigkeitsunterschied über einen längeren Zeitraum hinweg unaufholbar sein.

Selbst mit den besten Intentionen birgt eine Hemmung des technologischen Fortschritts, zum Beispiel in Form von strengen Regulierungen, katastrophale Folgen für die jeweiligen Nationen, indem sie wirtschaftlich zurückfallen und auf dem Weltmarkt weniger wettbewerbsfähig werden. Dies führt zu einem Verlust an Arbeitsplätzen, geringeren Investitionen und schließlich zu einem Rückgang der Innovationskraft. In Anbetracht dessen ist die Nutzung von Technologie für die Sicherstellung von sozialen Prinzipien eine valide, wenn nicht schon notwendige Lösung. Selbstverständlich kann man sich auf keine Technologie zu hundert Prozent verlassen, man kann vielen Bedenken jedoch damit antworten, dass die Lösung auf einem transparenten und klar definierten Konzept basiert, welches von unabhängigen Expert*innen gewartet und überwacht wird.





Editorial

EduPin

AI Allies

GridHarmony

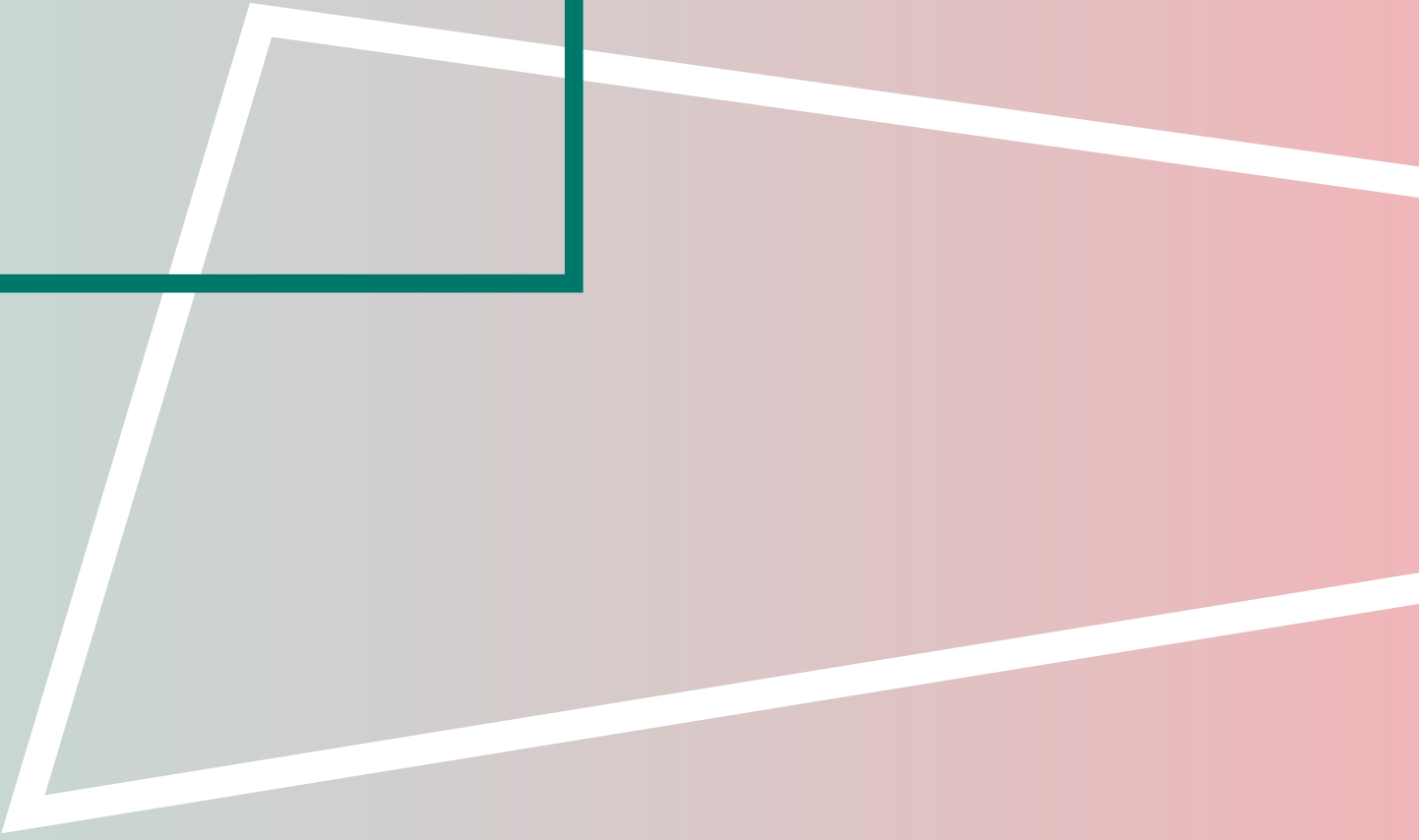
PowerBalance

Beyond
TheBarrier

Dialogreihe

Anhang

Dialogreihe



KI for Future? Ökologische Nachhaltigkeitsbewertung von künstlicher Intelligenz



Die Digital Future Challenge wurde durch die dreiteilige virtuelle Dialogreihe „Digital Responsibility“ begleitet, in der Expert*innen zu ökologischen, sozialen und Governance-Perspektiven auf den verantwortungsvollen Umgang mit KI eingeladen wurden und mit ihnen und den Teilnehmenden diskutiert wurde.

Die erste Veranstaltung betonte die Notwendigkeit, die ökologischen Auswirkungen des KI-Trainings und der Nutzung zu berücksichtigen und verantwortungsbewusst zu gestalten. Dabei rückten zwei Projekte in den Fokus: Zum einen **SustAI**, das ein Selbstbewertungstool entwickelt hat, um die Nachhaltigkeit von KI-Systemen zu bewerten und von **Dr. Anne Mollen** (Universität Münster und Algorithm Watch) vorgestellt wurde. Zum anderen **Escade**, ein Projekt für ressourcenschonende Rechenzentren, das von **Dusan Dokic** vom Deutschen Forschungszentrum für Künstliche Intelligenz präsentiert wurde.

Die Schlüsselerkenntnisse der Veranstaltung

- **Bewertungstools und Bewusstsein für nachhaltige KI:**
Das Selbstbewertungstool SustAI ermöglicht die Überprüfung der Nachhaltigkeit von KI-Modellen und schafft Bewusstsein für deren ökologische Auswirkungen.
- **Kriterien für eine nachhaltige KI-Definition:**
Eine nachhaltige KI sollte planetare Grenzen respektieren, problematische ökonomische Dynamiken nicht verstärken und den gesellschaftlichen Zusammenhalt nicht gefährden.
- **Notwendigkeit von individuellem und organisatorischem Mindshift:**
Bei der Entwicklung einer nachhaltigen KI sind nicht nur technische Komponenten, sondern auch individuelle und organisatorische Mindshifts sowie geeignete Rahmenbedingungen von entscheidender Bedeutung.



Ein Jahr mit ChatGPT: Welche Auswirkungen hat die Anwendung auf die Digitale Gesellschaft



Zum einjährigen „Geburtstag“ von ChatGPT in der offenen Nutzung wurde der Scheinwerfer auf die dynamischen Entwicklungen gelegt, die durch die Anwendung in der digitalen Gesellschaft aufkommen sind. Viele Expert*innen haben zu Fragen rund um das Lagebild zu Veränderungen auf Gesellschafts-, Wirtschafts- und Kompetenzebene Impulsvorträge gehalten.

So ging es etwa darum, wie verbreitet ChatGPT ist (**Sandy Jahn, Initiative D21**), welche Rolle Wahrhaftigkeit, Desinformation und Quellen bei der Nutzung von ChatGPT spielen (**Dr. Johanna Sprondel, Urania**) oder welche Möglichkeiten und Grenzen in einer Kulturerbeeinrichtung existieren (**Clemens Neudecker, Staatsbibliothek zu Berlin**). Darüber hinaus zeigte **Thomas Weber (McKinsey)** auf, in welcher Form der Einsatz in Unternehmen ausgestaltet sein kann, und **Prof. Dr. Martin Ehlert (Wissenschaftszentrum Berlin)** ging aus sozialwissenschaftlicher Perspektive auf Änderungen des Arbeitsmarktes ein. **Stefan Schönwetter** von der **Deutschen Kinder- und Jugendstiftung** berichtete darüber, welche Kompetenzen – insbesondere im Fokus auf Schulbildung – betrachtet werden sollten. Mit Blick auf Mitarbeitende gab **Sindy Leffler-Krebs** einen Einblick in die Vorgehensweise der **Deutschen Telekom** zu Kompetenzaufbau und Compliance-Perspektiven.

Zum Abschluss der Veranstaltung wurde ein Promptathon (**Dr. Marie Blachetta, Initiative D21**) durchgeführt, bei dem es insbesondere um die kompetente und verantwortungsvolle Nutzung von ChatGPT ging.

Die Schlüsselerkenntnisse der Veranstaltung

- **Multiperspektivität und Kompetenzaufbau sind zentral:**
Der Umgang mit KI-Tools wie ChatGPT erfordert die Berücksichtigung verschiedener Perspektiven und den Aufbau digitaler Kompetenzen, wobei Maßnahmen wie Promptathons helfen könnten.
- **Gestaltungsdimensionen von Upskilling-Maßnahmen:**
Upskilling-Maßnahmen müssen so gestaltet werden, dass sie nicht zu einer Spaltung in der Gesellschaft führen und gleichzeitig ethische Aspekte und veränderte Kompetenzanforderungen im Kontext von KI angemessen berücksichtigen.

Is this fAIr? Der Einsatz von KI im Human Resource Management



Dr. Alina Köchling (Deloitte) präsentierte Forschungsergebnisse, die zeigen, dass KI-Tools, wie Videosoftware und Chatbots, zunehmend in der Rekrutierung genutzt werden. Trotz Vorteilen wie höherer Produktivität und Risikominimierung können KI-Tools gesellschaftliche Vorurteile reproduzieren. Insbesondere Anwendungen wie asynchrone Video-Interviews stoßen bei Bewerbenden auf negatives Feedback. **Roberta Barone (Stepstone)** zeigte anschließend auf, welcher digital-ethische Zusammenhang – insbesondere im HRM – besteht. Dabei bezog sie sich sowohl auf den AI Act als auch auf vier Prinzipien für einen verantwortungsvollen Einsatz von KI im HRM.



Die Schlüsselerkenntnisse der Veranstaltung

- **Breite Nutzung von KI im HRM:**
Bis zu 86 Prozent der Arbeitgeber*innen verwenden technologiegestützte Videosoftware für Interviews. Die Anwendungsbereiche umfassen das algorithmische Prüfen von Lebensläufen, Telefon- und Chatinterviews sowie asynchrone Video-Interviews.
- **Negative Reaktionen der Bewerbenden:**
Obwohl KI im HRM zunehmend populär ist, erleben Bewerbende negative Auswirkungen, insbesondere in Bezug auf wahrgenommene Performance-Möglichkeiten, „perceived creepiness“ und Datenschutzbedenken.
- **Notwendigkeit von ethischen Grundsätzen und Regulierungen:**
Der „Aufstieg der KI-Ethik“ wird betont, und verantwortungsvolle KI-Prinzipien wie Unabhängigkeit, Fairness, Schadensprävention und Erklärbarkeit werden als notwendig erachtet. Die EU plant mit dem AI Act, Regulationen für den Einsatz von KI zu schaffen, und teilt Anwendungen in Risikokategorien ein. Unternehmen sollten sich bewusst sein, wie ihre Technologien auf ethische Weise genutzt werden können, um negative Auswirkungen zu minimieren.



Editorial

EduPin

AI Allies

GridHarmony

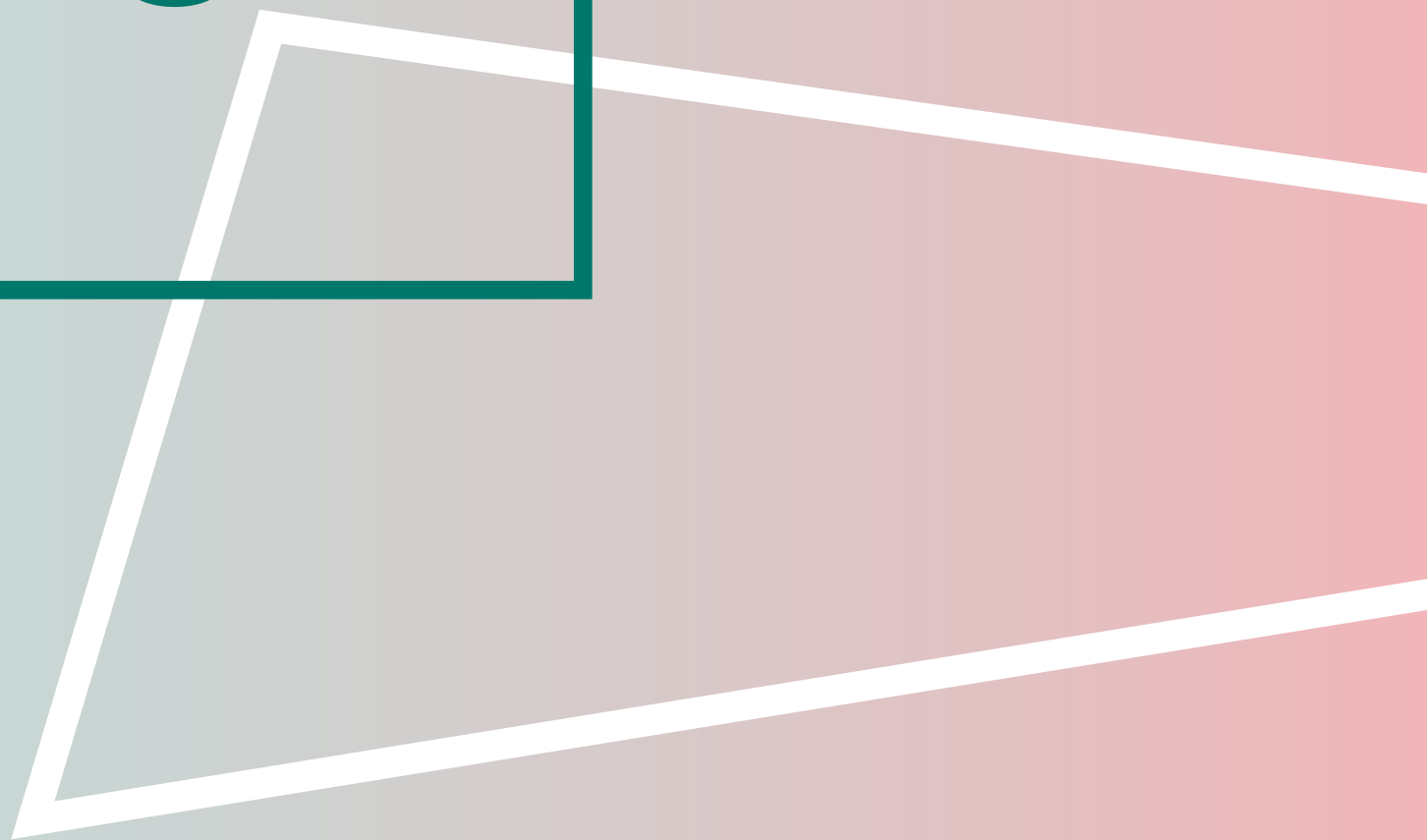
PowerBalance

Beyond
TheBarrier

Dialogreihe

Anhang

Anhang



ESG-Kriterien

Die Use Cases der Digital Future Challenge richten den Fokus auf die drei ESG-Kriterien (Environmental, Social & Governance) und kategorisieren sich in den verantwortungsvollen Umgang mit KI aus **ökologischer, sozialer** und **Governance-Perspektive**.



Unter die Kategorie der **ökologischen Perspektive** zählten folgende Fragestellungen:

- **Bertelsmann Stiftung:** Zusammentreffen von Klimawandel und Digitalisierung: Welche Rolle spielen KI und New Work?
- **Honda Research Institute GmbH:** Faires KI Energiemanagement und ICH
- **ReFrastructure – Stiftung für digitale Mehrweginfrastruktur gGmbH:** Kann eine KI einen Mehrwert hinsichtlich der Ökobilanz für eine Logistiklösung schaffen oder nicht?



Mit der **sozialen Perspektive** der ESG-Kriterien beschäftigten sich die folgenden Use Cases:

- **IBM Deutschland GmbH:** Human Friendly Automation – Analyse der Auswirkungen auf die Betroffenen von Automatisierung
- **PD – Berater der öffentlichen Hand GmbH:** Zielgerichtete Datensammlung für den Einsatz von KI in der datenbasierten Verkehrsplanung
- **O₂ Telefónica Germany:** Zugängliche und chancengleiche Conversational AI-Systeme in der Interaktion zwischen Mensch und Technologie



Die **Governance-Perspektive** wurde von den nachfolgenden Anwendungsfällen behandelt:

- **adesso SE:** Kann und darf KI schon allein programmieren?
- **ING Deutschland:** KI basiertes Transaktionsmonitoring als Lösung für Geldwäsche-Prävention
- **Deutsche Telekom AG:** Wie können Kund*innen eindeutig darüber informiert werden, dass sie es mit einer KI zu tun haben?

Sustainable Development Goals

Darüber hinaus können die entwickelten Lösungskonzepte den Sustainable Development Goals (SDGs) der United Nations (UN) zugeordnet werden. Die SDGs bestehen aus 17 übergeordneten Zielen (vgl. Abb. 12) für eine nachhaltige Entwicklung sowie den dazugehörigen 169 Unterzielen. Sie bilden einen umfassenden programmatischen Rahmen zur Förderung einer weltweit nachhaltigen Gesellschaft.



Abb. 12 – 17 Sustainable Development Goals der UN

Die Lösungskonzepte, die von den Studierenden erarbeitet wurden, tragen zur Umsetzung der von der UN definierten 17 Zielen bei. Zur Veranschaulichung sind hier exemplarisch je zwei Beispiele pro Ausarbeitung dargestellt:

EduPin – der smarte Verkehrsbegleiter für inklusive Kommunen

- SDG 4 – Hochwertige Bildung durch die Förderung der Verkehrserziehung
- SDG 11 – Nachhaltige Städte und Gemeinden durch die verbesserte Sicherheitsplanung des Straßenverkehrs

AI Allies – Conversational AI für alle: Barrierefreiheit und Gleichberechtigung in O₂ Telefónicas „Aura“

- SDG 10 – Weniger Ungleichheiten durch inklusive Angebote wie z. B. Gebärdensprache
- SDG 12 – Nachhaltige/r Konsum und Produktion durch inklusive Programmierung und das Training der Conversational AI

Grid Harmony – Für eine faire Energiezukunft: Die smarte Lösung mit GridHarmony!

- SDG 7 – Bezahlbare und saubere Energie durch faire Verteilung von Energie in Zeiten von Knappheit
- SDG 13 – Maßnahmen zum Klimaschutz durch Anreize im Verbraucherverhalten

PowerBalance – DER Gamechanger für die Energiewende!

- SDG 7 – Bezahlbare und saubere Energie durch faire Verteilung von Energie in Zeiten von Knappheit
- SDG 13 – Maßnahmen zum Klimaschutz durch Anreize im Verbraucherverhalten

BeyondTheBarrier – Wie ermöglicht man eine vorurteils- und barrierefreie Interaktion mit Conversational AI?

- SDG 10 – Weniger Ungleichheiten durch Berücksichtigung von gesellschaftlichen Unterschieden und sozialer Herkunft, wie etwa sprachlicher Barrieren in der Mensch-Maschine-Interaktion
- SDG 12 – Nachhaltige/r Konsum und Produktion durch inklusive Programmierung und Training der Conversational AI

Wir danken allen Unterstützer*innen



Impressum und Informationen

Über die Deloitte-Stiftung

Unter dem Leitmotiv „Gemeinsam Zukunft unternehmen“ bündelt die Deloitte-Stiftung ihr Engagement für Bildung und Wissenschaft. Mit ihren Aktivitäten fördert die Deloitte-Stiftung die Aus- und Weiterbildung junger Menschen, unterstützt die Entwicklung neuer, zukunftstauglicher Bildungsmodelle und leistet einen Beitrag zur Verbesserung des Innovationstransfers von der Hochschule in die Wirtschaft. Die Deloitte-Stiftung wurde von den Gesellschaftern und Führungskräften von Deloitte als eine nicht-rechtsfähige Stiftung in der treuhänderischen Verwaltung des Stifterverbandes gegründet.

Anschrift und Kontakt:

Deloitte-Stiftung im Stifterverband
c/o Deutsches Stiftungszentrum GmbH
Baedekerstr. 1
45128 Essen
info@deloitte-stiftung.de

Vorstand des Stifterverbandes (i.S.d.V.):

Prof. Dr. Michael Kaschke (Vorsitzender)
Dr. Simone Bagel-Trah
Dr. Martin Bruder Müller
Dr. phil. Nicola Leibinger-Kammüller
Dr. Cornelius Riese
Dr. rer. oec. Reinhard Christian Zinkann
Dr. Volker Meyer-Guckel

Vorstand der Deloitte-Stiftung:

Thomas Northoff (Vors.)
Dorothea Schmidt
Prof. Dr. Stefan Stolte

Über die Initiative D21 e.V.

Die Initiative D21 ist Deutschlands größtes gemeinnütziges Netzwerk für die digitale Gesellschaft. Sie wurde 1999 mit dem Ziel gegründet, die digitale Spaltung der Gesellschaft zu verhindern, und setzt sich bis heute zum Ziel, dass alle Menschen bestmöglich von der Digitalisierung profitieren. Gemeinsam mit einem großen Netzwerk bestehend aus rund 140 Mitgliedsunternehmen und -organisationen aus allen Branchen, der Wissenschaft, der Zivilgesellschaft sowie mit politischen Partner*innen aus Bund und Ländern beleuchtet die Initiative D21 die gesellschaftlichen Herausforderungen des digitalen Wandels. Sie liefert jährliche Lagebilder und stößt Debatten an, um die Zukunft der digitalen Gesellschaft sinnvoll zu gestalten.
Mehr unter: www.initiativesd21.de

Anschrift und Kontakt:

Initiative D21 e.V.
Reinhardtstraße 38
10117 Berlin
T +49 30 7675853-50
kontakt@initiativesd21.de

Präsidium des Vereins:

Marc Reinhardt, Valentina Daiber, Christiane Fritsch,
Dr. Irina Eckardt, Thomas Langkabel, Timm Lutter,
Prof. Barbara Schwarze

Hauptverantwortliche der DFC bei Deloitte-Stiftung und Initiative D21:

Dr. Marie Blachetta, Dr. Alina Köchling, Lena-Sophie Müller,
Thomas Northoff, Annabel Rust, Rebecca van der Meyden

Redakteur*innen der Publikation:

Dr. Marie Blachetta und Dr. Alina Köchling

Lektorat:

Cynthia Erhardt

Design:

Kantar, Natalie Birnbaum

Vielen Dank an Kantar, welches die Umsetzung der Publikation pro bono übernommen hat.

Bei den ausgearbeiteten Lösungen der Teams sind die Ideen, Meinungen und Lösungsansätze der Studierenden wiedergegeben. Die Positionen der Deloitte-Stiftung, der Initiative D21 oder des Bundesministeriums für Digitales und Verkehr können davon abweichen.



Die Publikation zur Digital Future Challenge #4, eine Publikation der Deloitte-Stiftung und der Initiative D21, ist lizenziert unter einer „Creative Commons Namensnennung 4.0 International“-Lizenz.

