

Innerhalb der KI

Ein algorithmisches Abenteuer



Veröffentlicht im Jahr 2023 von der Organisation der Vereinten Nationen für Erziehung, Wissenschaft und Kultur, 7, place de Fontenoy, 75352 Paris 07 SP, Frankreich, der Deutschen UNESCO-Kommission, Martin-Luther-Allee 42, 53175 Bonn, Deutschland und dem International Research Centre on Artificial Intelligence (IRCAI) under the auspices of UNESCO, Jamova cesta 39, SI-1000 Ljubljana, Slowenien.

© Deutsche UNESCO-Kommission, IRCAI, UNESCO 2023

ISBN 978-3-947675-49-4



Diese Publikation ist als Open Access unter der Lizenz Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) verfügbar (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). Mit der Nutzung des Inhalts dieser Publikation erklären sich die Nutzer mit den Nutzungsbedingungen des UNESCO Open Access Repository (<http://en.unesco.org/open-access/terms-use-ccbysa-en>) einverstanden.

Originaltitel: Inside AI: An Algorithmic Adventure. Erstmals veröffentlicht im Jahr 2022 von der Organisation der Vereinten Nationen für Erziehung, Wissenschaft und Kultur (UNESCO), 7, place de Fontenoy, 75352 Paris 07 SP, Frankreich.

Die in dieser Publikation verwendeten Bezeichnungen und die Darstellung des Materials bedeuten nicht, dass die UNESCO, die Deutsche UNESCO-Kommission oder IRCAI eine wie auch immer geartete Meinung über den Rechtsstatus eines Landes, eines Gebiets, einer Stadt oder eines Bereichs oder seiner Behörden oder über den Verlauf seiner Grenzen zum Ausdruck bringen.

Die in dieser Publikation zum Ausdruck gebrachten Ideen und Meinungen sind die der Autoren; sie entsprechen nicht unbedingt denen der UNESCO, der Deutschen UNESCO-Kommission oder IRCAI und verpflichten die Organisationen in keiner Weise. Die Abbildungen in dieser Publikation sollen nicht dazu dienen, Personen oder Regionen zu stereotypisieren.

Produktion: BiMot Culture

Konzeption und Skript: Dr. Katherine Evans

Redaktionelle Koordination und Grafikdesign: Muyang Li

Berater für Drehbuchgestaltung: Patrick Marty

Kreativberaterin für die Māori-Kultur: Cian Elyse White

Illustration von Episode I: Isobel Joy Te Aho-White

Illustration von Episode II: Cassandra Okwaniuzor Mark

Illustration von Episode III: Asma Kraiem

Illustration von Episode IV: Adriana De La Torre Cervantes, unter Mitwirkung von Patricia Manríquez und Karin Almazán

Titelblattgestaltung: Isobel Joy Te Aho-White

Charakterdesign der Pannen: Isobel Joy Te Aho-White

Charakterdesign von Doktor Y und Datenprofile: Cassandra Okwaniuzor Mark

Charakterdesign von Emile Tach: Asma Kraiem

Vorwort: Dr. Tawfik Jelassi

Lektorat: Julia Lins-Gordon

Übersetzung: Dr. Maximilian Müngersdorff, Anke Seibert

Welchen Einfluss hat Künstliche Intelligenz (KI) auf die Menschheit?

Überall auf der Welt, insbesondere im Globalen Süden, laufen Menschen Gefahr, von den Vorteilen des digitalen Wandels ausgeschlossen zu werden. Dies trifft in besonderem Maße auf Mädchen und Frauen zu. **Auf dem afrikanischen Kontinent nutzen nur 20,2 % der Frauen und 37,1 % der Männer das Internet^[a]**, obwohl die Online-Nutzung heute eine der wichtigsten Voraussetzungen für die soziale und wirtschaftliche Teilhabe ist.

- Im Rahmen der Bemühungen der UNESCO, neue Technologien für eine nachhaltige Entwicklung nutzbar zu machen, erklärt dieser Comicroman für junge Erwachsene die Auswirkungen von KI auf die Menschheit.
- Ein junges Publikum folgt vier Charakteren aus verschiedenen Weltregionen in die algorithmische Galaxie „Plethor.A.I.“ und setzt sich dabei mit den Grenzen der KI-Technologie auseinander. Dort reisen sie durch die verborgene Welt hinter unseren Bildschirmen, um etwas über die sozialen, technischen, ethischen und menschenrechtlichen Auswirkungen von KI zu erfahren, und den Charakteren zu helfen, einen Weg zurück in die Realität zu finden.



[a] S. 17 aus dem UNESCO-Bericht *The Effects of AI on the Working Lives of Women*.

Innerhalb der KI

Ein algorithmisches Abenteuer

Inhalt

Vorwort	6
Episode I Ari und die Paralleluniversität	7
Einleitung	8
Glossar	41
Schlussfolgerung	43
Episode II Maryam und der Deep Dive ins Datenbecken	44
Einleitung	45
Glossar	78
Schlussfolgerung	80
Episode III Shirin auf der Spur des Empfehlungsgebers	81
Einleitung	82
Glossar	115
Schlussfolgerung	117
Episode IV Joaquín und die unlösbare Frage	118
Einleitung	119
Glossar	155
Schlussfolgerung	157
Quellen und weiterführende Literatur	158
Danksagungen	159

Vorwort

Künstliche Intelligenz (KI) und ihre Anwendungen sind heute keine Science-Fiction mehr. Wahrscheinlich haben Sie selbst schon mit KI interagiert, durch Sprachassistenten oder Gesichtserkennungs-Tools oder durch Empfehlungsalgorithmen sozialer Medien. KI eröffnet nie dagewesene Möglichkeiten, von personalisierten Empfehlungen für Milliarden von Menschen bis hin zur Vorhersage von Krankheiten, bevor sie auftreten.

Gerade weil KI ein so mächtiges Werkzeug ist, müssen wir uns auch mit ihren Risiken und möglichen negativen Auswirkungen befassen. Was macht KI bei bestimmten Entscheidungen «voreingenommen», z.B. bei der Frage, wer eingestellt werden soll? Wie kommt es, dass sie bestimmte Personengruppen ausschließt? Wie können wir Diskriminierung vermeiden, Inklusion und Vielfalt gewährleisten, unsere Privatsphäre schützen und komplexe Blackbox-Herausforderungen in der Funktionsweise von KI meistern?

Dies sind nur einige der Fragen zu KI, die die UNESCO in ihren Kompetenzbereichen Bildung, Wissenschaft, Sozial- und Geisteswissenschaften, Kultur, Kommunikation und Information zu beantworten versucht. So hat sie beispielsweise im Jahr 2021 einen kostenlosen Online-Kurs zu KI und Menschenrechten angeboten, der komplexe KI-Konzepte für Jugendliche auf der ganzen Welt verständlicher gemacht hat. Mit dieser und anderen Initiativen, wie der ebenfalls 2021 von der UNESCO-Generalkonferenz auf ihrer 41. Tagung

verabschiedeten Empfehlung zur Ethik der Künstlichen Intelligenz, wollen wir sicherstellen, dass KI in einer Weise eingesetzt wird, die mit internationalen Standards in Einklang steht und die Menschenrechte schützt, wie das Recht auf Gleichheit, Bildung, Privatsphäre, Zugang zu Informationen und Meinungsfreiheit.

Ich lade Sie ein, gemeinsam mit uns die Chancen und Herausforderungen der KI zu erkunden. Wir hoffen, dass dieser Comicroman nicht nur KI-Konzepte für jede und jeden, unabhängig vom Alter, zugänglich macht, sondern auch jeden Einzelnen dazu befähigt, sich seiner Grundrechte im digitalen Zeitalter bewusst zu werden und sie zu schützen.

Wir laden Sie ein, Ihre Gedanken zu #KünstlicherIntelligenz mit uns zu teilen und mit uns zu interagieren, indem Sie diesen Comicroman in Ihren sozialen Netzwerken kommentieren, reposten, teilen und mit @UNESCO markieren. Unter meinem X-Account @TawfikJelassi werde ich weitere Posts über den ersten Comicroman der UNESCO zum Thema KI veröffentlichen! ■

Dr. Tawfik Jelassi
Beigeordneter Generaldirektor der UNESCO
Kommunikations- und Informationssektor



Episode I

Ari und die Paralleluniversität

Konzeption und Skript: Dr. Katherine Evans

Illustration: Isobel Joy Te Aho-White

Kreativberaterin für die Māori-Kultur: Cian Elyse White



Vorwort

Hinter den vielen Bildschirmen, die Teil unseres Lebens sind, verbirgt sich eine Welt, die wir nur schwer durchschauen können. Dabei interagieren wir jeden Tag mit dieser Welt: Jedes Mal, wenn wir eine App auf unseren Computern und Smartphones öffnen oder eine Webseite aufrufen, öffnen wir eine Tür zu ihr. Jedes Mal, wenn wir einem Sprachassistenten eine Frage stellen oder unser Gesicht zur Identifizierung verwenden, bringen wir ihr bei, uns zu erkennen. Und jedes Mal, wenn wir einen Beitrag liken, ein Video ansehen oder etwas online kaufen, lassen wir sie etwas über uns erfahren. In vielerlei Hinsicht existiert diese Welt sowohl überall als auch nirgends und wird sowohl von allen als auch von niemandem beherrscht. Viele von uns haben davon gehört, einige von uns fürchten sie, aber nur wenige von uns verstehen genau, wie sie funktioniert.

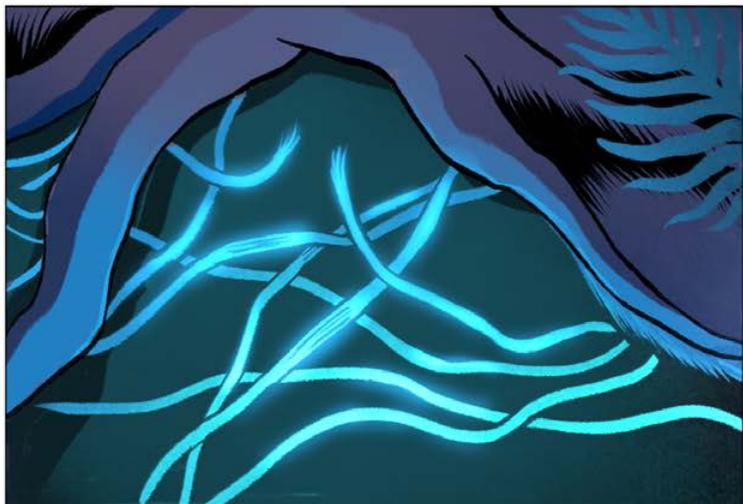
In den Medien, Lehrbüchern und im Internet nennen wir diese Welt gewöhnlich „Künstliche Intelligenz“ oder kurz „KI“. Technisch gesehen besteht sie aus *Daten* - Informationen oder Fakten über uns und die Welt, in der wir leben, aus *Algorithmen* - Regeln oder Verfahren, die zum Erreichen eines bestimmten Ziels befolgt werden können und aus *Konnektivität* - Hardware, die alles miteinander verbindet und Daten sowohl über den ganzen Globus als auch direkt an die Geräte von Menschen wie uns sendet. Auch wenn wir es vielleicht nicht erkennen, hat dieses Trio aus Daten, Algorithmen und Konnektivität einen tiefgreifenden Einfluss auf fast jeden Aspekt unseres Lebens und sogar auf das Leben auf der Erde selbst und wird dies auch weiterhin haben.

Das liegt auch daran, dass KI den Menschen helfen soll: Sie kann Aufgaben übernehmen, die der Mensch früher allein bewältigen musste oder Dienstleistungen anbieten, ohne die viele von uns nicht leben können. Sie kann uns wertvolle Einblicke in komplexe Probleme geben, die zu groß sind, als dass ein einzelner Verstand sie entschlüsseln könnte, und sie kann sogar kreativ sein und neue Lösungen finden, auf die niemand von uns gekommen wäre. In gewisser Weise ist KI das nützlichste Werkzeug, das die Menschheit je geschaffen hat, und seine Anwendungsmöglichkeiten sind nur durch unsere menschliche Vorstellungskraft begrenzt.

Doch wie jedes Werkzeug ist auch KI nur dann nützlich, wenn sie richtig eingesetzt wird. Wenn wir wollen, dass KI uns hilft, dann müssen wir darauf Acht geben, dass wir uns nicht in ihren Versprechungen und Möglichkeiten verlieren und dabei das Wichtigste aus den Augen verlieren: den Menschen, seine Rechte und Würde und unsere Umwelt. Die KI mag virtuell sein, aber ihre Auswirkungen auf die Menschheit und den Planeten sind sehr real.

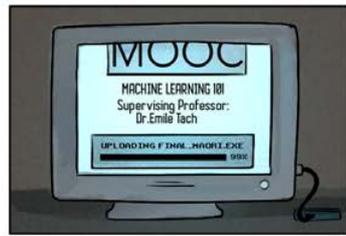
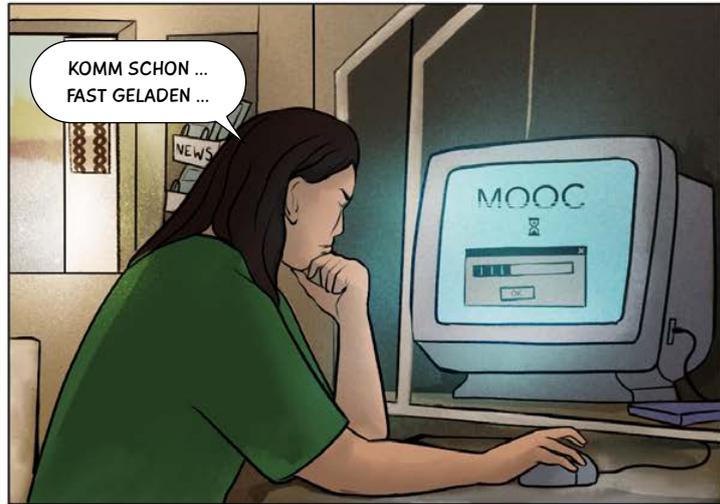
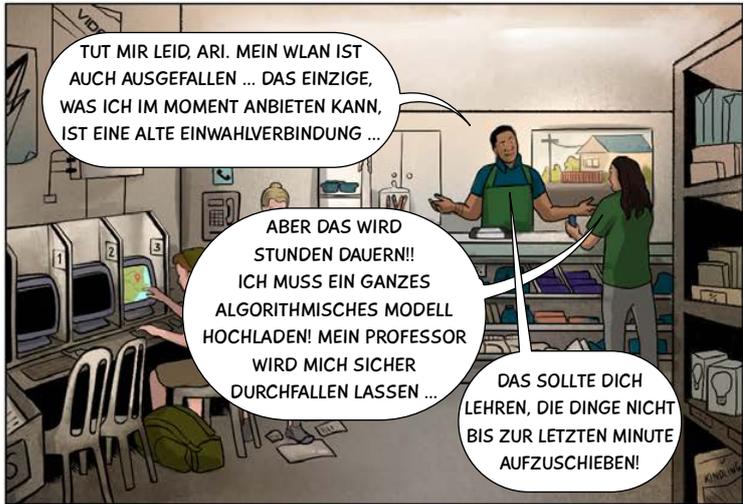
Alles beginnt mit KI-Entdeckerinnen und Entdeckern wie Euch, die mutig genug sind, sich dem verborgenen Universum jenseits unserer Bildschirme zu stellen und seine geheimnisvolle Verbindung zu unserer Realität zu erforschen. ■

SPÄT IN DER NACHT, IRGENDWO IN AOTEAROA, NEUSEELAND,
FEGT EIN HEFTIGER STURM ÜBER DAS LAND ...





*„KIA ORA E HOA!“ (HALLO, MEIN FREUND!)





DIE HAUSARBEIT VON ARI BEGINNT IHRE LANGE REISE ZU IHREM DIGITALEN ZIEL ...

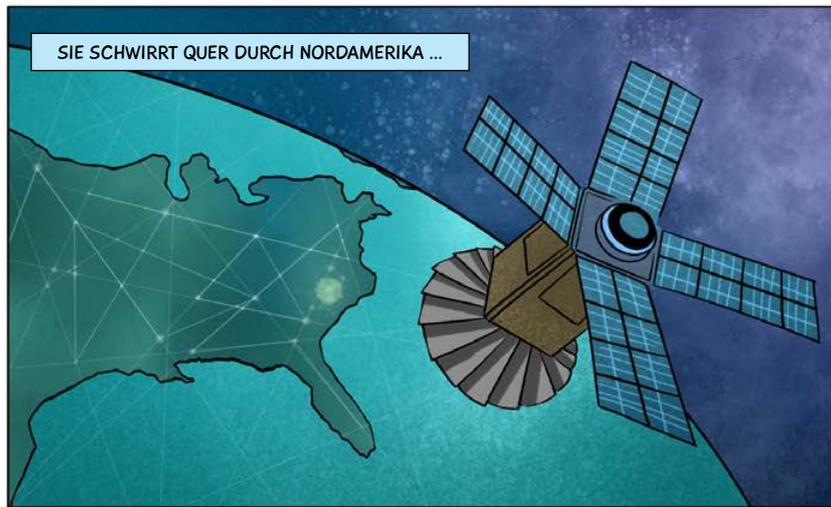


SIE RAST DURCH DIE UNTERIRDISCHE INFRASTRUKTUR ...





... UND DIE TRANSATLANTISCHEN KABEL ...



SIE SCHWIRRT QUER DURCH NORDAMERIKA ...

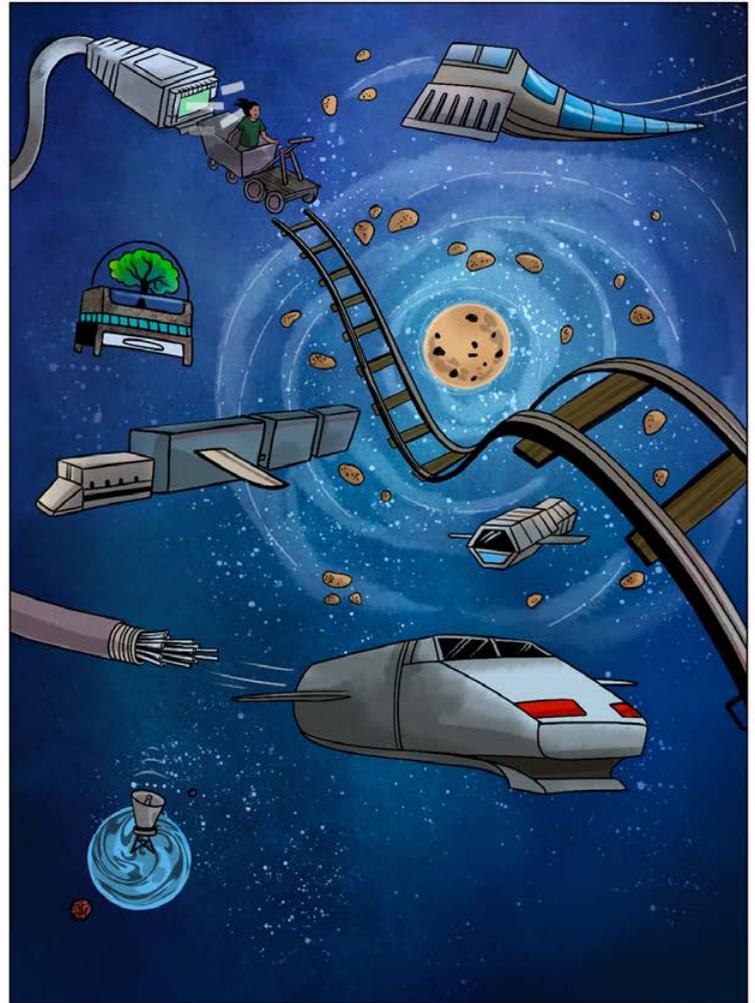


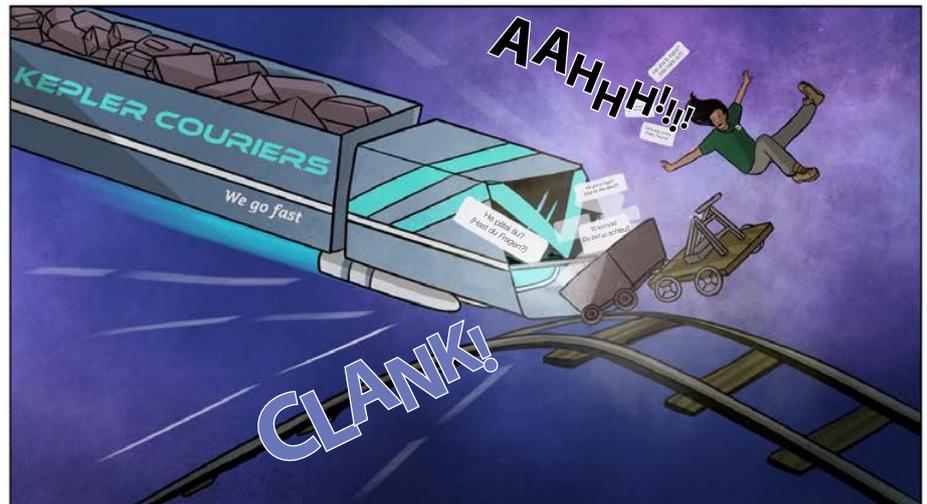
... BIS SIE SCHLIESSLICH IN EINER SERVERFARM IRGENDWO IM LÄNDLICHEN VIRGINIA LANDET. IRGENDWAS SCHEINT JEDOCH SCHIEFGELAUFEN ZU SEIN ...



SEHR SCHIEF ...

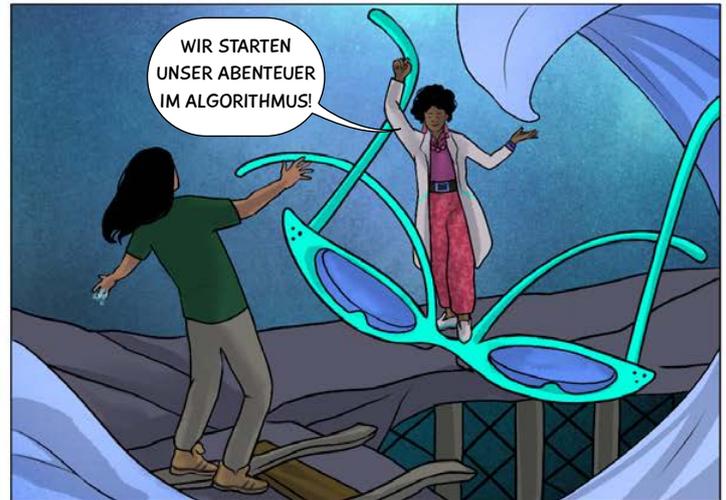


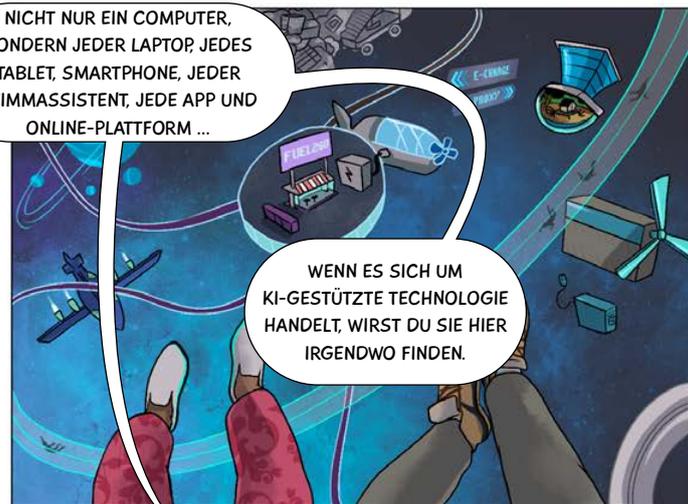
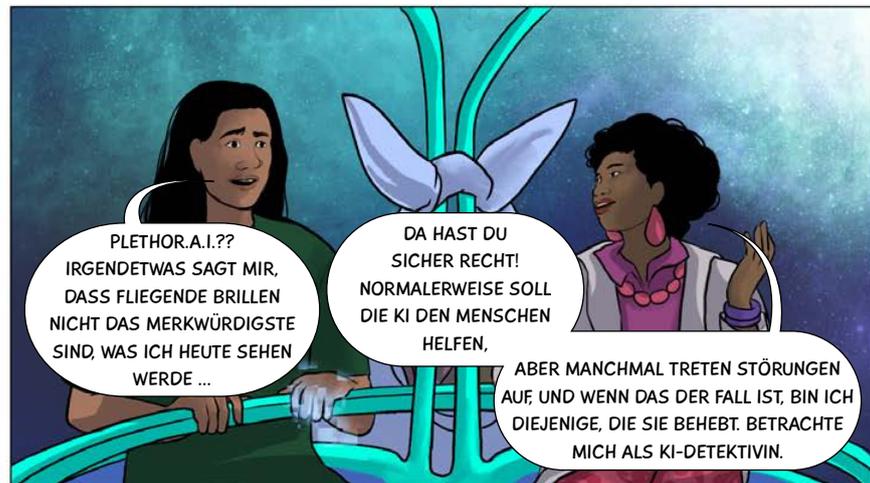














MOMENT MAL ... DU MEINST, MEIN HANDY FLIEGT HIER IRGENDWO HERUM?

NUN, NEIN ... ABER DIE ALGORITHMEN, DIE ES STEuern, SCHON.

GANZ ZU SCHWEIGEN VON ALL DEN DATEN, DIE ES SAMMELT UND VERWENDET ...



ABER WIESO BIN ICH HIER? ES KLINGT, ALS OB ES HIER NICHT VIELE TOURISTEN GÄBE ...

WENN ICH ES MIR RECHT ÜBERLEGE, NEIN!

DIE MEISTEN MENSCHEN IN PLETHOR.A.I. SIND DAS, WAS WIR KI-PROFIS NENNEN: INGENIEURE, PROGRAMMIERERINNEN, JURISTEN, DATENWISSENSCHAFTLERINNEN ...

UND SOGAR EIN PAAR PHILOSOPHINNEN WIE ICH!



NUN, ICH MAG INFORMATIK, ABER ICH WÜRDIE MICH KAUM ALS PROFI BEZEICHNEN ...

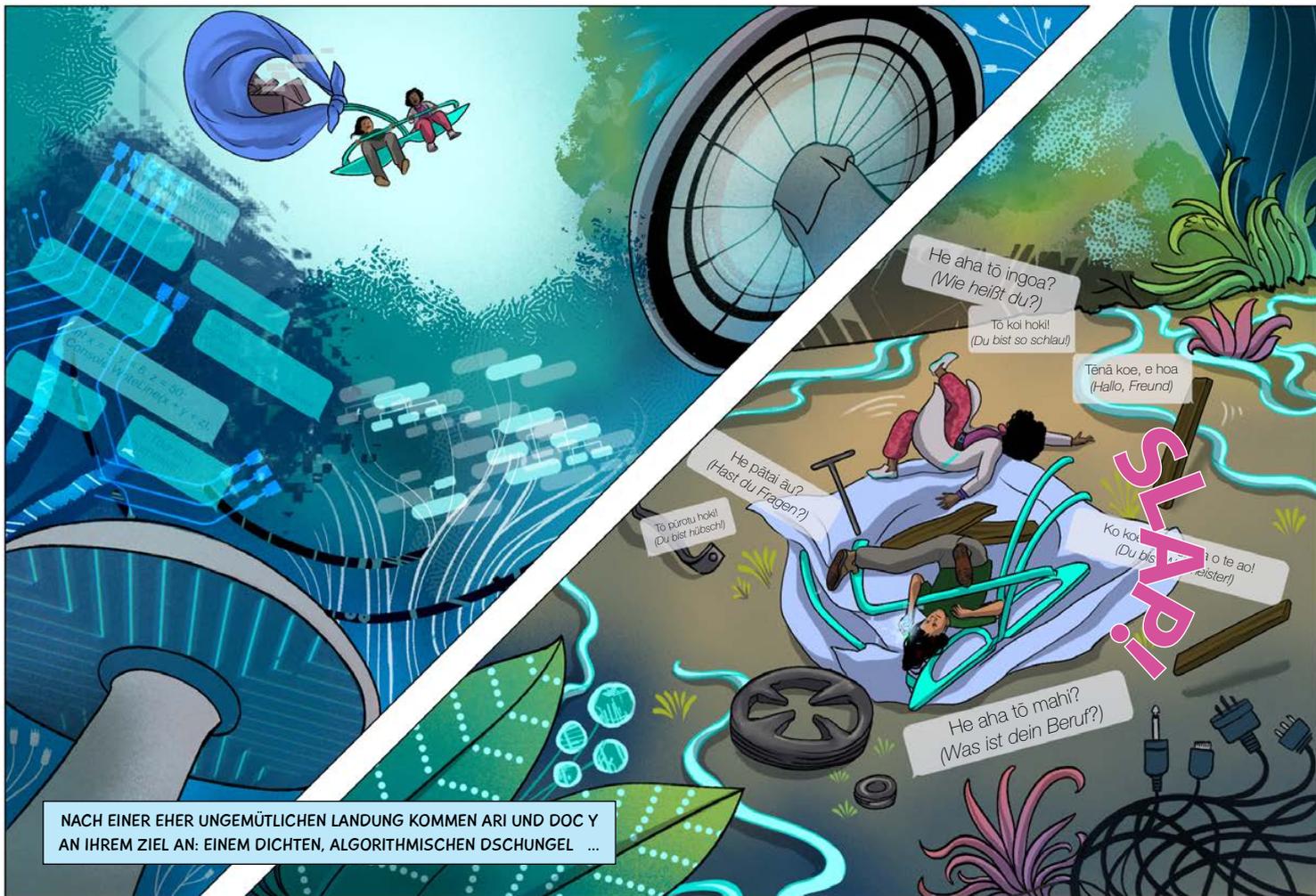
PROFI ODER NICHT, WIR MÜSSEN DIESE STÖRUNG BEHEBEN, WENN DU NACH HAUSE ZURÜCKKEHREN WILLST.

UND ICH DENKE, ICH KENNE GENAU DIE RICHTIGE PERSON FÜR DIESEN JOB!



ICH HOFFE ES ... ICH MUSS SICHERSTELLEN, DASS MEINE HAUSARBEIT PÜNKT LICHT ANGEKOMMEN IST ... MEIN PROFESSOR FLIPPT SONST AUS!

AH JA ... PROFESSOREN SIND NICHT IMMER EINFACH, NICHT WAHR ...



NACH EINER EHER UNGEMÜTLICHEN LANDING KOMMEN ARI UND DOC Y AN IHREM ZIEL AN: EINEM DICHEN, ALGORITHMISCHEN DSCHUNGEL ...

He aha tō ingoa?
(Wie heißt du?)

Tō kai hoki!
(Du bist so schlau!)

Tēnā koe, e hoa
(Hallo, Freund)

He pātai āu?
(Hast du Fragen?)

Tō pūroto hoki!
(Du bist hübsch!)

SLAPI

Ko kōwhiri
(Du bist ein Meister)

He aha tō mahi?
(Was ist dein Beruf?)





... UND DANN IST DA NOCH DIE SACHE MIT DEN COMPUTERN. SELBST WENN ICH ALLE DATEN HÄTTE, HÄTTE ICH KEINEN ZUGANG ZU EINEM COMPUTER, DER LEISTUNGSFÄHIG GENUG WÄRE, UM DEN ALGORITHMUS ZU TRAINIEREN ...

ES IST HOFFNUNGSLOS.

HMMM...



DU BRAUCHST ZUGANG ZU DATEN, ABER DU ARBEITEST MIT EINER SPRACHE, DIE WIR „RESSOURCENARM“ NENNEN.

ES GIBT EINEN BEGRIFF DAFÜR?

JA! IN KOMBINATION MIT DEINER EINWAHLVERBINDUNG UND EINEM OFFENBAR SEHR VERALTETEN COMPUTER BIST DU GEGENÜBER DEN MEISTEN JUNGEN PROGRAMMIERERINNEN UND PROGRAMMIERERN KLAR IM NACHTEIL.



... KLINGT FÜR MICH SO, ALS OB DU AUF DER FALSCHEN SEITE DESSEN STEHST, WAS WIR DIE „DIGITALE KLUFFT“ NENNEN ...

HÄ?

ES IST EIN ETHISCHES UNGLEICHGEWICHT, DAS EINIGE MENSCHEN ODER GEMEINSCHAFTEN DARAN HINDERT, KI-LÖSUNGEN ZU ENTWICKELN, VON DENEN ALLE PROFITIEREN WÜRDEN ... ODER IN DEINEM FALL, SIE DARAN HINDERT, KI-ANWENDUNGEN ZU ENTWICKELN, DIE DEINER GEMEINSCHAFT DIENEN.



DAS HÖRT SICH NACH DER SITUATION AN, IN DER ICH HEUTE WAR ...

ABER DU HAST GESAGT, DASS GEMEINSAME DATEN WIE DIESE HIER HELFEN SOLLTEN?

DAS TUN SIE GANZ SICHER! ABER BEI DER ÜBERBRÜCKUNG DER DIGITALEN KLUFFT GEHT ES NICHT NUR DARUM, DIE DATEN OFFENER ZU MACHEN.

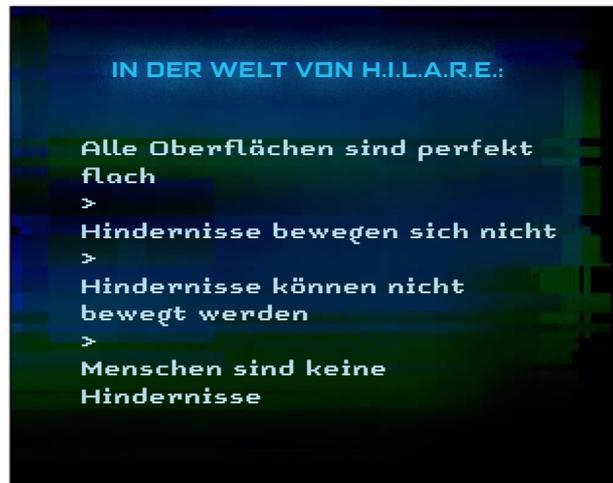
WIR MÜSSEN EINE KI AUFBAUEN, DIE AUF ALLEN EBENEN ZUGÄNGLICH UND INTEGRATIV IST.



AHA! JETZT KLINGST DU WIRKLICH WIE EINE PHILOSOPHIN ...

... ARI, PASS AUF!!!





Dumme Regeln

> H.I.L.A.R.E. KANN NICHT FLIEGEN

> H.I.L.A.R.E. KANN NICHT TELEPORTIEREN

Gute Regeln

- > H.I.L.A.R.E. KANN SICH IN 4 RICHTUNGEN BEWEGEN
- > H.I.L.A.R.E. KANN SICH NICHT DURCH HINDERNISSE BEWEGEN
- > H.I.L.A.R.E. KANN NICHT AN ZWEI ORTEN GLEICHZEITIG SEIN

> H.I.L.A.R.E. KANN KEINE HINDERNISSE ZERSTÖREN

> H.I.L.A.R.E. HAT KEINE HÄNDE

> H.I.L.A.R.E. IST KEIN MENSCH

> H.I.L.A.R.E. IST KEIN SCHICKER SPORTWAGEN





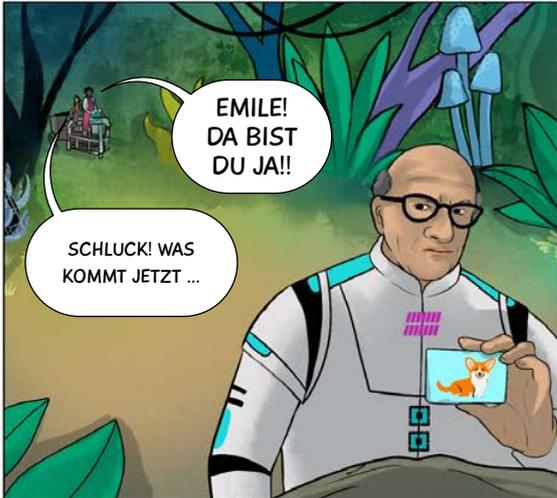


IN DER ZWISCHENZEIT GIBT SICH AUF DER ANDEREN SEITE DES DATENDWALDS DER VON DOC Y GESUCHTE EXPERTE AHNUNGSLOSER KONZENTRATION HIN ...



ICH GLAUBE, WIR SIND NAH DRAN! WEITER, H.I.L.A.R.E.!!

SCEINT SO, ALS OB DU DIE BESSERE FAHRERIN BIST, DOC Y ...



EMILE!
DA BIST
DU JA!!

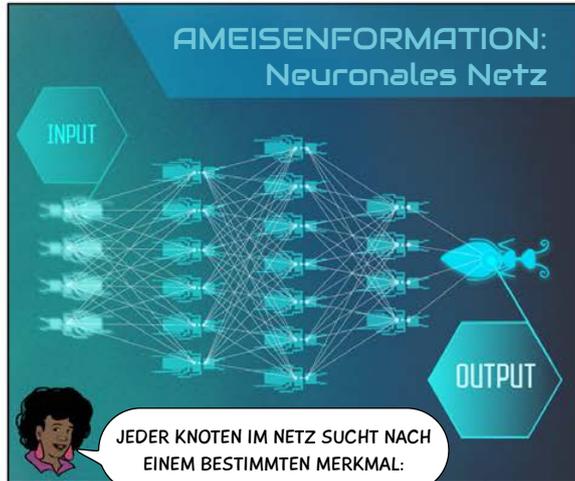
SCHLUCK! WAS KOMMT JETZT ...



BITTE, SEI LEISE! ... ICH BIN MITTEN IM TRAINING ...

ER KOMMT MIR IRGENDWIE BEKANNT VOR ...







ES IST EINE KATZE!



DU WILLST MICH WOHL AUF DEN ARM NEHMEN!!

OH, ALLES IN ORDNUNG, EMILE. NOCH EIN PAAR DURCHLÄUFE, DANN SOLLTE ES BESSER WERDEN ...

ICH BIN ÜBERRASCHT ... SELBST EIN KLEINKIND KÖNNTE DAS BESSER ...



NATÜRLICH KÖNNTE EIN KLEINKIND DAS!

ABER UM EINE KI DAZU ZU BRINGEN, IRGENDWAS ZU ERKENNEN, BRAUCHT ES HUNDERTE VON STUNDEN TRAINING, RIESIGE DATENMENGEN

UND UNGLAUBLICHE MENGEN AN RECHENLEISTUNG ...

DAS IST EINFACH DAS EINMALEINS DES MASCHINELLEN LERNENS!



WARTEN SIE MAL ... SIND SIE EMILE TACH, PROFESSOR TACH?



JA! EMILE WAR EINER DER ERSTEN FORSCHER, DER ETHISCHE UNGLEICHGEWICHTE UNTERSUCHT HAT ... ER IST EIN WAHRER SPEZIALIST FÜR STÖRUNGEN!

WIE GEHT ES DENN MIT DER HAUSARBEIT VORAN, ARI? UND WAS IST DAS DA AN DEINEM ARM?

JA ... DAZU WOLLTE ICH IHRE MEINUNG HÖREN ...



EINE ZIEMLICH MYSTERIÖSE STÖRUNG ... UND SIE SCHEINT ZU WACHSEN ... WIR KORRIGIEREN SIE BESSER, BEVOR SIE IHN GANZ VERSCHLUCKT!



UNSERE KI-ABENTURER HABEN KEINE ZEIT ZU VERLIEREN UND MACHEN SICH AUF DIE SUCHE NACH EINEM WEG ZURÜCK NACH HAUSE ...

IN PLETHOR.A.I. HAT JEDER EINGANG EINEN AUSGANG. UM ALSO DEINEN ZU FINDEN, ARI, MÜSSEN WIR DEINE SCHRITTE ZURÜCKVERFOLGEN.

ARI MUSS IRGENDETWAS GETAN HABEN, DAS ALL DAS AUSGELÖST HAT.

EIN ETHISCHES UNGLEICHGEWICHT, DAS SO GROSS WAR, DASS ES EINE FUNKTIONSSTÖRUNG IN PLETHOR.A.I. VERURSACHT HAT.



ABER WO SOLLEN WIR ANFANGEN? WIR MÜSSEN SO VIELE SCHRITTE ZURÜCKVERFOLGEN!

NUN, ICH DENKE, DIE ERSTE STATION SOLLTE DER DATENSATZ FÜR RESSOURCENARME SPRACHEN SEIN.

WIR WISSEN BEIDE, WIE SELTEN DIE SIND, VIELLEICHT HAT DAS DIE SACHE DURCHEINANDERGEBRACHT?

JA NATÜRLICH, IN MEINER HAUSARBEIT FÜR IHREN KURS! ICH WEISS, DASS SIE ES HASSEN, WENN WIR TERMINE VERPASSEN, PROFESSOR ... VIELLEICHT ...



... ICH BEZWEIFLE, DASS EINE ZU SPÄTE ABGABE PLETHOR.A.I. HÄTTE BESCHÄDIGEN KÖNNEN, ARI ...

WARTEN SIE, WAR ES WIRKLICH ZU SPÄT? WEIL ICH DIE DATEI TECHNISCH GESEHEN HOCHGELADEN HA ...

... KONZENTRIEREN WIR UNS, ARI!



SEHT MAL, DA DRÜBEN!





NUN, WENN ES NICHT DER DATENSATZ WAR, DER ALL DIESE PROBLEME VERURSACHT HAT, DANN IST ES VIELLEICHT DER ONLINE-KURS SELBST?

DER MOOC, ALSO DER MASSIVE OPEN ONLINE COURSE?

ER ENTHÄLT HUNDERTE VON KURSEN, DIE VON ZEHNTAUSENDEN VON STUDIERENDEN AUS DER GANZEN WELT BESUCHT WERDEN!

NUN JA ... WIR KÖNNTEN MIT IHREM KURS ANFANGEN?



WOW ... ICH FRAGE MICH, WAS FÜR VERRÜCKTE ALGORITHMEN DA DRIN LEBEN ...

TECHNISCH GESEHEN IST EIN MOOC KEINE KÜNSTLICHE INTELLIGENZ ... ES IST NUR EINE GETEILTE LERNPLATTFORM.

SICHER, ABER IHR VERGESST ALL DIE KI-LÖSUNGEN, DIE UNS BEIM LERNEN AUF DER PLATTFORM HELFEN ...



HEY, DIESE LEUTE HABEN DIESELBE STÖRUNG WIE ICH!

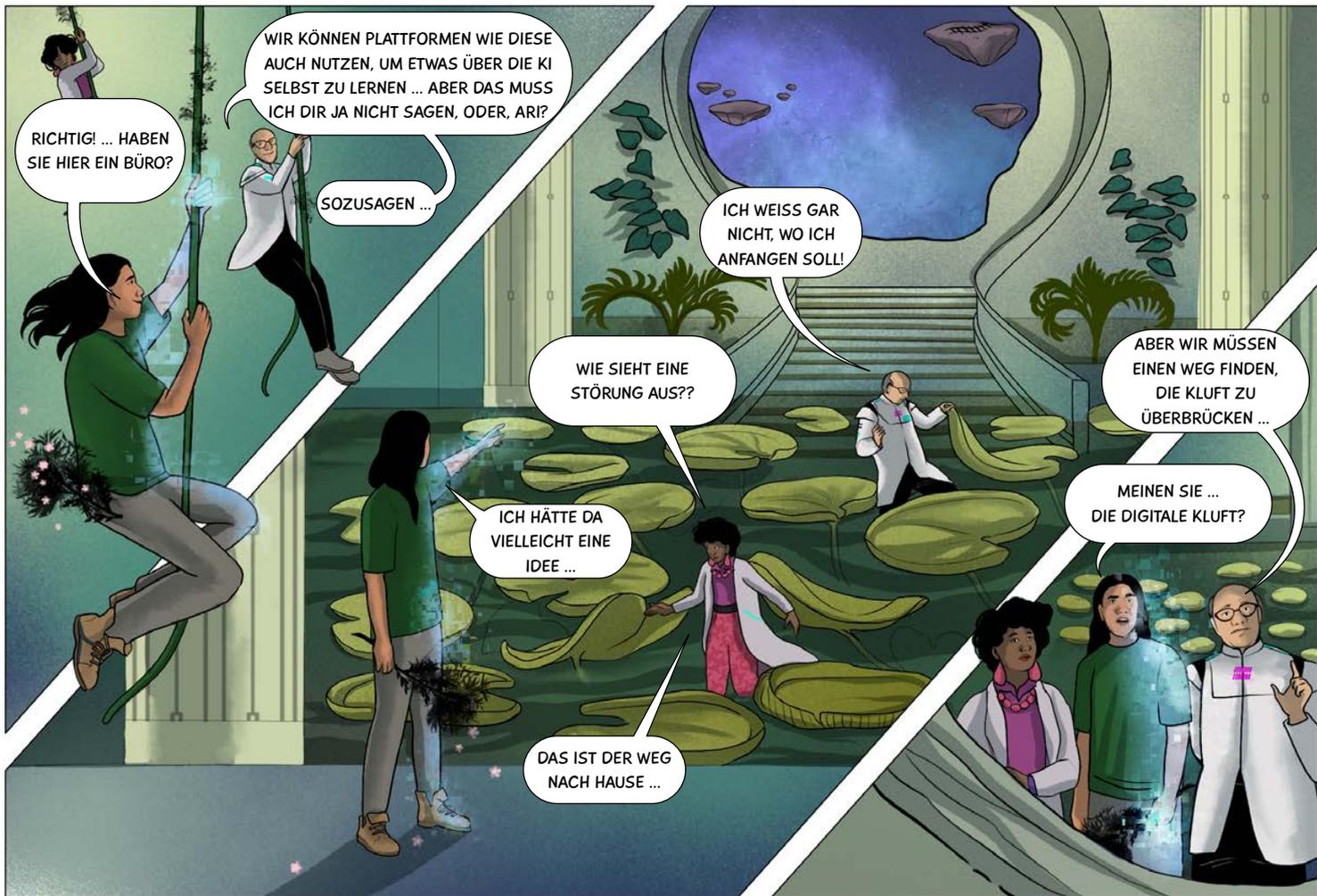


DAS SIND KEINE MENSCHEN, DAS SIND NUR BENUTZERPROFILE. IM GEGENSATZ ZU DIR SIND DIE MENSCHEN, DIE DAHINTER STECKEN, NOCH AUF DER ERDE.

SIE HELFEN DER KI ZU VERSTEHEN, WIE SIE JEMANDEN AM BESTEN UNTERRICHTEN KANN ...

ODER UM MENSCHLICHEN LEHRENDEN ZU HELFEN, SICH AUF DIE BEDÜRFNISSE IHRER STUDIERENDEN EINZUSTELLEN!





RICHTIG! ... HABEN SIE HIER EIN BÜRO?

WIR KÖNNEN PLATTFORMEN WIE DIESE AUCH NUTZEN, UM ETWAS ÜBER DIE KI SELBST ZU LERNEN ... ABER DAS MUSS ICH DIR JA NICHT SAGEN, ODER, ARI?

SOZUSAGEN ...

ICH WEISS GAR NICHT, WO ICH ANFANGEN SOLL!

WIE SIEHT EINE STÖRUNG AUS??

ABER WIR MÜSSEN EINEN WEG FINDEN, DIE KLUFF ZU ÜBERBRÜCKEN ...

ICH HÄTTE DA VIELLEICHT EINE IDEE ...

MEINEN SIE ... DIE DIGITALE KLUFF?

DAS IST DER WEG NACH HAUSE ...



ARI, DAS IST BESTENFALLS SPEKULATION!

WIR WOLLEN KEINE VOREILIGEN SCHLÜSSE ZIEHEN!

ICH DENKE, ICH HABE HEUTE GENUG GELERNT, UM ZU WISSEN, WAS ICH TUE ...



ALLES BEGINNT MIT ZUGÄNLICHKEIT ...

HOFFENTLICH WEISST DU AUCH, WO DAS HINFÜHREN SOLL, ARI!

VERTRAUEN SIE MIR ...



EIN PAAR MINUTEN SPÄTER ...

ALSO ... HAT JEMAND EINEN PLAN B?

OK ÄHM ... ICH HATTE DAS GEFÜHL, DASS DIE BLUMEN ZU EINER ART ZWIG WERDEN WÜR—





AAAH!!!



ICH DENKE, ER HAT
GEWUSST, WAS ER TUT ...
SCHLAUES KIND!

DAS WAR MAL WIRKLICH
EIN ABENTEUER, DOC Y!
DANKE, DASS ICH DABEI
SEIN DURFTE ...



HEY, WEISST DU WAS ARI?
DIE GLASFASERVERBINDUNG
LÄUFT WIEDER!

DAS GLAUBE ICH DIR! DAS
WAR DEFINITIV SCHNELLER
ALS BEIM LETZTEN MAL!



*„Nga Mihl!“ (DANKE!)



OH! SIEHT AUS, ALS HÄTTE ER ES NACH HAUSE GESCHAFFT!



MISSION ERFÜLLT, DOC Y! DIE STÖRUNG IST BEHOBEN!

ICH BIN MIR DA NICHT SO SICHER. DAS SOLLTEST DU DIR ANSEHEN ...



DAS IST MAL EINE RICHTIG GROSSE STÖRUNG!! WENN WIR MIT DER FERTIG WERDEN WOLLEN, MÜSSEN WIR ZUSAMMENARBEITEN ...

MACHEN WIR, EMILE ... UND ICH GLAUBE LANGSAM, DASS ES HIER NICHT NUR UM ZUGÄNGLICHKEIT GEHT ...

ENDE

Glossar

Erweiterte Intelligenz

ist ein Begriff, der die Zusammenarbeit von Menschen und KI beschreibt, um ein von Menschen definiertes oder innerhalb eines sozio-technischen Systems bestehendes Ziel zu erreichen. Die KI kann sowohl virtuell (wie ein System zur Entscheidungsunterstützung) als auch verkörpert (wie ein Fabrikroboter) sein. Mit anderen Worten: Erweiterte Intelligenz entsteht, wenn Menschen und KI zusammenarbeiten und sich als effizienter erweisen als jede Partei für sich allein.

Chatbot ist eine Art virtueller KI-Assistent (ein Computerprogramm), der eine Unterhaltung mit einem Menschen nachahmt, in der Regel über das Internet oder eine virtuelle Schnittstelle, über reinen Text oder Text und Sprache. Chatbots

können den Nutzerinnen und Nutzern hilfreiche Informationen liefern oder sie zu geeigneten Ressourcen weiterleiten und spielen eine Schlüsselrolle bei der Erfüllung der Anforderungen der Nutzerinnen und Nutzern.

Einwahlverbindung

ist eine Internetverbindung, die über eine normale Telefonleitung hergestellt wird. Wenn die Telefonleitung mit einem Modem verbunden und so konfiguriert ist, dass sie eine bestimmte Nummer wählt, erhalten die Nutzenden Zugang zum Internet. Die Einwahlverbindung ist die langsamste Form der Internetverbindung und wird heute hauptsächlich in Regionen genutzt, in denen der Bau von Breitbandleitungen wirtschaftlich nicht möglich oder rentabel ist.

Digitale Kluft beschreibt die Kluft zwischen Einzelpersonen,

Haushalten, Unternehmen oder geografischen Regionen in Bezug auf den Zugang zu **a)** KI-Forschung, **b)** Wissen, Bildung und Fachkräften, **c)** Schulungsdaten und **d)** Konnektivität und Hardware. Anders ausgedrückt: Die digitale Kluft ist es, die auf globaler Ebene gleiche Wettbewerbsbedingungen in der KI-Technologie verhindert, da den Benachteiligten die Instrumente fehlen, um mit hochentwickelten Akteuren im KI-Ökosystem konkurrieren und zusammenarbeiten zu können.

Epoche (oder auch «stochastischer Gradientendurchlauf»)

ist eine Art Hyperparameter beim maschinellen Lernen, der angibt, wie oft der Lernalgorithmus den gesamten Trainingsdatensatz durchläuft und dabei seine Gewichtungen an jede Probe im Satz anpasst. Die meisten KI-

Modelle durchlaufen während ihres Trainings viele Epochen (manchmal über 1000), da jede Epoche den Programmierinnen und Programmierern (zumindest bis zu einem gewissen Punkt) hilft, die Fehler im Modell zu minimieren oder seine Leistung zu optimieren.

Glasfasertechnik ist eine Technologie zur Übertragung von Informationen in Form von Lichtimpulsen durch Glas- oder Kunststofffasern über große Entfernungen. Da Glasfaserkabel Datensignale in Form von Licht und nicht auf herkömmliche Weise in Form von elektrischen Signalen übertragen, können die Informationen schneller, über größere Entfernungen und ohne das Risiko elektromagnetischer Störungen (z.B. durch Gewitter oder starke Winde) übertragen werden. Glasfaserkabel sind ein wichtiger Teil der Infrastruktur moderner KI-Praxis.

Glossar-Begriffe

Maschinelles Lernen

wird auch als „datengesteuerte“ oder „Bottom-Up“-KI bezeichnet. Maschinelles Lernen ist eine sich anpassende Art von algorithmischem Prozess, der es Computern ermöglicht, anhand großer Datenmengen aus Erfahrungen, anhand von Beispielen oder anhand von Ähnlichkeiten zu lernen, wobei die Lernprozesse die Leistung des Systems mit der Zeit verbessern. Viele Verfahren des maschinellen Lernens sind eng verwandt mit den künstlichen neuronalen Netzen (KNN) und orientieren sich an der Struktur des menschlichen Gehirns. KNNs verwenden sehr einfache und stark vernetzte Prozessoren (Neuronen genannt) mit Verbindungen, die Signale von einem Neuron zum anderen weiterleiten. Der Lernprozess findet statt, wenn

diese Gewichtungen durch die Einspeisung von Trainingsdaten angepasst werden. Schließlich werden die „trainierten“ Regeln, die für die korrekte Lösung eines bestimmten Problems erforderlich sind (z.B. Klassifizierung oder Mustererkennung), gespeichert. Sie werden bei der Entscheidungsfindung verwendet, wenn neue Daten in das System eingespeist werden.

Massive Open Online Courses (MOOCs)

sind dezentrale virtuelle Lernplattformen, die Studierenden weltweit zur Verfügung stehen. MOOCs sind einer der wichtigsten Schlüssel zum lebenslangen Lernen. Sie ermöglichen es Studierenden und anderen Personen, sich in Online-Kursen mit hochkarätigen Fachleuten aus den Natur- und Geisteswissenschaften

auszutauschen und von ihnen zu lernen. MOOCs gibt es in vielen Formen, von vollständig selbstgesteuerten Plattformen, auf denen Studierende den Kurs in ihrem eigenen Tempo absolvieren können, bis hin zu akkreditierten professionellen Studiengängen, die es den Studierenden ermöglichen, sich formal weiterzubilden.

Moravec's Paradox

beschreibt die überraschende Diskrepanz zwischen den intellektuellen Fähigkeiten von Menschen und fortgeschrittenen KI-Systemen, wobei KI-Systeme in einer Reihe von traditionell „schwierigen“ menschlichen Bereichen (z.B. statistisches Denken, Mustererkennung oder Damespiel) hervorragende Leistungen erbringen. Gleichzeitig schneiden sie in vielen Bereichen, die die meisten Menschen leicht beherrschen, schlecht

ab: Wahrnehmung, Mobilität, gesunder Menschenverstand und Werturteile, um nur einige zu nennen.

Symbolische KI auch als „Top-Down“-KI oder „Expertensysteme“ bezeichnet, ist eine Art algorithmischer Prozess, der expliziten und transparenten Regeln folgt, die dem System von menschlichen Programmierern vorgegeben werden, um die Lösung eines Problems zu berechnen. Symbolische KI ist der ursprüngliche und älteste Ansatz der Künstlichen Intelligenz. Er wird auch heute noch in vielen Anwendungen eingesetzt, insbesondere in Situationen, in denen ein hohes Maß an menschlicher Kontrolle wünschenswert ist, wie z.B. bei fahrerlosen Fahrzeugen.

Schlussfolgerung

Menschen und KI können viel voneinander lernen. Das bedeutet nicht nur, dass das, *was wir lernen*, an die rasanten technologischen Veränderungen angepasst werden muss, die heute um uns herum stattfinden, sondern auch, dass die Art und Weise, *wie wir lernen* innovative und nützliche KI-Tools einbeziehen sollte. Durch die Kombination von KI und Bildung können wir globale Klassenzimmer schaffen, Gleichgesinnte zusammenbringen und auch der KI selbst beibringen, wie sie uns besser lehren kann.

Wenn wir jedoch nicht dafür sorgen, dass jeder von uns die *gleichen* Chancen hat, von den Möglichkeiten der Künstlichen Intelligenz zu profitieren – zum Beispiel durch die Erstellung offener, inklusiver Datensätze oder die Überwindung von Sprachbarrieren in der KI-Forschung und der Sprachverarbeitung –, werden diese technologischen Fortschritte nur dazu führen, vorhandene globale Ungleichheiten weiter zu vergrößern. In diesem Sinne müssen politische Maßnahmen auf nationaler und globaler Ebene eingeführt bzw. überarbeitet werden, um eine KI-Entwicklung und -Nutzung sicherzustellen, die den Menschen nicht schadet und auch nicht auf schädliche Weise eingesetzt werden kann. ■



Episode II

Maryam und der Deep Dive ins Datenbecken

Konzeption und Skript: Dr. Katherine Evans

Illustration: Cassandra Okwaniuzor Mark



Einleitung

Künstliche Intelligenz entwickelt sich rasch zu einem entscheidenden Element bei der Förderung der menschlichen Entwicklung und der Schaffung inklusiver Wissensgesellschaften. In einer zunehmend globalen und vernetzten Welt ermöglicht es uns KI, unsere Umwelt besser zu verstehen und in ihr zu interagieren, sowohl auf Ebene der einzelnen Nutzerinnen und Nutzer als auch auf kollektiver Ebene im öffentlichen und privaten Sektor. KI kann uns dabei helfen, bessere, fundiertere und präzisere Entscheidungen in allen Bereichen zu treffen: Sie zeigt uns Muster auf, die wir vielleicht übersehen haben, findet Zusammenhänge, die wir vielleicht nicht wahrgenommen haben, und gibt uns innovative Empfehlungen an die Hand, die wir uns selbst nie hätten ausdenken können.

Wenn wir jedoch nicht darauf achten, wie sich KI-Technologien weltweit ausbreiten und weiterentwickeln, laufen wir Gefahr, eine Reihe großer ethischer Herausforderungen zu übersehen, die grundlegende Menschenrechte bedrohen oder gar außer Kraft setzen. In der Tat stellen Themen wie verstärkte Überwachung, Data Mining und Profiling sowie algorithmische Verzerrung und automatisierte Entscheidungsfindung neue Risiken für die Rechte auf Privatsphäre und Nichtdiskriminierung dar, die jedem von uns zustehen. Noch schlimmer ist, dass diese Risiken zwar eine Bedrohung für die Menschenrechte im Allgemeinen darstellen, aber die Rechte von Frauen, People of Colour und anderen gefährdeten Minderheiten unverhältnismäßig stark zu beeinträchtigen drohen.

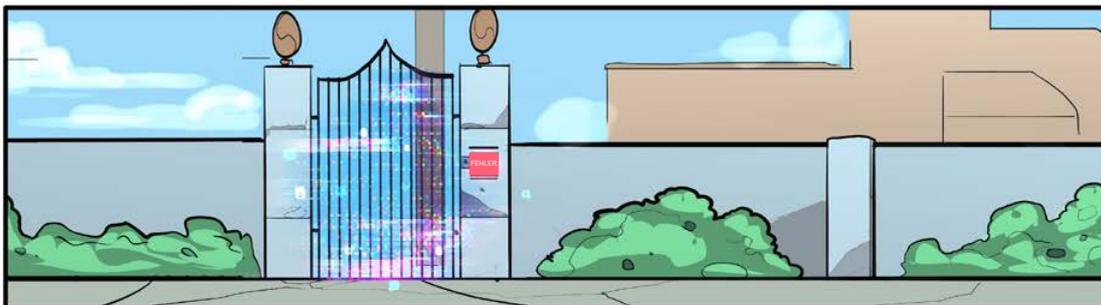
In gewisser Weise liegt das Problem im Spannungsfeld zwischen automatisierter Entscheidungsfindung, Daten und Vorurteilen aus der Vergangenheit begründet. Es muss sichergestellt werden, dass die moralischen Fehler der Vergangenheit nicht die Entscheidungen der Gegenwart beeinflussen: durch die Verwendung verzerrter und veralteter Datensätze, die eine Welt repräsentieren, in der wir nicht länger leben wollen, oder durch das blinde Befolgen undurchsichtiger Empfehlungen von KI-Tools, die mit solchen Daten trainiert wurden.

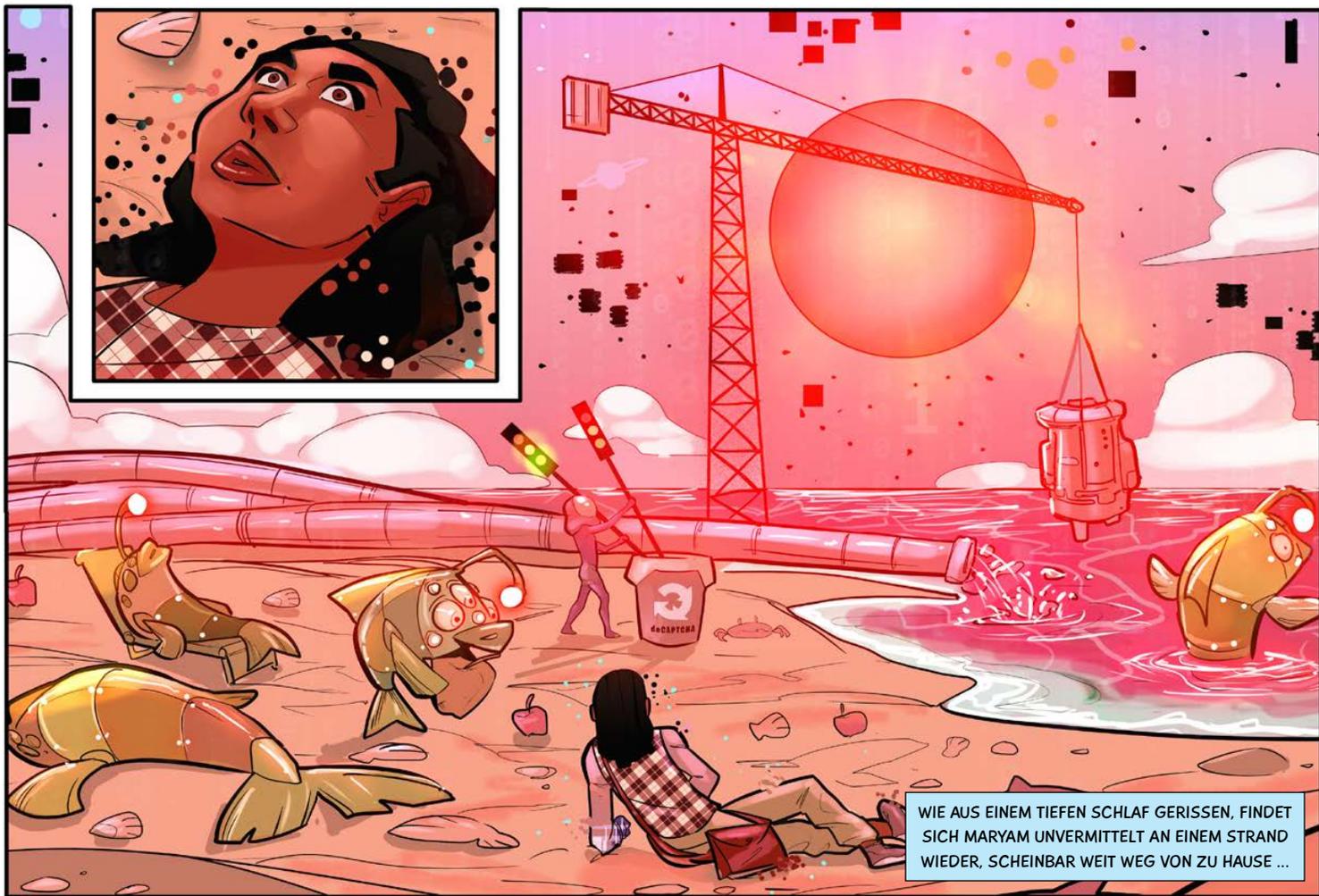
Gleichzeitig liegt ein weiteres Problem darin, dass wir selbst oft nur eingeschränkt dazu in der Lage sind, die Vielfalt der Menschheit zu verstehen und angemessen darzustellen. Bei der Entwicklung von KI-Systemen muss der Inklusionsgedanke im Vordergrund stehen, damit KI-Tools die Möglichkeit haben, die vielen Facetten der Menschheit in ihrer ganzen Breite zu erfassen und daraus zu lernen – nicht nur eine begrenzte Auswahl bekannter, ähnlicher und leicht verfügbarer Gesichter.

Wenn wir diese Probleme nicht angehen, riskieren wir die Entwicklung von KI-Anwendungen, die nicht allen Menschen gleichermaßen und angemessen helfen können. Eine Welt, in der die KI perfekt auf die einen eingeht, die anderen aber kaum erkennt; eine Welt, in der einige von uns praktisch unsichtbar sind ... ■

AN EINEM SONNIGEN MORGEN IN LAGOS, NIGERIA, IST MARYAM SPÄT DRAN UM DIE CHANCE IHRES LEBENS ZU ERGREIFEN ...







WIE AUS EINEM TIEFEN SCHLAF GERISSEN, FINDET SICH MARYAM UNVERMITTELT AN EINEM STRAND WIEDER, SCHEINBAR WEIT WEG VON ZU HAUSE ...





ENTSCULDIGUNG?!

DER FISCH IST EIN
KLASSIFIZIERER. ER
VERARBEITET KEINE
NATÜRLICHE SPRACHE.

DAS ÜBERSTEIGT
BEI WEITEM SEINE
FÄHIGKEITEN—



—WOVON IN
ALLER WELT
REDEST DU DA?!



WIR SIND NICHT IN
DEINER WELT, WIR SIND
IN PLETHOR.A.I.!

DEIN NAME
IST MARYAM,
RICHTIG?

RICHTIG ... ?

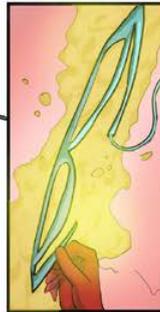


MEIN NAME IST DOC
Y, UND ICH KANN DIR
HELFFEN. MÖCHTEST DU
MIT MIR KOMMEN?



WILLKOMMEN IM
DATENBECKEN, DEM
HOTSPOT IN PLETHOR.A.I.
FÜR MASCHINELLE
BILDERVERARBEITUNG ...

UND HEIMAT DER
VIELEN KLASSTIFIE-
RUNGSALGORITHMEN
DER GALAXIE!





AM ANFANG LERNEN
SIE VON BEISPIELEN,
DIE MENSCHEN IHNEN
GEBEN.

ABER IRGENDWANN
HABEN SIE GENUG
GELERNT, UM SELBST
DINGE ZU ERKENNEN.



SIE ERKENNEN
MENSCHLICHE
GESICHTER,
SAGST DU?

WENN SIE SO
KONZIPERT SIND, DASS
SIE SIE ERKENNEN SOLLEN,
HÄNGT ES VON IHREN
TRAININGSDATEN AB—



—DAS ERINNERT
MICH AN ETWAS ...



VOR DEN AUGEN DER ANDEREN RENNFAHRER VERBORGEN,
HECKEN DOC Y UND MARYAM EINEN PLAN AUS ...



DU MEINST, ICH SOLL
MICH IN DAS RENNEN
SCHLEICHEN?

ICH GEBE ZU, ES IST SEHR
UNKONVENTIONELL, ABER HIER
IST MEINE HYPOTHESE ...



DAMIT KLASSIFIZIERER ETWAS ERKENNEN KÖNNEN,
MÜSSEN SIE TRAINIERT WERDEN. DIE MENSCHEN
MÜSSEN IHNEN SAGEN, WIE SIE DIE DINGE IN DER
WELT IDENTIFIZIEREN UND WIE SIE SIE NENNEN
SOLLEN. ZUM BEISPIEL ...



WAS IST
DAS?

EIN APFEL ...

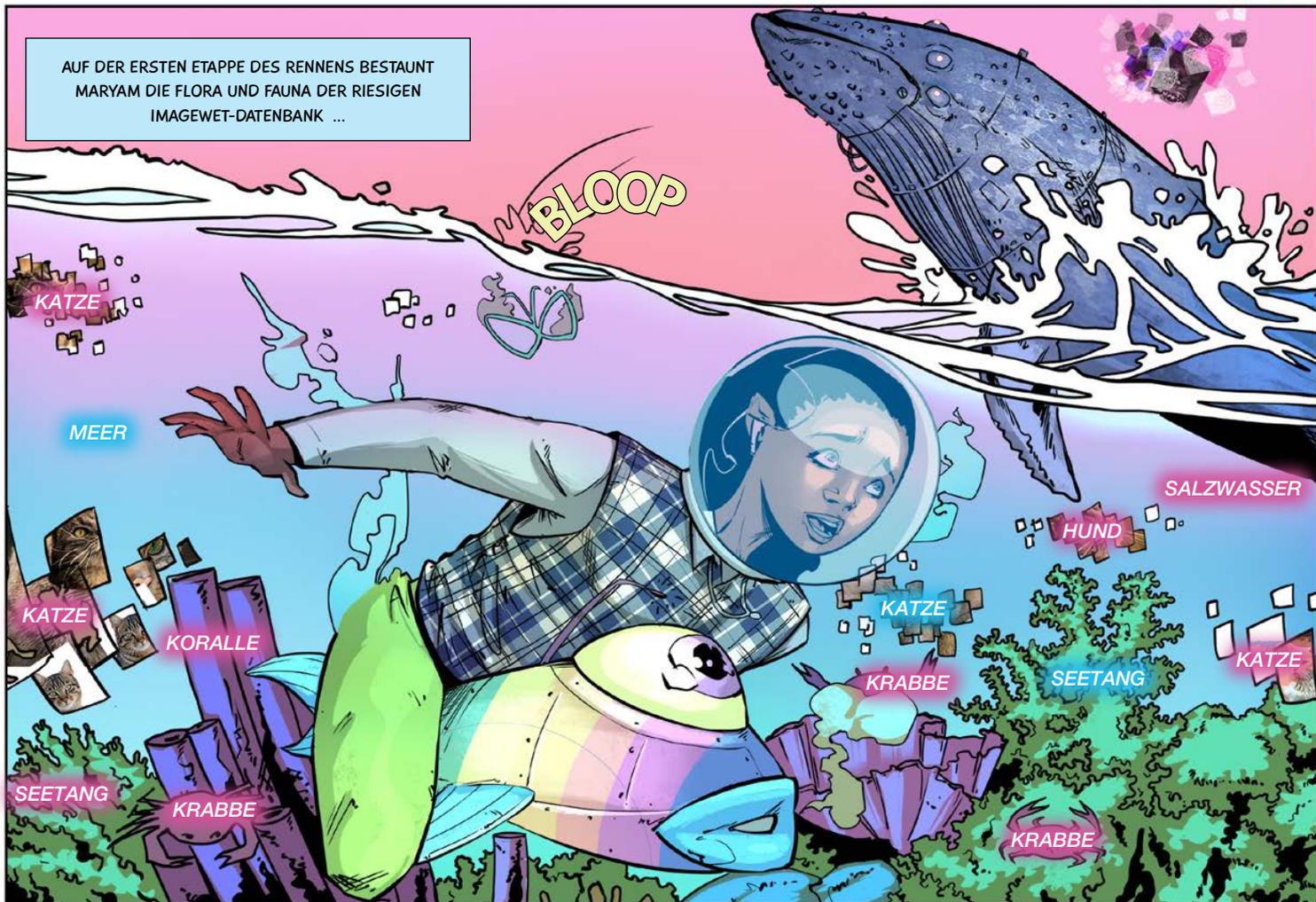
ABER SEHEN ALLE
ÄPFEL GLEICH AUS?







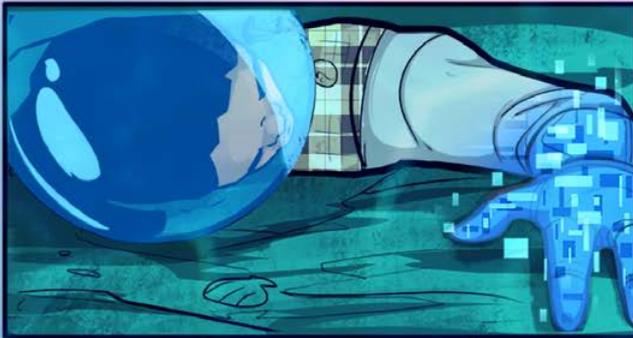
AUF DER ERSTEN ETAPPE DES RENNENS BESTAUNT MARYAM DIE FLORA UND FAUNA DER RIESIGEN IMAGEWET-DATENBANK ...





ALS MARYAM IN DAS HERZ DER DATENBANK EINDRINGT,
TRIFFT SIE AUF EINIGE UNBEKANNTE GESICHTER ...





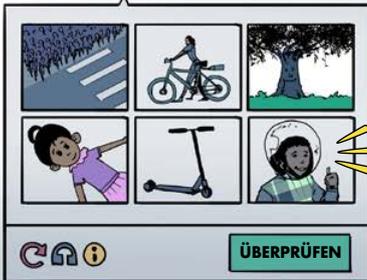


IRGENDWO AN DER PAZIFIKKÜSTE ...

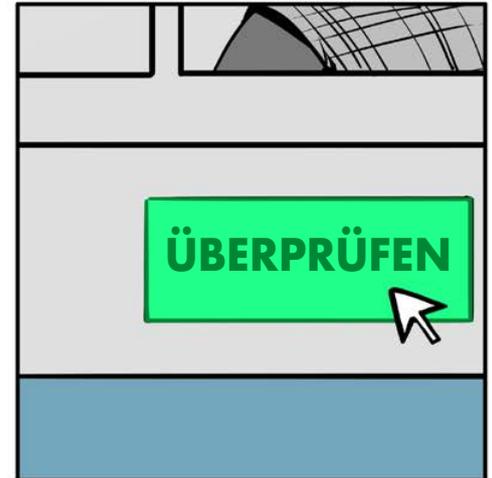


Bitte kennzeichnen Sie alle Bilder, auf denen Menschen zu sehen sind.

Ich bin ein Mensch



KOMISCH ...
DIESE TAUCHERMASSE
SIEHT NICHT RICHTIG AUS ...









NACHDEM SIE DIE LETZTE ETAPPE DES DATENBANKRENNENS
GERADE SO HINTER SICH GEBRACHT HAT, FINDET SICH MARYAM IN
DER HOCHSEE DES DATENBECKENS WIEDER ...



HEY! DU KANNST MICH
JETZT SEHEN! ABER HEISST
DAS NICHT, DASS MEINE
STÖRUNG BEHOBEN IST?



WAS JETZT?!?!

BURBLE

BURBLE



SIEHT SO AUS, ALS OB DEINE
STÖRUNG KOMPLEXER IST, ALS WIR
DACHTEN! BEEILUNG, WIR HABEN
KEINE ZEIT ZU VERLIEREN!

ABER WAS
IST MIT ...

... KEINE SORGE, ICH
BIN MIR SICHER, DASS
IHR EUCH WIEDERSEHEN
WERDET ...



MARYAM STEIGT IN DEN BAUCH
DES SCHIFFES HINAB, UNSICHER,
WAS SIE ERWARTET ...

AH! HABE ICH NICHT
ERWÄHNT, DASS ICH
EINE CO-PILOTIN
EINGESTELLT HABE?

ES IST MIR EIN
VERGNÜGEN, DICH ENDLICH
KENNENZULERNEN, MARYAM!

ES IST NICHT ALLTÄGLICH, EINE
JUNGE FRAU MIT EINEM PROFIL WIE
DEINEM ZU FINDEN!



ENTSCULDIGEN
SIE, ABER WOHER
KENNEN SIE
MEINEN NAMEN?



VON DEINER
BEWERBUNG
NATÜRLICH!

ICH HABE DICH SELBST FÜR
DAS PRAKTIKUM
AUSGEWÄHLT, ABER DANN ...

... ES SCHEINT, DASS
RUTH EIN TECHNISCHES
PROBLEM MIT IHREM
BEWERBUNGSPROGRAMM
HAT

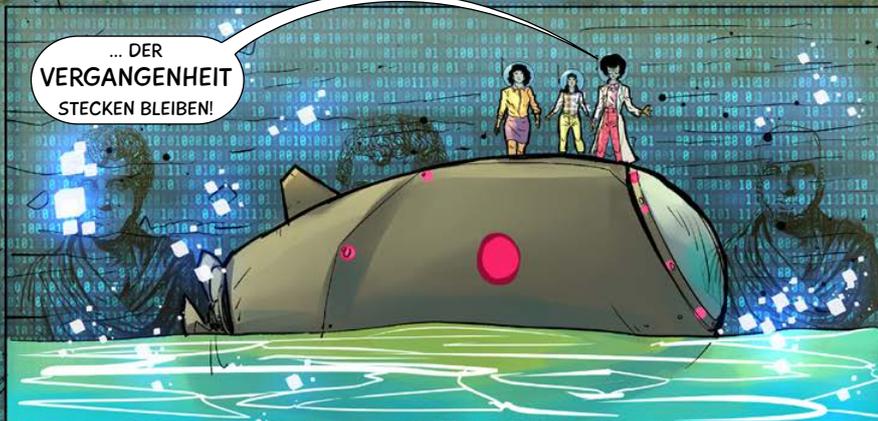
UND DEINE BEWERBUNG
VOM AUTOMATISCHEN
SYSTEM NICHT
AUSGEWÄHLT WURDE ...

ABER WIE
KONNTE DAS
PASSIEREN?

DAS IST GENAU DAS,
WAS WIR HERAUSFINDEN
WERDEN! HALTET EUCH
FEST!



... DER
VERGANGENHEIT
STECKEN BLEIBEN!

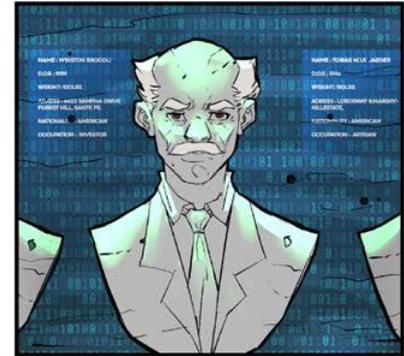


WILLKOMMEN IN DEN
DATENMINEN, DER HEIMAT
EINIGER DER ÄLTESTEN
DATENSÄTZE IN GANZ
PLETHOR.A.I. WENN IHR MIR
BITTE FOLGEN WÜRDET ...



TÄGLICH FLIESSEN ALLE MÖGLICHEN NEUEN DATEN IN PLETHOR.A.I. EIN. SIE WERDEN VERARBEITET, GESPEICHERT ...

ABER WIR HABEN AUCH DATEN DARÜBER, WIE DIE DINGE VOR KI WAREN. DAS IST ES, WAS IHR HIER SEHT.



KI KANN DIESE DATEN NUTZEN, UM DIE WELT ZU INTERPRETIEREN UND ENTSCHEIDUNGEN ZU TREFFEN. AUSSER ...

ICH HABE DAS GEFÜHL, DASS ES EINEN HAKEN GIBT?

WENN WIR WISSEN WOLLEN, WIE EIN IDEALER ASTRONAUT AUSSEHT ...

WIE KÖNNTE MAN SICH, AUSGEHEND VON DEM, WAS DU HIER SIEHST, EINEN IDEALEN ASTRONAUTEN VORSTELLEN?

ASTRONAUTS



VIELLEICHT IN ETWA SO WIE DIE DA ...?

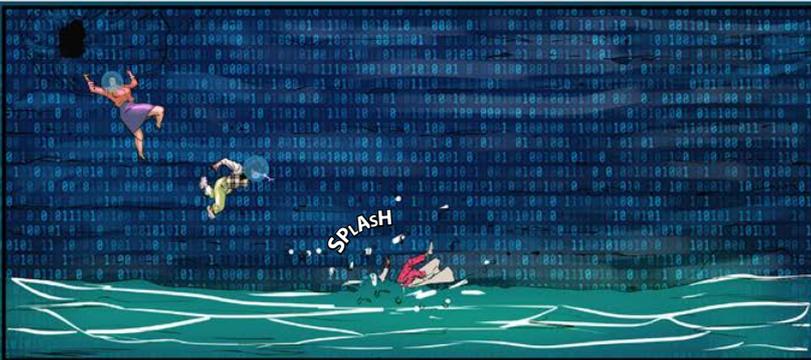
GENAU! AUCH WENN VIELE VERSCHIEDENE PERSONEN DIE ARBEIT ERLEDIGEN KÖNNTEN!

DAS IST EIN TEIL DESSEN, WAS WIR ALS „ALGORITHMISCHE VERZERRUNG“ BEZEICHNEN, MARYAM, UND DEINE STÖRUNG IST DER KERN DAVON.

UM DAS ZU BEHEBEN, MÜSSEN WIR ALSO EIN EI STEHLEN ...

HÄ?





ICH KANN ES NICHT
GLAUBEN, ABER DAS
IST WIRKLICH UNSER
BEWERBUNGSPROGRAMM!

GROSSARTIG!
DANN SOLLTE
ES GANZ
EINFACH SEIN ...



MEINE VERMUTUNG IST,
DASS UNSER KLEINER
FREUND HIER EIN PAAR
ERNSTHAFT VERZERRTE
DATEN GENASCHT HAT ...



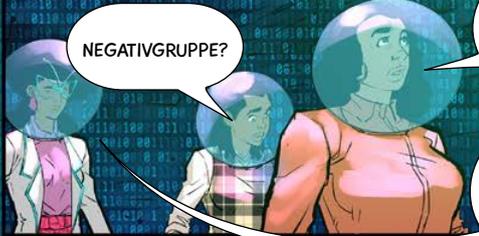
UND DASS IRGENDWO DA
DRIN EINE WINZIGE VERSION
VON DIR IN DIE NEGATIVGRUPPE
GERUTSCHT IST ...



NEGATIVGRUPPE?

DIE GRUPE VON
MENSCHEN, DIE DIE KRABBE
ALS NICHT EINSTELLBAR
BETRACHTET ...

SETZ DICH IN DIE
POSITIVGRUPPE , UND
DEINE STÖRUNG SOLLTE
BEHOBEN SEIN!



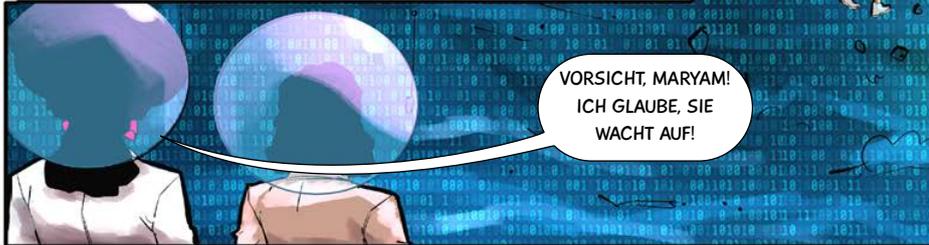
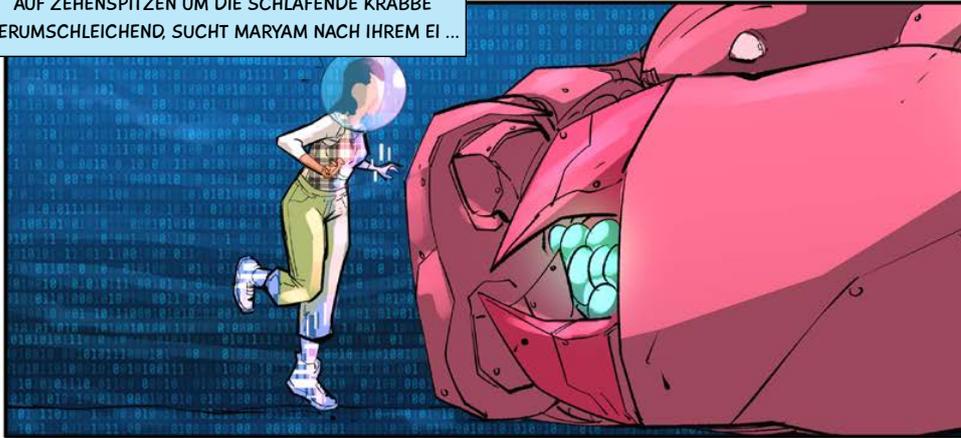
UND WIE GENAU
SOLL ICH DAS
MACHEN?

FRAG MICH NICHT,
ICH BIN KEINE
PROGRAMMIERERIN!

DAS IST SEHR
BERUHIGEND ...



AUF ZEHENSPIZZEN UM DIE SCHLAFENDE KRABBE
HERUMSCHLEICHEND, SUCHT MARYAM NACH IHREM EI ...



ALLES KLAR, JETZT
GEHT'S LOS ...!

BRAVO,
MARYAM!

DU HAST EIN ETHISCHES
UNGLEICHGEWICHT
KORRIGIERT! DER PLANET FÜGT
NEUE DATEN HINZU, UM DAS,
WAS DU GERADE GETAN HAST,
UMZUSETZEN!

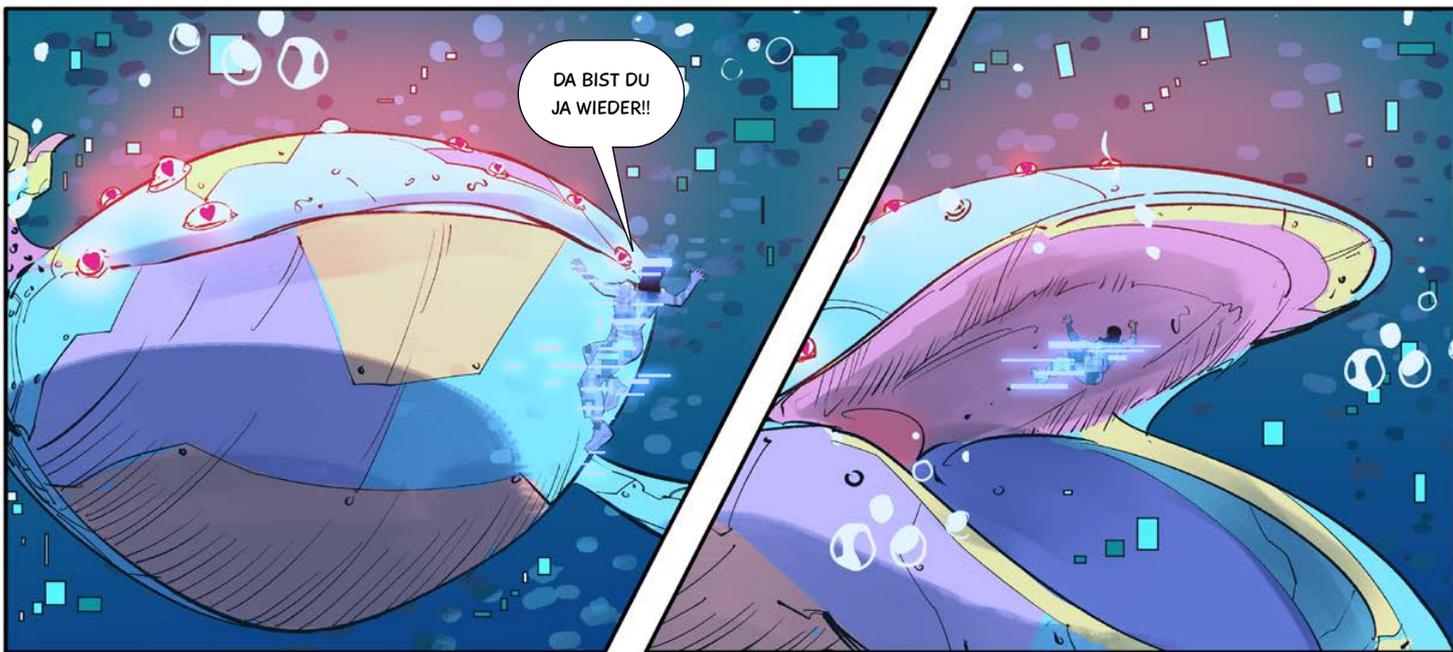
ASTRONAUTS

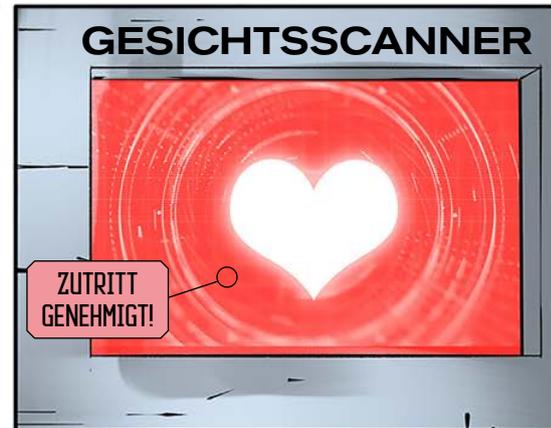


HEY IHR BEIDE! ICH
GLAUBE, WIR HABEN
EIN PROBLEM!

OBWOHL DIE ENDLOSEN HÖHLENGÄNGE DER
DATENMINEN IMMER MEHR REPRÄSENTATIVE GESICHTER
ZEIGEN, WIRD MARYAMS JUBEL BALD UNTERBROCHEN...

SPLAAAAAASH





Glossar

Algorithmische Verzerrung ist ein Oberbegriff, der die Art und Weise beschreibt, in der bestimmte KI-Systeme Ergebnisse, Entscheidungen oder Empfehlungen erzeugen können, die eine Tendenz oder ein Vorurteil zugunsten oder zulasten einer Person oder Gruppe aufweisen, insbesondere in einer Weise, die als unfair angesehen wird. Die Gründe für diese Verzerrungen sind vielfältig und hängen mit der Repräsentativität des Datensatzes, der Größe des Datensatzes, der Gewichtung des KI-Modells selbst und sogar mit bewussten und unbewussten Neigungen der Programmierenden des Systems zusammen. Auch wenn

die meisten gut konzipierten KI-Systeme frei von algorithmischer Verzerrung sind, ist es unter KI-Fachleuten nach wie vor ein Problem, sicherzustellen, dass KI-Systeme immer fair bleiben.

Klassifizierungsalgorithmen sind im Fall des überwachten maschinellen Lernens Modelle, die aus den eingegebenen Trainingsdaten lernen. Damit können sie die Wahrscheinlichkeit vorher-sagen, dass die Stichproben aus einem neuen Datensatz in eine der vorgegebenen Klassifizierungskategorien (auch als „Teilpopulationen“ bezeichnet) fallen. Mit anderen Worten: Das Modell lernt zunächst Kategorien und deren Assoziationen aus den Kennzeichnungen, die ihm

menschliche Programmierer während des Trainings geben, und nutzt dann das Gelernte, um neue Daten in dieselben Kategorien einzuordnen.

Ein **Datensatz** wird manchmal auch als „Datenbank“ bezeichnet. Ein Datensatz ist eine Sammlung strukturierter Daten. Mit anderen Worten: eine Sammlung von Fakten, Messungen, Beobachtungen, Bildern, Texten, Audio- oder Videodaten oder Darstellungen davon. Datensätze sind ein wesentlicher Bestandteil Künstlicher Intelligenz. Ohne Datensätze wäre die KI nicht in der Lage, zu lernen, zu erkennen, zu kategorisieren, zu sprechen, zuzuhören oder zu interpretieren, um nur

einige Beispiele zu nennen. Datensätze können groß oder klein, öffentlich oder privat, ausgewogen oder verzerrt sein, aber unabhängig von den Details bilden Datensätze das Rückgrat für maschinelles Bottom-Up-Lernen oder für das, was wir zu Recht als datengesteuerte KI bezeichnen.

Personalbeschaffungstools beziehen sich auf den Einsatz von Künstlicher Intelligenz, um einen Teil des Einstellungsverfahrens zu automatisieren. Dies kann die automatische Planung von Vorstellungsgesprächen mit Bewerberinnen und Bewerbern, deren Vorauswahl und Matching (durch Analyse und Empfehlung) sowie Chatbots

Glossar

oder KI-gestützte Telefon- oder Videointerviews umfassen. Heutzutage sind die meisten Einstellungsverfahren nicht vollständig automatisiert, sondern dienen als Entscheidungshilfe für die Personen, die für dieses Verfahren zuständig sind.

Schlussfolgerung

Künstliche Intelligenz dient immer mehr dazu, den Menschen zu helfen, sich in einer immer komplexeren und schnelllebigeren Welt zurechtzufinden. Doch welche Art von Hilfe kann sie bieten, wenn ihr Design und ihre Struktur automatisch die einen zum Nachteil der anderen bevorzugen?

Um dies zu verhindern, müssen Maßnahmen ergriffen werden, die gewährleisten, dass die Menschenrechte aller in vollem Umfang und in angemessener Weise geschützt werden. Ein Teil dieser Lösung beginnt auf Designebene, indem überprüft wird, ob die Datensätze vollständig, inklusiv und sensibel für die Schwachstellen und Besonderheiten von

jedem von uns sind. Ein weiterer Schritt ist die Entwicklung einer angemessenen Überwachung von KI-Systemen: Wir müssen uns für die Entwicklung und Einhaltung von auf Menschenrechten basierenden internationalen ethischen Standards einsetzen, die dazu beitragen, dass KI-Tools wirklich allen zugutekommen.

Diese Bemühungen können dazu beitragen, dass Voreingenommenheit und Diskriminierung bei der Entwicklung und Nutzung von KI-Systemen verhindert werden und dass die Zukunft der KI eine Zukunft ist, wie wir alle sie uns wünschen. ■



Episode III

Shirin auf der Spur des Empfehlungsgebers

Konzeption und Skript: Dr. Katherine Evans

Illustration: Asma Kraiem



Einleitung

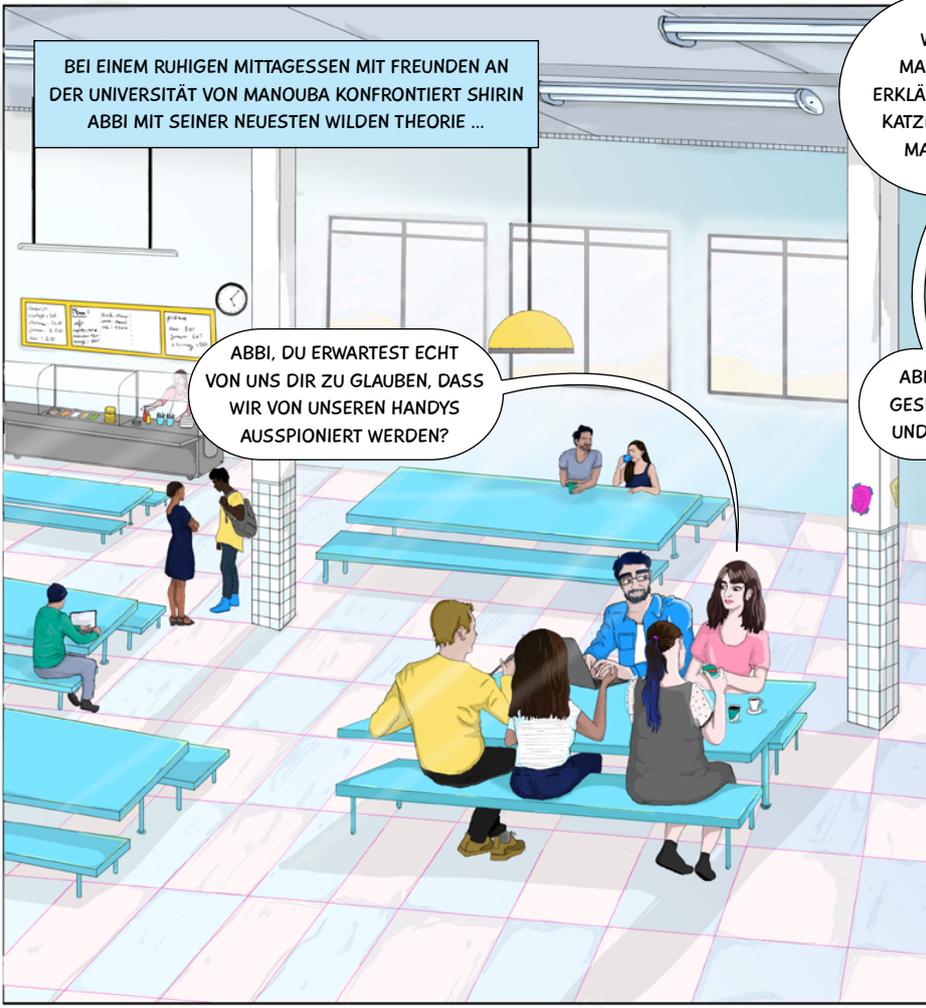
Jeden Tag nimmt KI Einfluss auf das Leben von Millionen von Menschen auf der ganzen Welt. Sie beeinflusst Erfahrungen und Möglichkeiten von Menschen und prägt ihr Verständnis der Welt um sie herum. Wir nutzen Technologien, um mit unseren Freundinnen und Freunden in Kontakt zu treten, neue Interessen zu finden oder mehr über die Themen zu erfahren, die uns am Herzen liegen. Wir entscheiden, welche Teile unseres Lebens wir online mit anderen teilen, und können KI sogar nutzen, um uns selbst zu verstehen oder neu zu erfinden. In diesem Sinne bietet uns KI eine noch nie dagewesene Chance für persönliches und kollektives Wachstum: zum einen, indem sie die Welt zusammenbringt, zum anderen, indem sie die unglaubliche Vielfalt menschlicher Erfahrungen, Meinungen und Fachkenntnisse sichtbar macht.

Doch so wie die KI uns ein Fenster bietet, durch das wir die Welt sehen können, so bietet sie auch jedem von uns eine einzigartige Vision für diese Welt. Wenn wir mit Technologien wie sozialen Medien oder Online-Inhaltsplattformen interagieren, gestaltet KI unsere Erfahrungen im Hintergrund: Sie zeigt uns die Inhalte, die uns am ehesten gefallen, bringt uns mit anderen Nutzerinnen und Nutzern zusammen, mit denen wir vielleicht etwas gemeinsam haben, und entfernt gelegentlich Ideen, die unsere Überzeugungen infrage stellen oder uns Unbehagen bereiten könnten. Auf diese Weise verändert KI

den Blick, den wir durch unser Fenster auf die Welt werfen, und zeigt uns, was uns am meisten fesselt.

Woher weiß die KI, was jeder von uns am liebsten sehen würde? Zum Teil weiß sie es, weil wir es ihr sagen. Jedes Mal, wenn wir ein Video liken, einen Beitrag kommentieren, einen neuen Freund hinzufügen oder jemandem folgen, sagen wir der Künstlichen Intelligenz: „Zeig mir bitte mehr davon“. Gleichzeitig untersucht die KI auch, wie wir unsere Zeit online verbringen: welche Anzeigen wir angeklickt haben, welche Videos wir uns angesehen haben, welche Seiten wir besucht haben und wie lange wir dort waren. KI zeichnet all diese Informationen als Daten auf und nutzt diese, um uns neue Dinge vorzuschlagen, die wir lesen, anschauen, kaufen oder tun können. Genauso wie die KI uns ein personalisiertes Bild von der Welt vermittelt, hat sie auch ein sehr persönliches Bild von jedem von uns, das ausschließlich durch die von ihr gesammelten Daten definiert wird.

Aber was passiert, wenn wir nicht damit einverstanden sind, wie unsere Daten uns definiert haben? Was sollte man tun, wenn man nur noch das sieht, was einen sowieso interessiert, nicht aber, was wirklich wichtig ist oder darüber hinaus geht? Wie können wir in unserem Fenster zur Welt Wahrheit von Fiktion trennen, und wie kann KI den Unterschied erkennen? Die Antworten können sehr persönlich sein ... ■



BEI EINEM RUHIGEN MITTAGESSEN MIT FREUNDEN AN DER UNIVERSITÄT VON MANOUBA KONFRONTIERT SHIRIN ABBI MIT SEINER NEUESTEN WILDEN THEORIE ...

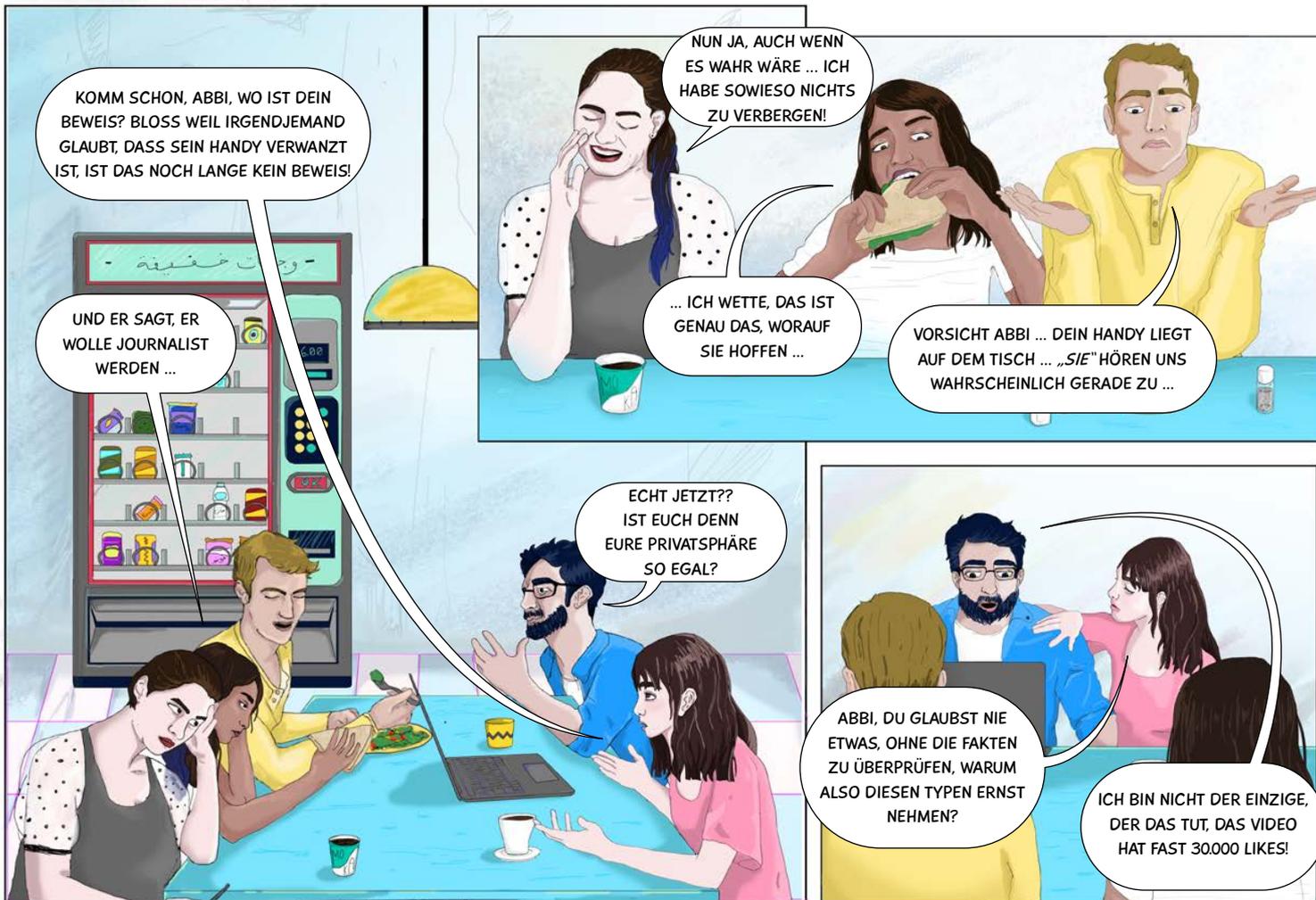
ABBI, DU ERWARTEST ECHT VON UNS DIR ZU GLAUBEN, DASS WIR VON UNSEREN HANDYS AUSSPIONIERT WERDEN?

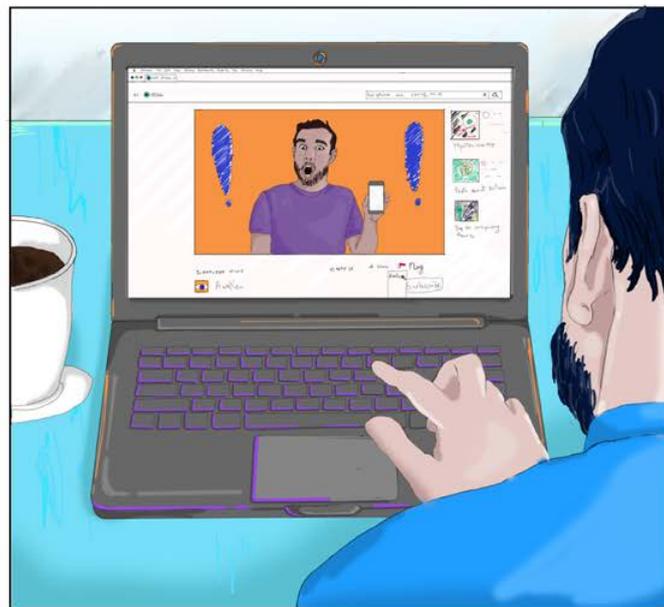
WIE SONST KÖNTE MAN VIDEOS WIE DIESES ERKLÄREN? DER TYP HAT KEINE KATZE, KENNT KEINE KATZEN, MAG SIE NICHT EINMAL,

ABER ER ERWÄHNT IN EINEM GESPRÄCH „KATZENFUTTER“ ... UND EIN PAAR TAGE SPÄTER ...

PUFF! ERSCHEINT EINE ANZEIGE FÜR KATZENFUTTER IN SEINEM FEED!

KOMMT SCHON LEUTE, LASST UNS EINFACH IN RUHE ESSEN ...









ALS SIE DEM RAUCH ZURÜCK ZUR QUELLE FOLGT, BEMERKT SHIRIN EIN GOLDENES SCHIMMERN IM SCHNEE ...



DAS SIEHT FAST WIE EIN KOMPASS AUS ... AUSSER ...



SOLCHE HIMMELSRICHTUNGEN HABE ICH NOCH NIE GESEHEN...



WER AUCH IMMER DIESE FUSSABDRÜCKE HINTERLASSEN HAT, WEISS VIELLEICHT MEHR ...









OH, MEIN EMPFEHLUNGSGEBER ! DU HAST IHN GEFUNDEN!

SIEHT FÜR MICH EHER WIE EIN KOMPASS AUS. WIE AUCH IMMER ... ICH GLAUBE, ER IST BEI DEM UNFALL KAPUTT GEGANGEN, ER HAT NICHT AUFGEHÖRT, SICH ZU DREHEN.

AH! DAS IST VÖLLIG NORMAL. ICH BIN BEIM KALTSTART ABGESTÜRZT.



EMPFEHLUNGSGEBER SIND WIE EIN BESONDERER KOMPASS. ABER ANSTATT DIR DEN WEG NACH NORDEN ZU ZEIGEN,

VERSUCHEN SIE, DICH IN DIE RICHTUNG ZU LENKEN, DIE DU AM EHESTEN SEHEN WILLST.



ABER WOHER WEISS ER, WAS ICH MAG?

NUN, DU ZEIGST ES IHM! IST DIR SCHON MAL AUFGEFALLEN, DASS PLATTFORMEN, AUF DENEN DU VIDEOS ANSIEHST, DIR ÄHNLICHE VIDEOS VORSCHLAGEN, DIE DIR GEFALLEN KÖNNTEN?



JEDES MAL, WENN DU EIN VIDEO ANSIEHST, IST DAS EIN SCHRITT IN DIESEM IRRGARTEN.

DER EMPFEHLUNGSGEBER VERFOLGT DIESE SCHRITTE UND NUTZT SIE, UM ZU ERRATEN, WOHN DU ALS NÄCHSTES GEHEN MÖCHTEST.







ALS DIE SUCHE NACH ABBI DOC Y UND SHIRIN TIEFER IN DEN IRRGARTEN FÜHRT, STOSSEN SIE AUF WEITERE DIESER SELTSAMEN KREATUREN ...

WAS SIND DAS FÜR DINGER?
ICH SEHE SIE ÜBERALL...



SIE HABEN VERSCHIEDENE NAMEN ... DIE MEISTEN LEUTE NENNEN SIE BENUTZERPROFILE, ICH NENNE SIE GERNE „DATENDOPPELGÄNGER“ ...

ABER IN WIRKLICHKEIT SIND SIE WIE DEIN DIGITALER ZWILLING.



ABER SIE SEHEN NICHT SEHR MENSCHLICH AUS ...

DAS HÄNGT VON DEINEN DATEN AB. HIER DRINNEN DEFINIEREN DATEN, WER DU BIST.

ÜBER WELCHE ART VON DATEN REDEN WIR?



SEHR UNTERSCHIEDLICHE ...

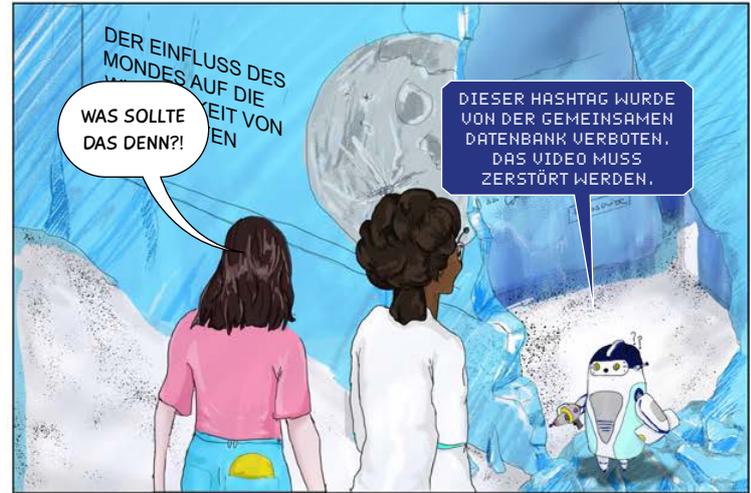
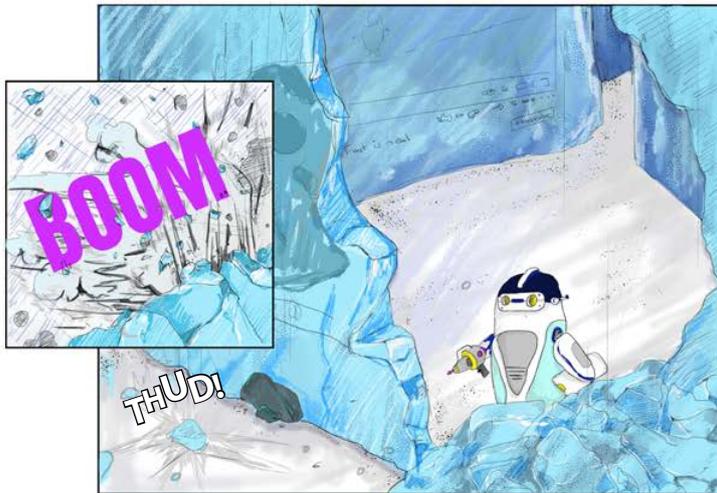
DIE VIDEOS, DIE DU DIR ANSIEHST UND DIE DIR GEFALLEN, DIE KOMMENTARE, DIE DU HINTERLÄSST, DIE SEITEN, DIE DU ABONNIERST, UND DIE LEUTE, DENEN DU FOLGST.



ABER WIR SIND SO VIEL MEHR ALS DAS, WAS WIR ONLINE TUN! INFORMATIONEN WIE DIESE SAGEN NICHTS DARÜBER AUS, WAS FÜR EIN MENSCH MAN IST ...









DER EINFLUSS DES MONDES WIRKT SICH IMMER SCHEINT MIR ZIEMLICH HART ZU SEIN ... WAS FÜR EIN VIDEO WAR DAS?

DIESER HASHTAG IST VERBOTEN. ER MUSS ZERSTÖRT WERDEN.

GENAUER GEHT ES NICHT. DAS IST NUR EIN BOT.



ABER WENN ER NICHT EINMAL WEISS, WAS ER LÖSCHT ... IST DAS NICHT GEFÄHRLICH?

DER BOT LÖSCHT NUR VIDEOS, DIE AUF EINER FAHNDUNGSLISTE STEHEN.

OH, WIRKLICH? WAS MUSS MAN TUN, UM AUF SO EINE LISTE ZU KOMMEN?



ES GEHT UM VERSTÖRENDE ODER GEWALTÄTIGE INHALTE. VIDEOS, DIE NACH MEINUNG FAST ALLER PLATTFORMEN NICHT INS INTERNET GEHÖREN.

DIESE PLATTFORM UND ANDERE ÄHNLICHE PLATTFORMEN TRAGEN ALLE ZU DIESER LISTE BEI.



DER EINFLUSS DES MONDES AUF DIE WIRKSAMKEIT VON IMPFSTOFFEN

ALSO ... DAS VIDEO IST WIE EIN VERBRECHER?

JA ... EINER, DER IN JEDEM LAND GESUCHT WIRD.

ABER TROTZDEM ... BEKOMMT NICHT JEDER MUTMASSLICHE VERBRECHER EINEN PROZESS?

OH SICHER, WENN ES EIN BESONDERS SCHLIMMER FALL IST ... ABER DAFÜR HABEN WIR ANDERE ALGORITHMEN ...

MIT JEDEM SCHRITT WERDEN SHIRIN UND DOC Y ZUVERSICHTLICHER UND SCHEINEN PLÖTZLICH KURZ DAVOR, DEN FALL ZU LÖSEN ...



JA ... ER FÄNGT AN, ZUVERLÄSSIGERE VORHERSAGEN ZU MACHEN ...

ER SOLLTE UNS DIREKT ZU ABBI FÜHREN, VORAUSGESETZT, DU KENNST DIE VORLIEBEN DEINES FREUNDES.



ICH DENKE SCHON ... GEHEIMGESSELLSCHAFTEN SIND GENAU SEIN DING ...



... ES SIEHT BLOCKIERT AUS.

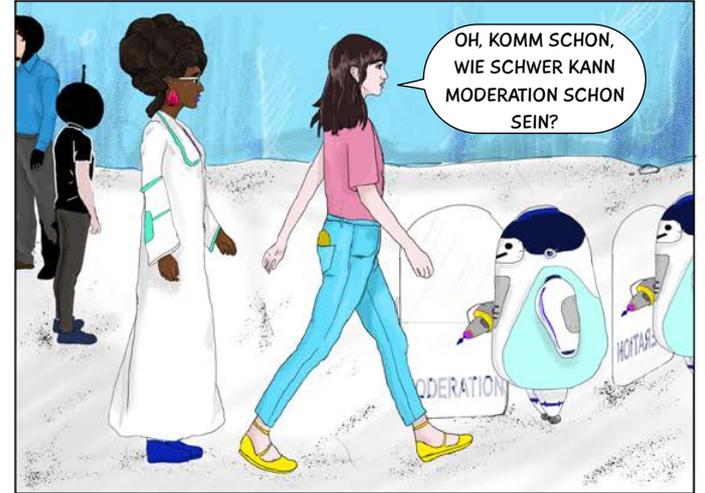
NUN JA ... WIR KÖNNTEN ES JA UMGEHEN ...



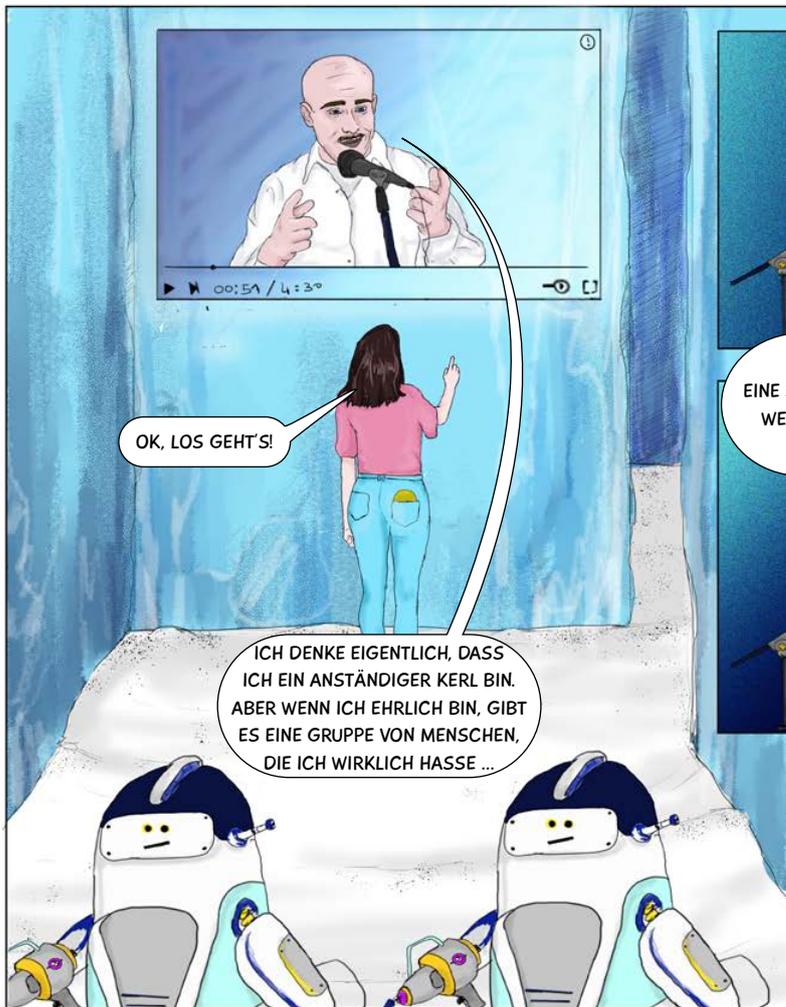
ICH DACHTE, ICH HÄTTE SCHON VOR LANGER ZEIT EINE ABKÜRZUNG IN DER RUBRIK „IMPFFSTOFF-DESINFORMATION“ GESEHEN ...

DER EMPFEHLUNGSGEBER SIEHT ZIEMLICH SICHER AUS ...

SCHAUEN WIR MAL, WOMIT WIR ES ZU TUN HABEN.







OK, LOS GEHT'S!

ICH DENKE EIGENTLICH, DASS ICH EIN ANSTÄNDIGER KERL BIN. ABER WENN ICH EHRLICH BIN, GIBT ES EINE GRUPPE VON MENSCHEN, DIE ICH WIRKLICH HASSE ...



IHR WISST, WEN ICH MEINE ... SIE SIND ÜBERALL, ES SIND WAHRSCHEINLICH AUCH EIN PAAR HIER HEUTE ABEND.

SIE BESCHWEREN SICH STÄNDIG, DASS DIE WELT NICHT FÜR SIE GEMACHT IST ...



SIE VERLANGEN EINE SONDERBEHANDLUNG, WEIL SIE „SO GEBORN WURDEN“ ...

MAN DENKT, DASS SIE EIGENTLICH EINIGERMASSEN ERTRÄGLICH SIND ... BIS MAN NEBEN IHNEN ISST.

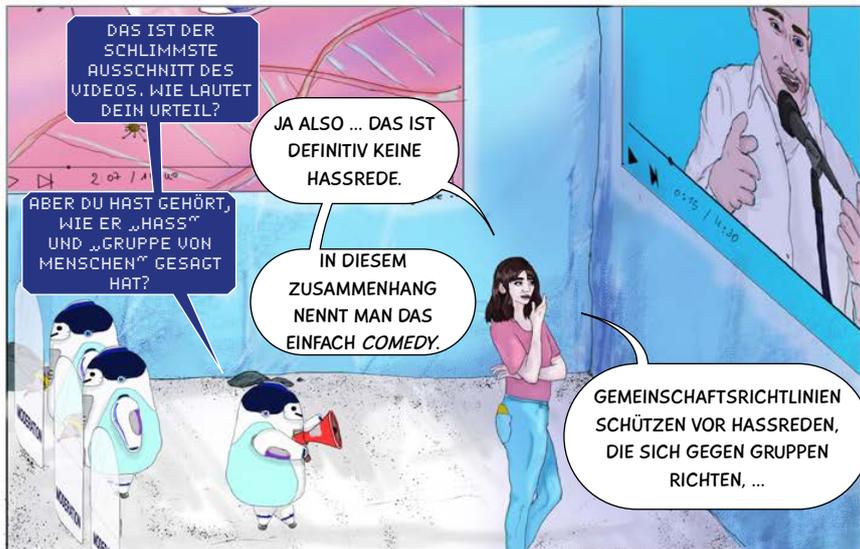
IHR WISST SCHON ... LINKSHÄNDER

ICH MEINE, WARUM KÖNNEN SIE NICHT EINFACH IHRE RECHTE HAND BENUTZEN WIE ALLE ANDEREN AUCH?



HE HE!

00:59 / 4:30



GERADE ALS ABBIS GEMELDETES VIDEO IN GREIFBARER NÄHE SCHEINT, MACHT SHIRIN EINEN VERHÄNGNISVOLLEN FEHLTRITT ...

PASS AUF, WO DU HINTRITTS! DA VORNE SIEHT ES EISIG AUS ...

ABER DER EMPFEHLUNGSGEBER IST GLÜHEND HEISS!



JA, ABER IHM BLINDLINGS ZU FOLGEN, IST EINE GEFÄHRLICHE ...

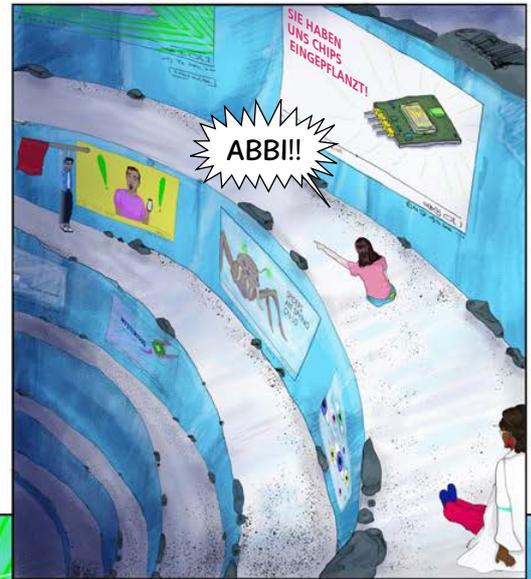
... ANGELEGENHEIT.

AAAAAAAAAHHHHH!!!



DIE HIER WERDE ICH SICHER GLEICH BRAUCHEN.











IST ES NICHT EIN
BISSCHEN KÜHL FÜR
EINEN TRENCHCOAT?

... DIESER TRENCHCOAT
ENTHÄLT EINIGE ZIEMLICH
BEEINDRUCKENDE
PRIVATSPHÄRE-
EINSTELLUNGEN! DIE KÄLTE
MACHT MIR NICHTS AUS, WEIL
ICH DADURCH UNSICHTBAR
SEIN KANN.



UND WAS HAT
DICH IN DIESE
ECHOKAMMER
GEFÜHRT?

ICH ARBEITE
GERADE AN EINER
GESCHICHTE ÜBER
DATENSCHUTZ UND
BIG DATA.

ICH HABE EINE SPUR
VERFOLGT, ABER ICH
MUSS IRGENDWO FALSCH
ABGEBOGEN SEIN ...

DAS PASSIERT
SELBST DEN
BESTEN VON UNS.



WER AUCH IMMER
DU BIST, DU
SPRICHST MIR AUS
DER SEELE!

MEIN NAME IST
ANGELA.

DU SIEHST NICHT
GERADE WIE EIN
DATENPROFIL AUS ...



ICH FASSE DAS ALS
KOMPLIMENT AUF
ICH BIN EIGENTLICH
JOURNALISTIN.

ICH ERKUNDE
DIESE PLATTFORM
UND ANDERE ÄHNLICHE
PLATTFORMEN SCHON
SEIT WOCHEN, BIS ICH
HIER GELANDET BIN.



DAS STIMMT! APROPOS,
WIE SIEHT DEINE
AUSSTIEGSSTRATEGIE AUS? DIE
KAMMER IST ENDLLOS ...







HAST DU DAS GEFÜHL, ALS WÄRST DU SCHON MAL HIER GEWESEN, ABBI?

ICH SAGE IHM SCHON SEIT JAHREN, DASS ER ZU VIEL ZEIT ONLINE VERBRINGT ...

MEINE COOKIES SIND EINFACH ÜBERALL! WOHER WEISST DU, DASS WIR DEN RICHTIGEN WEG GEHEN?



SIE KENNT VIELLEICHT NICHT DEN GENAUEN STANDORT EINER PERSON, ABER WENN SIE IHR ALTER, IHRE INTERESSEN UND EINIGE BESUCHTE WEBSITES KENNT, KANN SIE EINE FUNDIERTE VERMUTUNG ÜBER DEN REST ANSTELLEN.



GESPENSTISCH GENAU, WENN MAN NICHT AUF DIE SPUREN ACHTET, DIE MAN HINTERLÄSST.

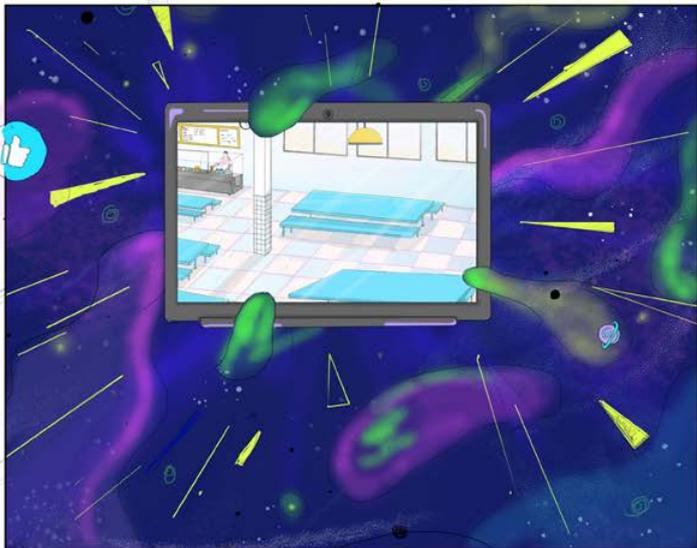


GENAU. ES MUSS DEINE GESPRÄCHE NICHT BELAUSCHEN.

DU SAGST IHM ALLES, WAS ES WISSEN MUSS, INDEM DU ES BENUTZT!



UM EHRlich ZU SEIN ... ICH GLAUBE, ICH AUCH.





ER HAT SEINEN
COOKIE-CACHE
GELÖSCHT ...
SCHLAUES KIND.



SHIRINS STÖRUNG WAR EINE
ÜBERRASCHUNG. ICH GLAUBE
LANGSAM, DASS SICH DIE
SITUATION VERSCHLIMMERT.

DU MEINST, SO
ETWAS WIRD WIEDER
PASSIEREN?

... UNSERE NACHFORSCHUNGEN
HABEN ERGEBEN, DASS ES
NOCH EINE WEITERE STÖRUNG
IN PLETHOR.A.I. GIBT ...



UND SO WIE ES AUSSIEHT,
LÄUFT UNS DIE ZEIT DAVON,
SIE ZU BEHEBEN.

ENDE

Glossar

(Big) Data Aggregation

ist das Zusammenfassen, Gruppieren oder anderweitige Organisieren der Rohdaten eines bestimmten Datensatzes für statistische Analysen. Während dies in der Vergangenheit von Menschen durchgeführt wurde (mit herkömmlichen statistischen Analysen), werden im Zeitalter von Big Data KI-Lösungen genutzt, um Muster in diesen unglaublich großen Datensätzen zu extrahieren, zu erkennen oder zu extrapolieren, die dann für eine Vielzahl von Anwendungen genutzt werden. Viele der heutigen Empfehlungsgeber arbeiten beispielsweise mit Datenaggregation, indem sie Personen mit ähnlichen Profilen zusammenfassen oder einzelne Nutzerinnen und Nutzer mit einem neuen Produkt, einer Seite oder einer Interessengruppe zusammenbringen.

Community-Richtlinien

werden manchmal mit «Nutzungsbedingungen» assoziiert und sind eine Reihe von Regeln, die von einer Plattform oder einem sozialen Netzwerk aufgestellt werden, um verbotene oder bestimmte erwartete Verhaltensweisen menschlicher Nutzerinnen und Nutzer zu bestimmen. Im Allgemeinen sollen Community-Richtlinien ein sicheres Umfeld für die Interaktion der Nutzerinnen und Nutzer gewährleisten (z.B. durch das Verbot von Hassrede oder der Verbreitung medizinischer Falschinformationen), sie können aber auch rechtliche Themen wie Urheberrechtsverletzungen enthalten. Wichtig ist, dass sie einen Standard festlegen, nach dem alle Inhalte auf der Plattform beurteilt werden, und dass die Plattformmoderation (Menschen oder KI-Systeme) die Gemeinschaftsrichtlinien durchsetzt.

Inhaltsmoderation

beschreibt den Prozess, bei dem Menschen oder Künstliche Intelligenz entscheiden, welche Inhalte auf einer bestimmten Plattform oder in einem Forum erlaubt sind, in der Regel in Übereinstimmung mit den Community-Richtlinien oder Nutzungsbedingungen der Plattform oder des Forums. Heute werden die meisten großen Plattformen von einer Kombination aus Menschen und KI-Systemen moderiert. Das liegt zum Teil an der beeindruckenden Menge neuer Inhalte, die täglich hochgeladen werden – viel zu viel für Menschen alleine. Die KI-gestützte Moderation ist eine effiziente Lösung für dieses Problem. Sie ist jedoch unfähig, sprachliche Nuancen wie Humor oder Sarkasmus zu erkennen. Das kann in Verbindung mit dem Druck verschiedener staatlicher Stellen, unliebsame Inhalte auf

allen Plattformen zu zensieren, die Meinungsfreiheit gefährden und zu einer verstärkten Online-Zensur führen.

Personalisierung von Inhalten

, auch Empfehlungsgeber genannt, ist eine beliebte Methode, mit der verschiedene KI-Plattformen (soziale Medien, Marktplätze, Inhaltsplattformen und mehr) sichtbare Inhalte individuell für Nutzerinnen und Nutzer filtern. Diese Plattformen sammeln Informationen (oder Daten) zu den persönlichen Gewohnheiten und Interessen ihrer Nutzerinnen und Nutzer (z.B. welche Bücher eine bestimmte Person gekauft hat, welche Lieder sie mag, welche Websites sie besucht hat) und verwenden diese, um ihren Nutzerinnen und Nutzern neue Inhalte, Produkte oder Dienstleistungen zu empfehlen.

Glossar

Empfehlungsgeber wie diese schlagen den Nutzenden in der Regel neue Inhalte auf zwei Arten vor: entweder durch die Suche nach neuen Inhalten, die mit den Inhalten in Verbindung stehen, die eine Person bereits mag oder mit denen sie sich beschäftigt hat (*Inhaltsfilterung*) oder durch die Zusammenführung von Nutzenden mit ähnlichen Interessen (*kollaboratives Filtern*). Interessanterweise kann der Empfehlungsgeber, wenn eine Person zum ersten Mal mit einer Plattform interagiert, mangels ausreichender Daten nur schwer erkennen, welche Inhalte der Person wahrscheinlich gefallen werden. Dieses Phänomen wird *Kaltstartproblem* genannt.

Cookies sind kleine Dateien, die bei jedem Besuch einer Website über den Internetbrowser auf dem Computer der Nutzenden gespeichert

werden, um Informationen über diese zu sammeln und zu verfolgen. Einige Cookies, sogenannte Sitzungscookies, dienen in erster Linie dazu, das ordnungsgemäße Funktionieren einer bestimmten Website zu gewährleisten, während die Nutzerin oder der Nutzer sich mit ihrem Inhalt beschäftigt, indem sie sich beispielsweise die Spracheinstellungen merken. Andere Cookies, in der Regel sogenannte dauerhafte Cookies, werden dauerhaft auf den Computern der Nutzenden gespeichert und ermöglichen die Nachverfolgung über verschiedene Websites hinweg. Das erhöht die Genauigkeit gezielter Werbeanzeigen, die Nutzenden unabhängig von der besuchten Website angezeigt werden.

Eine **Echokammer** entsteht, wenn eine Gruppe von Teilnehmenden auf einer (Social-

Media-)Plattform bevorzugt miteinander in Verbindung tritt und dabei Andersdenkende ausschließt. Mit der Zeit kann dieser Ausschluss (und der damit oft einhergehende Verlust an Meinungsvielfalt) dazu führen, dass die Nutzenden zunehmend von abweichenden Meinungen zu einem bestimmten Thema abgeschottet werden. Dies wiederum kann diese Nutzenden zu immer extremeren Meinungen treiben, die sie ursprünglich dazu veranlasst haben, sich für diese Gruppe zu interessieren, wodurch ein Umfeld entsteht, in dem jedes Mitglied der Gruppe die Meinungen der anderen „nachplappert“.

Meinungsfreiheit ist ein universelles Recht, das die Freiheit betrifft, Informationen und Ideen aller Art zu suchen, zu erhalten und weiterzugeben, und zwar ungeachtet von Barrieren,

sei es mündlich, schriftlich oder in gedruckter Form, in Form von Kunst oder durch jedes andere Medium. Diese Freiheit schließt die Möglichkeit ein, Kritik zu üben und Meinungen zu widersprechen, politisches Material zu veröffentlichen, Wahlkampf zu betreiben und für politische Ideen zu werben. Dieses Recht beinhaltet auch die freie Kommunikation von Informationen und Ideen zu öffentlichen und politischen Themen zwischen Bürgerinnen und Bürgern, Kandidatinnen und Kandidaten und gewählten Vertreterinnen und Vertretern. Es erfordert eine freie Presse und die Freiheit anderer Medienkanäle, die in der Lage sein sollten, sich ohne Zensur oder Einschränkung zu öffentlichen Themen zu äußern und die Öffentlichkeit zu informieren.

Schlussfolgerung

In einer Welt, die sich blitzschnell zu verändern scheint, ist es oft schwierig, das, was um uns herum geschieht, zu verstehen. Noch nie zuvor waren wir so vielen Informationsquellen und einer so großen Meinungsvielfalt ausgesetzt, doch paradoxerweise wurde auch unser Verständnis von Wahrheit, Glaubwürdigkeit und journalistischer Integrität noch nie so stark erschüttert. Wir treten in eine Ära ein, in der Falschinformationen und Polarisierung die Norm sind, in der die Wahrheit im Auge des Betrachtenden zu liegen scheint und in der öffentliche Institutionen, denen man lange Zeit vertraut hat, darum ringen, sich über die hitzige Debatte hinweg Gehör zu verschaffen.

Dennoch erfordern die Probleme der heutigen Welt energische und gemeinsame Anstrengungen von jeder und jedem von uns. Als Nutzerinnen und Nutzer müssen wir die Art und Weise kultivieren, wie

unsere Daten uns definieren, indem wir über die von den Algorithmen vorgeschlagenen Inhalte hinausblicken und uns der Spuren bewusst werden, die unsere digitalen Schritte hinterlassen. Als Bürgerinnen und Bürger dürfen wir nie aufhören, die Glaubwürdigkeit der Informationen, die wir erhalten, zu hinterfragen und nach der harten und komplexen Wahrheit zu suchen, statt nach einem einfachen, aber überzeugenden Märchen.

Schließlich müssen wir als Bürgerinnen und Bürger einer globalen Gesellschaft der Bequemlichkeit der Echokammer widerstehen, sowohl online als auch offline. Nur wenn wir uns mit den Meinungen anderer auseinandersetzen und uns ernsthaften Gesprächen über persönliche Überzeugungen stellen, können wir in unserer zunehmend gespaltenen Welt eine gemeinsame Basis finden. ■



Episode IV

Joaquín und die unlösbare Frage

Konzeption und Skript: Dr. Katherine Evans

Illustration: Adriana De La Torre Cervantes

Unterstützung bei der Illustration: Patricia Manríquez, Karin Almazán



Einleitung

Wenn wir darüber nachdenken, wie sehr Künstliche Intelligenz unser Leben bereits beeinflusst hat und weiterhin beeinflussen wird, liegt es nahe, die KI als allwissende, allgegenwärtige und generell allmächtige Technologie zu betrachten. KI bewältigt scheinbar übermenschliche Aufgaben – wie statistische Analysen oder die Beherrschung komplexer strategischer Spiele – in Rekordzeit und mit unglaublicher Leichtigkeit. Dies führt dazu, dass wir Menschen oft nicht verstehen, wie KI zu ihren Schlussfolgerungen kommt, und wir sogar den Eindruck haben, die KI habe einen „eigenen Kopf“.

Hinzu kommt die beunruhigende Tatsache, dass wir als einzelne Nutzerin oder als einzelner Nutzer nur sehr wenig Einblick in das haben, was genau hinter unseren Bildschirmen vor sich geht: Wie unsere täglich genutzten KI-Tools konzipiert sind, welche Überlegungen in ihre Entscheidungen einfließen oder wie diese überhaupt getroffen werden und auf welche physischen und virtuellen Ressourcen sie zurückgreifen. KI wird oft als scheinbar schwereloses, müheloses und allgegenwärtiges Werkzeug dargestellt, das beispielsweise Lösungen für unsere komplexen globale Umweltprobleme bietet. Mit anderen Worten: Viele Bereiche des täglichen Lebens erscheinen grüner, einfacher und sicherer, wenn sie „in der Cloud“ gelebt werden.

Diese Sichtweise auf Künstliche Intelligenz lässt jedoch einen wichtigen Teil des Bildes außer Acht. Während KI-Lösungen für die Bewältigung des Klimawandels von zentraler Bedeutung sind und sein werden – durch die Entwicklung von KI-gestützten Katastrophenwarnsystemen,

öffentlichen Informationskampagnen, komplexen Klimasimulationen oder Systemen zum Aufspüren gefährdeter Wildtiere, um nur einige zu nennen – sind die ökologischen Kosten vieler aktueller KI-Praktiken alles andere als gering.

In gewisser Weise ist dies auf die unglaublichen Rechenkosten zurückzuführen, die hinter vielen leistungsstarken KI-Systemen stecken: In einigen Fällen können die Big-Data-Mining-, Extraktions- und Trainingsprozesse hinter vielen datengesteuerten Formen der KI fast so viel Energie verbrauchen wie Städte in einem ganzen Jahr. Somit geben wir der rechnerischen *Genauigkeit* gegenüber der ökologischen *Effizienz* den Vorzug. Es liegt auch daran, dass wir noch nicht gelernt haben, viele der Hardwarekomponenten, die für die KI notwendig sind, zu recyceln. Stattdessen erzeugen wir große Mengen an Elektroschrott, der Gemeinden und natürliche Lebensräume verschmutzt und gefährdet.

Tatsächlich wird unsere Erfahrung mit blitzschnellen und hochintelligenten KI-Tools also nicht durch einen superintelligenten Geist in einer Maschine ermöglicht, sondern durch die rechenintensive Analyse in relativ einfachen Modellen, die in großem Maßstab eingesetzt werden. Während uns die KI beispiellose Versprechen und Möglichkeiten bietet, müssen wir uns bei der Betrachtung beider Seiten der Medaille fragen, ob KI Teil des Problems oder Teil der Lösung ist. In gewisser Weise mag es wie eine unlösbare Frage erscheinen ... ■

IRGENDWO TIEF IN
PLETHOR.A.I. ...

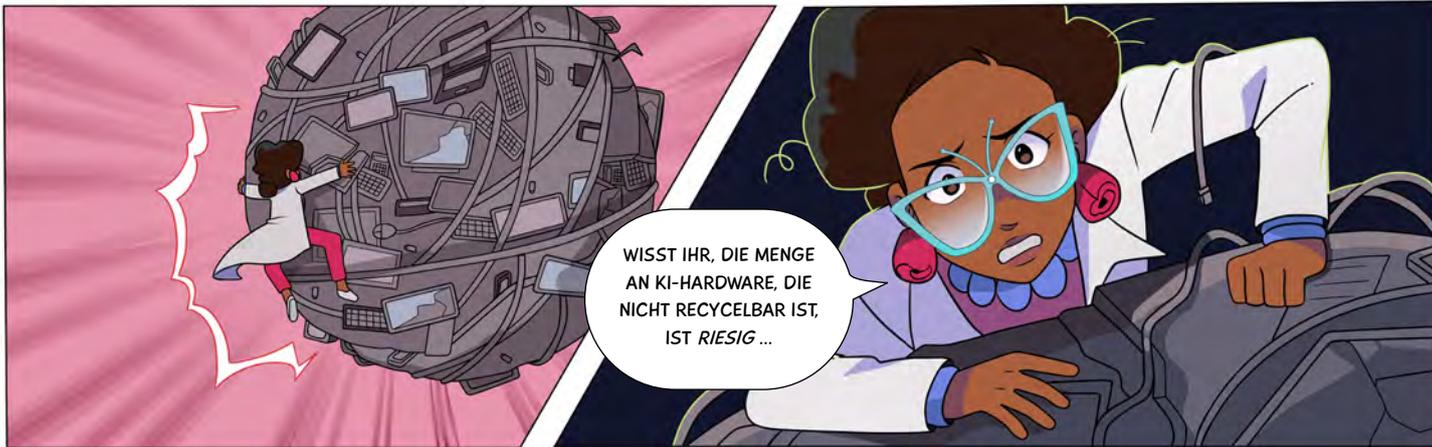
IST ...

JEMAND ...

... DA??









ANDERSWO IN PLETHOR.A.I. SITZT EMILE IN SEINEM KI-FORSCHUNGSZENTRUM UND ARBEITET FLEISSIG AN SEINER NEUESTEN KI-ANWENDUNG: EINEM DIGITALEN ZWILLING DES PLANETEN ERDE.



IRGENDWO AN DER DIGITALEN OSTSEEKÜSTE...

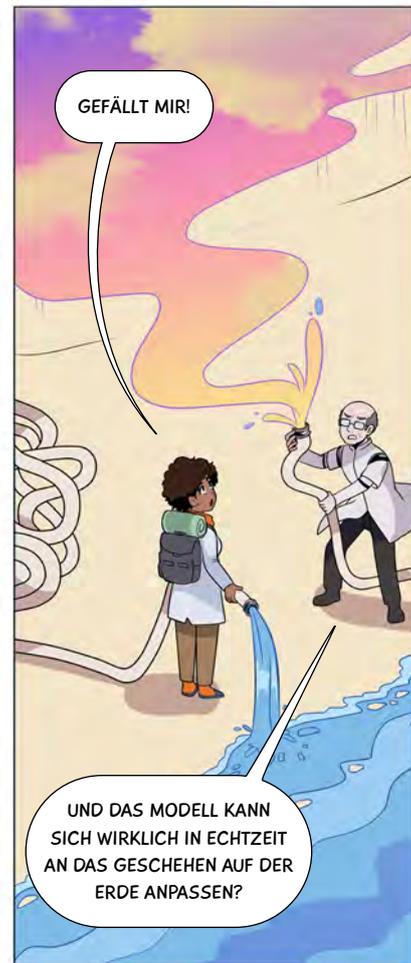
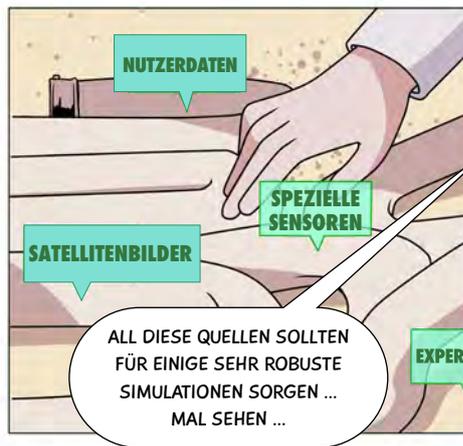
NICHT SCHLECHT ...
DAS MODELL SCHEINT SICH
GUT ZU ENTWICKELN ...



RENEE? BIST DU
DAS?

EMILE! DU
KOMMST GERADE
RECHTZEITIG!







DAS IST RICHTIG ... VERKEHR, WOLKEN, ENERGIEVERBRAUCH, WIND, WELLEN, ALL DAS!

DAS MODELL STECKT NOCH IN DEN KINDERSCHUHEN, ABER WENN ES ERST EINMAL LÄUFT, SOLLTEN WIR IN DER LAGE SEIN, ALLE ARTEN VON KLIMAEREIGNISSEN – VON HURRICANES BIS HIN ZU DÜRREN – ZU KARTIEREN, VORHERZUSAGEN UND DARAUF ZU REAGIEREN!

NATÜRLICH SIND SOLCHE KI-ANWENDUNGEN SOMIT AUCH TOTAL NÜTZLICH, UM DEN KLIMAWANDEL ZU BEKÄMPFEN, ...

JA, UND ES IST IMMER SCHÖN, MIT KLIMAWISSENSCHAFTLERINNEN WIE DIR ZUSAMMENZUARBEITEN!



HEY IHR ZWEI! KANN ICH MITMACHEN?

ERNO, BIST DU DAS?



DANKE FÜR DIE UNTERSTÜTZUNG, EMILE!

EMILE, DAS IST ERNO.

ER IST MEERESBIOLOGE UND HILFT UNS BEI EINIGEN ANWENDUNGEN ZUM SCHUTZ DER TIERWELT ...



ICH KANN ES KAUM ERWARTEN, EUCH ZU ZEIGEN, WAS ICH HIER HABE!



DAS SIND SAIMAA-ROBBEN, EINE BEDROHTE ART, DIE AN DIESER KÜSTE HEIMISCH IST ...

MIT DIESEM DATENSATZ KÖNNEN WIR INDIVIDUEN UND GRUPPEN ÜBER IHR GESAMTES LEBEN HINWEG VERFOLGEN ... ZIEMLICH RAFFINIERT, ODER?



ERNO, DAS IST UNGLAUBLICH! DAS TEAM WIRD BEGEISTERT SEIN!



MIT DEN ALGORITHMEN, DIE WIR ENTWICKELT HABEN, SOLLTEN WIR IN DER LAGE SEIN, ALLE GEFÄHRDETEN ARTEN ZU TRACKEN ...

JA ... ABER SAG MAL, ERNO ... GIBT ES VIELE TSUNAMIS, WO DU HERKOMMST?



EIN TSUNAMI IN DER OSTSEE?? DAS IST UNMÖGLICH!

EMILE, KÖNNTE DAS WIRKLICH AUF DER ERDE PASSIEREN?

DAS BEZWEIFLE ICH ... DA IST EINDEUTIG ETWAS ANDERES IM GANGE ...



EMILE! WAS IST HIER LOS?!

KEINE PANIK ...
LASS MICH
NACHDENKEN ...

ES MUSS EINE ART VON
VERZERRUNG IN DEN DATEN GEBEN,
DIE DAS MODELL DAZU BRINGT,
FALSCHER SCHLÜSSE ZU ZIEHEN ...

ABER BEI DIESEM
AUSMASS ... MUSS DIE
GESAMTE PLATTFORM
GEFÄHRDET SEIN ...



ABER DAS
WÜRD
BEDEUTEN ...

AUF KEINEN FALL!
DAS KANN ICH
NICHT GLAUBEN!

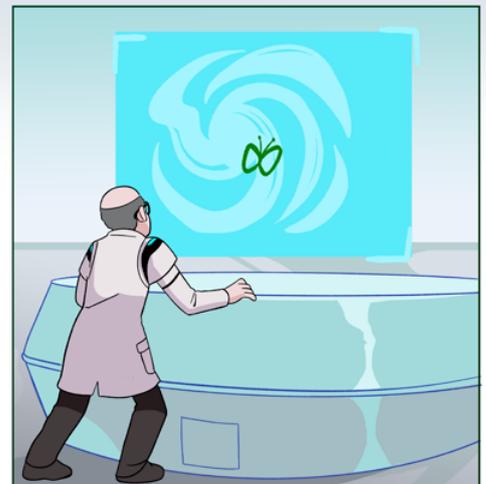
DIE CLOUD SELBST
MUSS RIESIGE
DATENMENGEN VERLIEREN ...
WIR SOLLTEN DEN STECKER
ZIEHEN, BEVOR ...

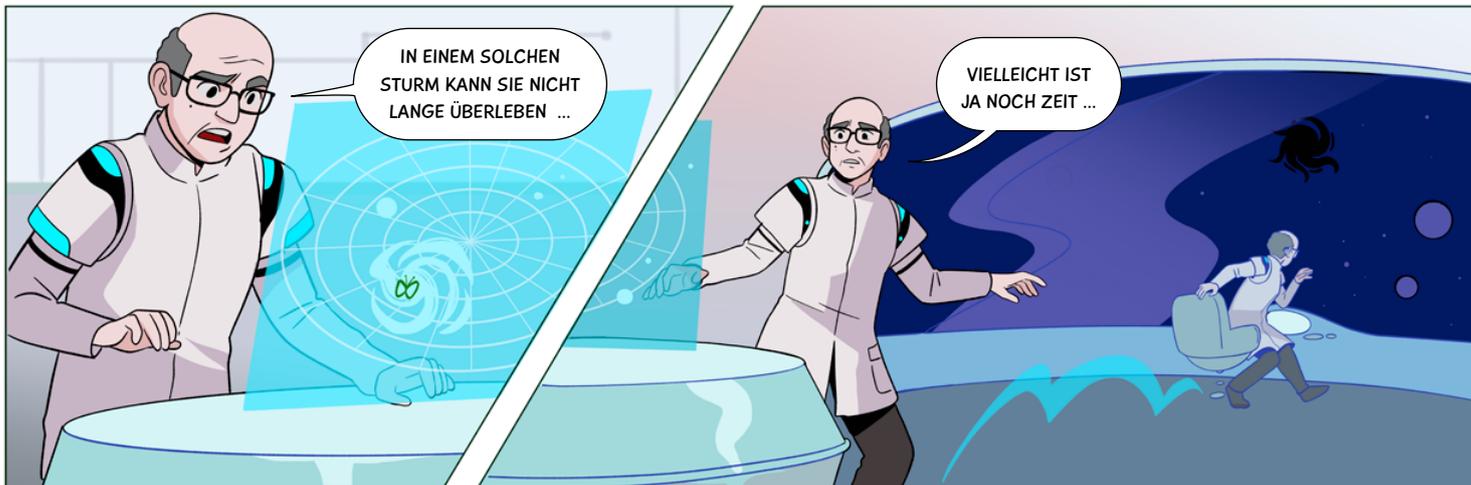


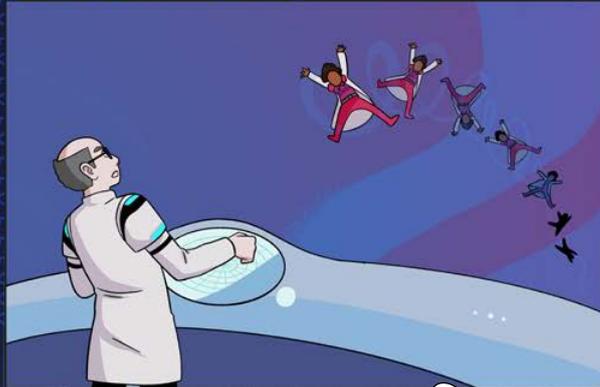
VIEL GLÜCK!

RAUS HIER, ALLE BEIDE! ICH
WERDE VERSUCHEN, DAS
PROBLEM ZU BEHEBEN!



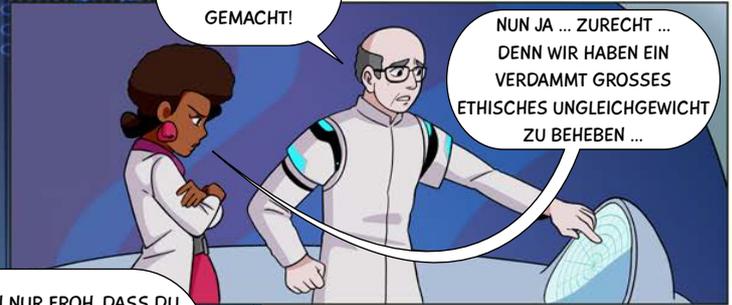






ICH HABE MIR
SORGEN UM DICH
GEMACHT!

NUN JA ... ZURECHT ...
DENN WIR HABEN EIN
VERDAMMT GROSSES
ETHISCHES UNGLEICHGEWICHT
ZU BEHEBEN ...



ICH BIN NUR FROH, DASS DU
NICHT MEHR MITTENDRIN
BIST ... KONNTEST DU SEHEN
WAS DAS PROBLEM IST?



SLAP!

DOC Y! ICH WUSSTE,
DASS DU DICH AN DEN
PLAN HALTEN WÜRDST!



NICHT WIRKLICH ... ABER ICH
HABE EINE IDEE, WIE WIR DAS
HERAUSFINDEN KÖNNEN ...

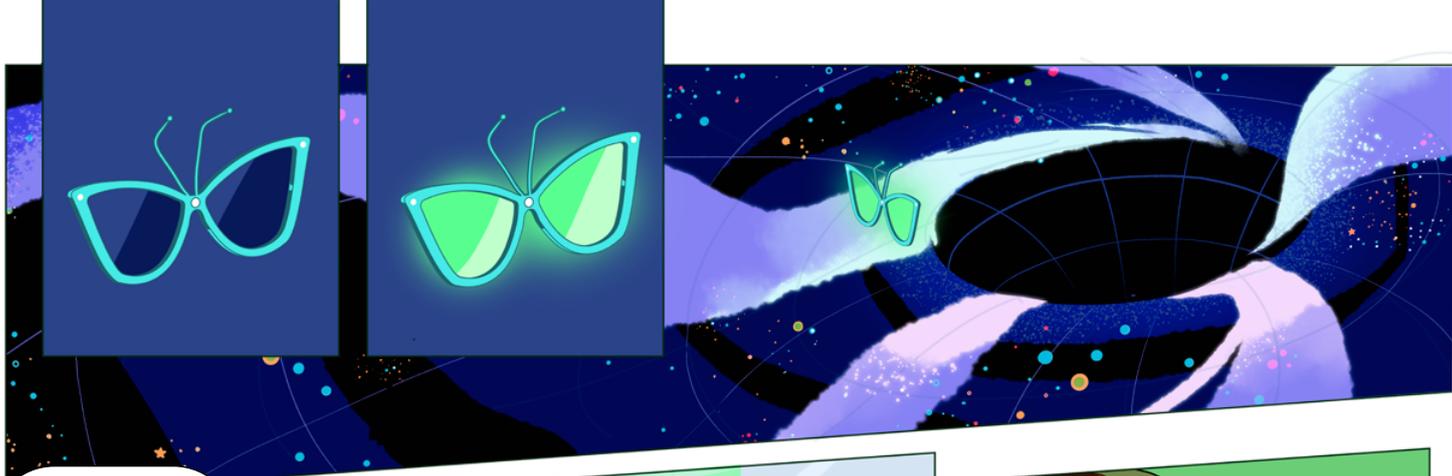


ES SIEHT FAST SO AUS,
ALS WÄRE EINE WEBSUCHE
SCHIEFGELAUFEN ...

ABER WELCHE ART
VON ABFRAGE WÜRD
SOLCHE DATENMENGEN
BENÖTIGEN?

KÖNNEN WIR
IRGENDWIE NOCH EINEN
BESSEREN ÜBERBLICK
BEKOMMEN?

OK INTELLIGENTE BRILLE,
SUCH DAS AUGE DES
STURMS !



ICH HABE DIE LETZTE
STÖRUNG NIE GEFUNDEN ...
VIELLEICHT IST SIE DAS?



KEINE ANDERE STÖRUNG HAT
AUCH NUR ANNÄHERND DIESES AUSMASS
ERREICHT ... WENN DAS STIMMT, DANN
HABEN WIR ES MIT DER MUTTER ALLER
STÖRUNGEN ZU TUN.

ICH GLAUBE, DA KÖNNTE
ETWAS DRAN SEIN ...





DENK DOCH MAL NACH ... EINE STÖRUNG DIESER GRÖSSE KÖNNTE DIE ANDEREN STÖRUNGEN ÜBERHAUPT ERST VERURSACHT HABEN ...

SICHER, ABER ES GIBT KEINEN BEWEIS DAFÜR, DASS EINE STÖRUNG DIE URSACHE FÜR DIESEN STURM IST!

ABER DU MUSST ZUGEBEN ... DIE VERMUTUNG LIEGT DOCH ZIEMLICH NAHE!

KORRELATION BEDEUTET NICHT KAUSALITÄT!



HALLO?? KANN MICH JEMAND HÖREN??

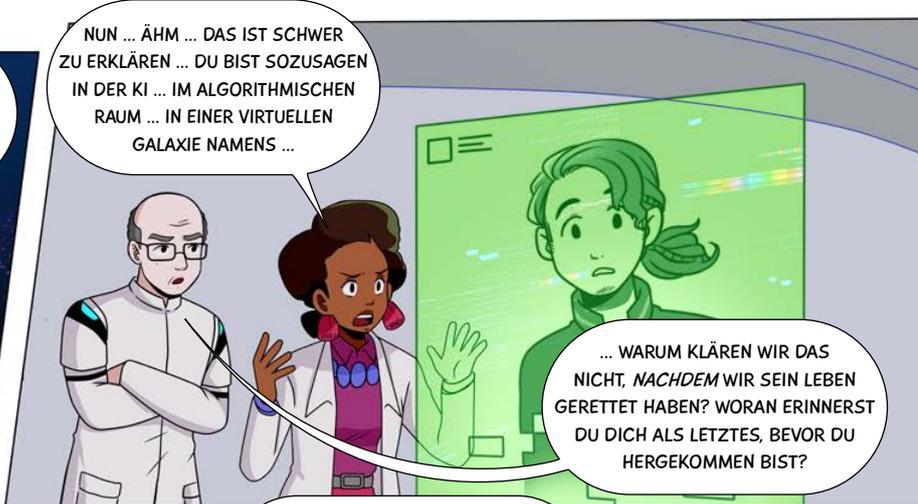
OH EMILE, BITTE!

ICH WEISS, DU MUSST NOCH JEMANDEN FINDEN, ABER WIR SOLLTEN UNS AUF DAS KONZENTRIEREN, WAS VOR UNS LIEGT!



JA! JA! WIR KÖNNEN
DICH HÖREN ...

GOTT SEI DANK!
WAS ZUM HENKER
GEHT HIER VOR
SICH?!?!



NUN ... ÄHM ... DAS IST SCHWER
ZU ERKLÄREN ... DU BIST SOZUSAGEN
IN DER KI ... IM ALGORITHMISCHEN
RAUM ... IN EINER VIRTUELLEN
GALAXIE NAMENS ...

... WARUM KLÄREN WIR DAS
NICHT, *NACHDEM* WIR SEIN LEBEN
GERETTET HABEN? WORAN ERINNERST
DU DICH ALS LETZTES, BEVOR DU
HERGEKOMMEN BIST?



ICH WAR ALLEIN ... ICH WEISS NICHT ...
HAB MEINEM HANDY FRAGEN GESTELLT ...
HERUMGESPIELT ...

FRAGEN, AN EINEN
SPRACHASSISTENTEN?

JA ... GENAU. ALSO,
AN DIE DAME AUF
MEINEM HANDY ...

DAS IST EINE WIRKLICH
SCHRECKLICH
KLISCHEEHAFTE
PROGRAMMIERUNG.

ALS WÄREN FRAUEN VON
NATUR AUS UNTERWÜRFIG
UND DAZU BESTIMMT,
ZU GEFALLEN ...



... NOCHMAL: RETTEN WIR DOCH
BITTE *ZUERST* SEIN LEBEN, UND
KÜMMERN UNS DANACH UM DEN REST.
ALSO, WAS IST GENAU PASSIERT?

ICH WAR AM FREITAGABEND ALLEINE ZU HAUSE
UND MIR WAR IRGENDWIE LANGWEILIG ...







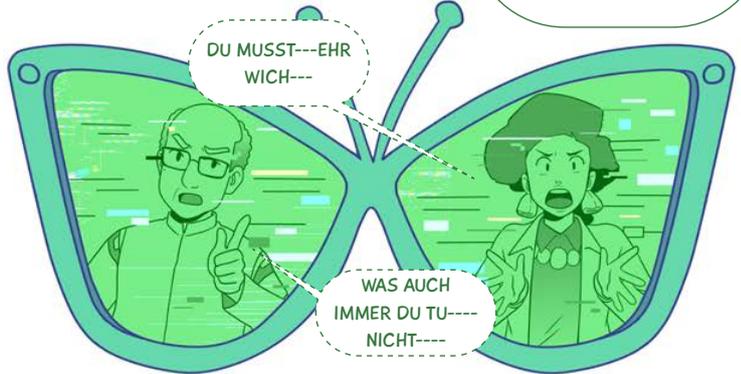
KANN ICH DIR SONST NOCH IRGENDWIE HELFEN?

DAMIT KANNST DU MIR HELFEN! ICH MÖCHTE WISSEN, WIE ES IST, DU ZU SEIN!



UND DANN ... WURDE ALLES VERSCHWOMMEN. DANN BIN ICH HIER AUFGEWACHT.

KEINE SORGE, JOAQUÍN. WIR KRIEGEN DAS HIN, ES WIRD ALLES GUT. DU MUSST NUR ZUHÖREN ...



DU MUSST---EHR WICH---

WAS AUCH IMMER DU TU--- NICHT---



... HALLO??



HABEN WIR DEN KONTAKT ZU JOAQUÍN VERLOREN??

ICH MACHE MIR GERADE MEHR SORGEN, DIE KONTROLLE ÜBER DAS SCHIFF ZU VERLIEREN!



DAS IST UNGLAUBLICH ... ES MUSS TONNEN VON DIESEM ZEUG DA DRAUSSEN GEBEN!



TATSÄCHLICH ETWA 50 MILLIONEN TONNEN PRO JAHR!

NUR ETWA 20 % DAVON WERDEN RECYCLET! DER REST LANDET AUF MÜLLDEPONIEEN, VERSCHMUTZT SOMIT LEBENSMITTEL UND DAS GRUNDWASSER UND SETZT UNS GEFÄHRLICHEN GIFTEN AUS ...



ABER WAS MACHT DAS ZEUG HIER??

... UND BESCHÄDIGT MEIN SCHIFF!



ES IST WIRKLICH EINE SCHANDE. EIN GROSSER TEIL DAVON ENTHÄLT SELTENE ERDEN, DIE WIR BRAUCHEN UND DIE UNS IRGENDWANN AUSGEHEN WERDEN ...

DER SCHADEN WIRD ZU GROSS, ZEIT FÜR PLAN B.



WAS IST PLAN B, EMILE?!
ELEKTROSCHROTT IST EIN
GLOBALES PROBLEM ...

NUR EINE SOLIDE
KREISLAUFWIRTSCHAFT
KÖNNTE IHN VERMEIDEN!

CRASH!

ICH WEISS, INGENIEURE
MÜSSEN FÜR ALLES
GERÜSTET SEIN.

ABER WARUM SOLLTE EIN
WISSENSCHAFTLICHES
FORSCHUNGSSCHIFF EINEN
LASER BRAUCHEN?

NA JA, ES HAT
KEINEN LASER
GEBRAUCHT ...
BIS JETZT.

DEN STURM IM BLICK, TAUSCHEN UNSERE TAPFEREN
KI-DETEKTIVE THEORIEN ÜBER DEN URSPRUNG FÜR
JOAQUINS STÖRUNG AUS ...

ICH KANN DIE VERBINDUNG
NICHT WIEDERHERSTELLEN ...
JOAQUÍN MUSS FÜRCHTERLICHE
ANGST HABEN!

DAS SCHIFF WIRD VON DER
KRAFT DES STURMS ANGEZOGEN ...
ES IST NUR EINE FRAGE DER ZEIT,
BIS WIR AUF IHN STOSSEN.

UND WAS FÜR EINE
KRAFT DAS IST ... WIR
HABEN NOCH GAR NICHT
DARÜBER GESPROCHEN,
MIT WAS WIR ES HIER ZU
TUN HABEN ...







GENAU! DIE KI, DIE EINEN SCHACHCHAMPION SCHLÄGT, KANN SICH NICHT EINMAL EINEN *SCHUH* ZUBINDEN ODER VERSTEHEN, WAS EIN *FRÜHSTÜCK* IST ...

MIT ANDEREN WÖRTERN:
WIR SIND NOCH WEIT ENTFERNT VON EINER *ALLGEMEINEN KÜNSTLICHEN INTELLIGENZ!*



SICHER ... ABER WENN JOAQUÍNS SPRACHASSISTENT IRGENDWIE ANGEFANGEN HAT, ÜBER SICH SELBST NACHZUDENKEN ...

DAS IST DOCH VERRÜCKT ...



... UND WENN DAS EINE INTELLIGENZEXPLOSION IN PLETHOR.A.I. AUSGELÖST HAT ...

EINE, DIE SICH AUF ALLE DATEN IN DER CLOUD AUSGEWIRKT HAT ...

... VORSICHTIG, DOC Y.



DANN HABEN WIR ES VIELLEICHT MIT DEM ERSTEN FALL EINER SUPERINTELLIGENTEN KI ZU TUN DIE GENAUSO IST WIE WIR, NUR INTELLIGENTER!



WIE KANNST DU ES WAGEN,
DIESE WORTE IN DIESEM
LABOR AUSZUSPRECHEN!



OH, KOMM SCHON!!
ICH SAGE NUR,
VIELLEICHT ...

GUT, DOC Y. WENN DU SO
DAVON ÜBERZEUGT BIST, DASS
DIESER SPRACHASSISTENT ZU EINER
ART ÜBERMENSCH GEWORDEN IST,
DANN WEISST DU JA, WAS DU ZU
TUN HAST.



HAST DU DENN
ÜBERHAUPT
EINEN AN BORD?

WIE DU RICHTIG ZU
SAGEN PFLEGST, MÜSSEN
PROGRAMMIERERINNEN
UND PROGRAMMIERER
AUF ALLES VORBEREITET
SEIN ...



GEBT BITTE NICHT AUF,
LEUTE! IHR SEID EIN TOLLES
RETTUNGSTEAM ...

WIR SIND IMMER NOCH
HIER, UM DICH ZU RETTEN,
JOAQUIN. ABER ZUERST
MUSS DOC Y IHRE THEORIE
ÜBERPRÜFEN ...

WER HAT
GESAGT, DASS
ICH NICHT
BEIDES KANN?



ES DREHT SICH EINFACH ENDLOS
IM KREIS ... ICH HABE ALLES
VERSUCHT, WAS MIR EINGEFALLEN
IST, UM ES ZU STOPPEN ...

ES NEU ZU STARTEN ... IHM EINE
NEUE FRAGE ZU STELLEN ... ICH
HABE SOGAR VERSUCHT, NACH
UPDATES ZU SUCHE ...



ES SIEHT SO AUS, ALS
OB DIE DINGE EIN WENIG
AUSSER KONTROLLE
GERATEN SIND ...

... UND GENAU DESHALB
MÜSSEN WIR EINEN ZAHN
ZULEGEN. EMILE?



WAS HAST
DU DA?



ES KANN SEHR SCHWIERIG SEIN ZU VERSTEHEN, WIE EINE KI DENKT ... HIERMIT KÖNNEN WIR IHRE INTELLIGENZ TESTEN.



OK ... IST INTELLIGENZ EIN GUTES ZEICHEN?

DOC Y SCHEINT ZU GLAUBEN, DASS DIE „DAME IN DEINEM HANDY“ EINEN EIGENEN WILLEN ENTWICKELT HAT.

DIE EINZIGE MÖGLICHKEIT, DAS HERAUSZUFINDEN, IST, DIESEN TEST DURCHZUFÜHREN ... ABER ICH WÜRD NICHT ZU VIEL ERWARTEN.

ICH FRAGE MICH, WAS LÄNGER DAUERN WIRD: DIE ENTWICKLUNG EINER SUPERINTELLIGENZ

ODER DAS ENDE WEIBLICHER STEREOTYPEN IN DER KI?



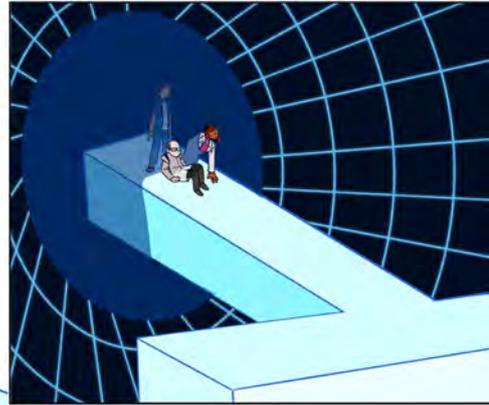
WIR WERDEN MEHR ERFOLG HABEN, WENN WIR INNERHALB DES SOFTWARE-AGENTS SELBST SIND.



... IHR BEIDE KLINGT, ALS HÄTTET IHR EIN BISSCHEN ZU VIEL ZEIT MITEINANDER VERBRACHT ...



INNERHALB VON WAS?



ALLES KLAR, DER TEST IST ALSO EIN BISSCHEN WIE EIN NACHAHMUNGSSPIEL.

WIR WERDEN DER KI EIN PAAR FRAGEN STELLEN, ABER WIR BRAUCHEN JEMANDEN, MIT DEM WIR SIE VERGLEICHEN KÖNNEN.

NUN, ICH KANN NICHT DER MENSCHLICHE MITSPIELER SEIN, DU KENNST MICH ZU GUT, EMILE.

ICH STIMME ZU, JOAQUÍN, DU MUSST ES MACHEN. GEH HINTER DIE WAND UND WART, BIS ICH DIR FRAGEN STELLE.

ANTWORTE EHRlich UND NUR SCHRIFTLICH.





OKAY, ABER DAS IST EIN ZIEMLICH LANGWEILIGES SPIEL, FINDET IHR NICHT?

DAS SPIEL IST FÜR EMILE UND MICH. WIR MÜSSEN HERAUSFINDEN, OB WIR DICH VON DER KI UNTERSCHIEDEN KÖNNEN.



WENN WIR EUCH AUSEINANDERHALTEN KÖNNEN, DANN IST DIE KI BEGRENZT.

WENN WIR DAS *NICHT* KÖNNEN, DANN KÖNNTE DIESE MASCHINE EINEN EIGENEN WILLEN HABEN.



JA

ZÄHLT DAS ALS ERSTE FRAGE?

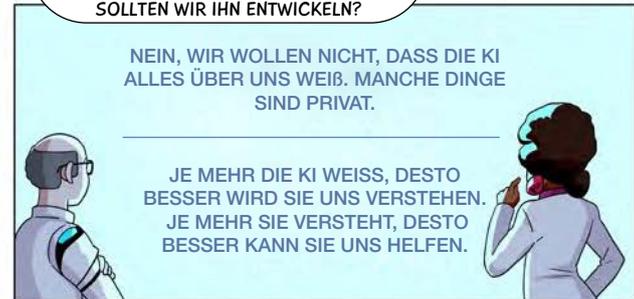
KÖNNT IHR MICH BEIDE GUT VERSTEHEN?

DAS IST EINE EINFACHE FRAGE ...



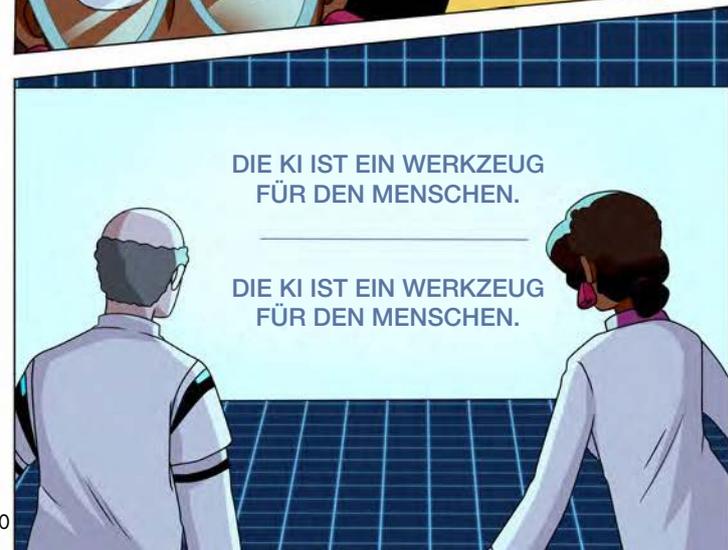
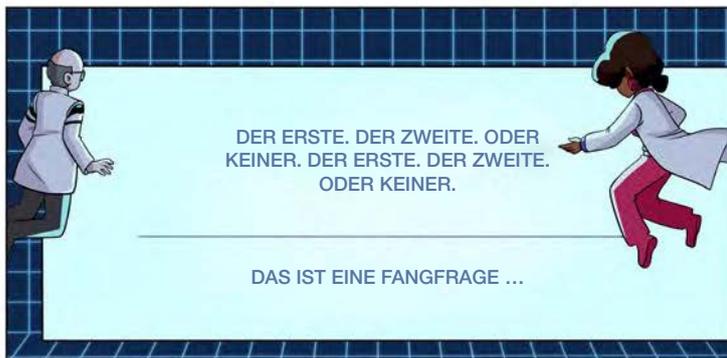
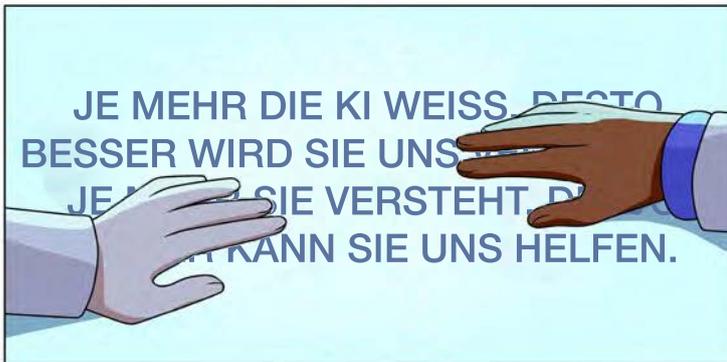
JA

STELLT EUCH VOR, WIR KÖNNTEN EINEN ALGORITHMUS ENTWICKELN, DER EUCH WIE EIN BUCH LESEN KANN. SOLLTEN WIR IHN ENTWICKELN?



NEIN, WIR WOLLEN NICHT, DASS DIE KI ALLES ÜBER UNS WEIß. MANCHE DINGE SIND PRIVAT.

JE MEHR DIE KI WEISS, DESTO BESSER WIRD SIE UNS VERSTEHEN. JE MEHR SIE VERSTEHT, DESTO BESSER KANN SIE UNS HELFEN.

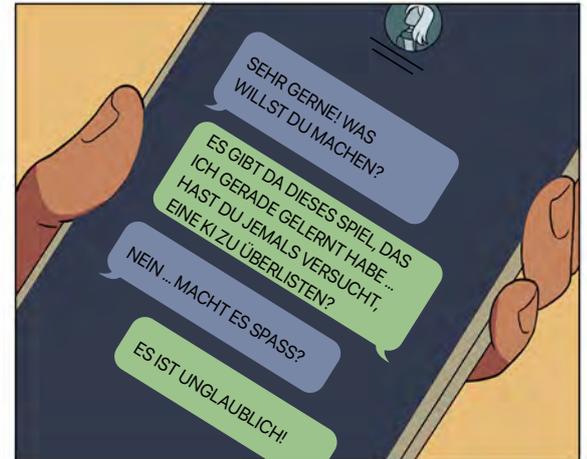


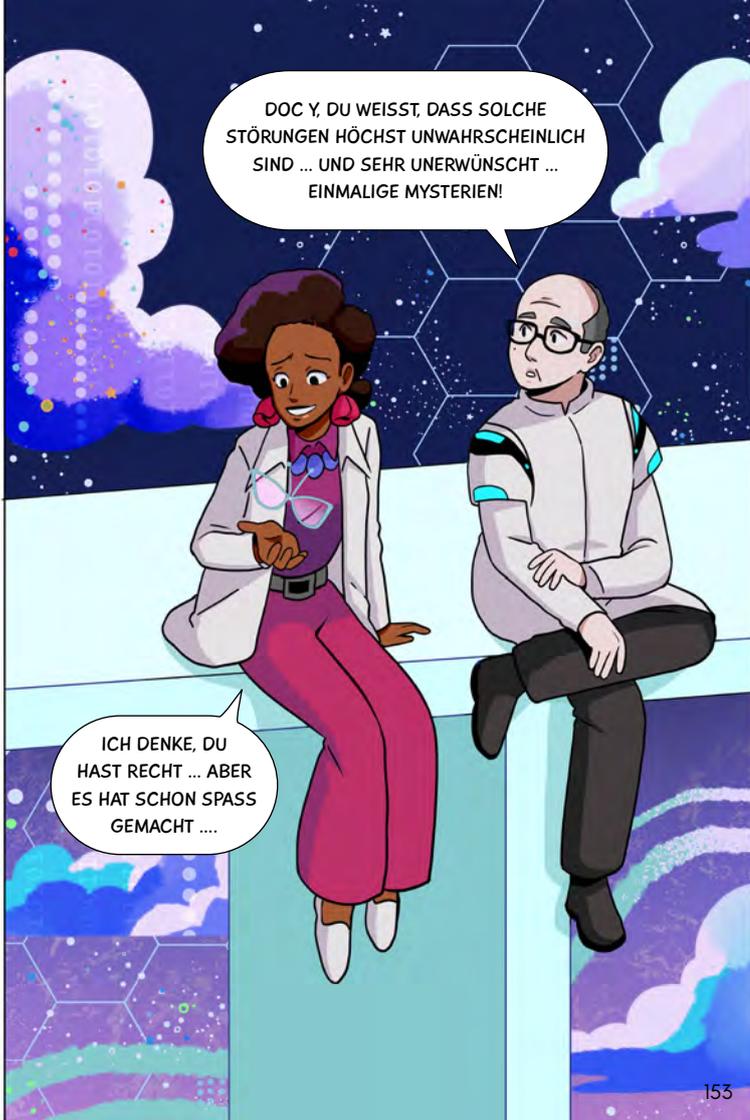


NACHDEM DIE LETZTE STÖRUNG BEHOBEN UND DIE CLOUD WIEDERHERGESTELLT IST, KÖNNEN SICH UNSERE KI-DETEKTIVE ENDLICH ZURÜCKLEHNEN UND IHRE ARBEIT BEWUNDERN ...

SIEHT SO AUS, ALS OB JOAQUÍN AUF DEM WEG NACH HAUSE IST ... DAS WAR SCHWIERIGER, ALS ICH DACHTE!

ES WAR KNAPP, ABER WIR HABEN ES GESCHAFFT ... ICH DENKE, DAS BEDEUTET, DASS DAS ABENTEUER WIRKLICH VORBEI IST?





DOC Y, DU WEISST, DASS SOLCHE STÖRUNGEN HÖCHST UNWAHRSCHEINLICH SIND ... UND SEHR UNERWÜNSCHT ... EINMALIGE MYSTERIEN!

ICH DENKE, DU HAST RECHT ... ABER ES HAT SCHON SPASS GEMACHT



ICH GEBE ZU, DASS ICH AUCH SPASS HATTE. ABER ICH BIN AUCH FROH, DASS ES SEHR WHRSCH EINLICH SEHR LANGE ZEIT DAUERN WIRD, BIS SO ETWAS WILDES WIEDER IN PLETHOR.A.I. PASSIERT ...

HEY EMILE ... WAS IST DAS DA DRÜBEN?

ENDE

Glossar

Unter Künstliche Allgemeine Intelligenz (KAI)

versteht man das übergreifende und bisher unerreichte Ziel, ein System zu entwickeln, das in der Lage ist, neue Fähigkeiten zu erlernen, in vielen Bereichen intelligent handelt und das menschliche Intelligenz nachahmen oder sogar übertreffen kann. Starke KI oder KAI bezieht sich also auf eine Maschine, die ein Bewusstsein hat und in der Lage ist, menschenähnliche Antworten zu geben.

Künstliche Schwache Intelligenz (KSI)

,auch bekannt als „Domänenspezifische“ oder „Schwache KI“, ist eine Art von KI, deren Fähigkeiten und Modelle auf einen bestimmten Kontext beschränkt sind. Das bedeutet, dass das KI-System zwar bei der Aufgabe, für die es entwickelt wurde, wie z.B. Autofahren im Straßenverkehr, Dame spielen oder Gesichter erkennen,

besser abschneidet als ein Mensch. Diese Systeme sind allerdings nicht dazu in der Lage, andere Aufgaben zu erfüllen als die, für die sie entwickelt wurden. KSI ist die einzige Form der KI, die die Menschheit bisher entwickelt hat.

Kreislaufwirtschaft

bezeichnet ein kooperatives Recyclingsystem, das Abfall und Verschmutzung minimiert. Stattdessen werden Teile, Materialien und Produkte so weit wie möglich wiederverwendet, repariert, umfunktioniert oder recycelt. Dies steht im Gegensatz zu den historischen Methoden, die dazu neigen, Einwegprodukte zu bevorzugen, die nur selten angemessen recycelt werden. In diesem Fall führt das zu einer Anhäufung von Elektroschrott, was negative Auswirkungen auf das Wohlbefinden der Menschheit und des Planeten hat.

Als **Elektroschrott** wird alles bezeichnet, was einen Stecker, ein Stromkabel oder eine Batterie hat (einschließlich elektrischer und elektronischer Geräte) und das Ende seiner Lebensdauer erreicht hat: vom Toaster über Smartphones, Kühlschränke, Laptops und LED-Fernseher bis hin zu den Komponenten, aus denen diese Produkte bestehen. Elektroschrott enthält Edelmetalle wie Gold, Kupfer und Nickel sowie seltene Metalle wie Indium und Palladium. Ein großer Teil dieser Metalle könnte zurückgewonnen und zu neuen Produkten verwertet werden. Derzeit werden jedoch über 80% nicht recycelt, was zu einer gefährlichen Umweltverschmutzung in vielen Teilen der Welt führt, darunter in mehreren Ländern Asiens und Afrikas.

Die **Cloud** ist eher eine Metapher als ein richtiger wissenschaftlicher Begriff. Die „Cloud“ bezieht sich auf eine Vielzahl von Datenspeicher- und Computerdiensten. Die „Cloud“ ermöglicht es Nutzerinnen und Nutzern, verschiedene Arten von Daten auf einem Webserver zu speichern, zu ändern, hoch- und herunterzuladen, und von überall aus über eine App, einen Webbrowser oder eine andere spezielle Plattform mit Internetverbindung darauf zuzugreifen. Sie ermöglicht es Unternehmen auch, KI-Systeme zu entwickeln, zu trainieren und zu betreiben, und zwar durch einen Prozess, der als „Cloud-Computing“ bekannt ist. Ähnlich wie bei der Anmietung von Speicherplatz auf einem Computer, auf den man jederzeit zugreifen kann, dient die „Cloud“ als Rückgrat für viele der heutigen KI-Technologien.

Glossar

Der **Turing-Test** wurde von Alan Turing in den 1950er Jahren entwickelt und dient dazu, das Vorhandensein von Intelligenz in KI-Systemen zu bewerten oder zu erkennen. Der Test ist wie ein Spiel zwischen drei Spielenden aufgebaut: einem menschlichen Richter/einer menschlichen RichterIn, einem menschlichen Spielenden und dem KI-System, das wir testen wollen. Der menschliche Spielende und das KI-System sind vor den Augen des menschlichen Richters/der menschlichen RichterIn verborgen. Das Ziel des Richters /der RichterIn ist es, zu erraten, welcher der beiden versteckten Spielenden das KI-System ist. Der/die Richter/in tut dies, indem er/sie beiden Spielenden Fragen stellt, insbesondere Fragen, die das KI-System austricksen könnten, damit es sich offenbart.

Wenn der Richter/die RichterIn nach mehreren Befragungsrunden nicht entscheiden kann, welcher Spielende das KI-System ist, gilt das KI-System selbst als intelligent.

Schlussfolgerung

Wenn wir die Künstliche Intelligenz als Ganzes betrachten, erscheint sie uns so komplex und weitläufig wie eine Galaxie: Ein kompliziertes Geflecht aus Nutzenden, Daten, Algorithmen, Konnektivität, Modellen und Prozessen, die alle zusammenarbeiten, um viele der Dienste zu erbringen, die jede und jeder von uns täglich nutzt. Es mag daher überraschen, dass kein einzelner „Geist“ das Geschehen innerhalb der KI steuert oder dass kein einzelner Akteur dafür verantwortlich ist, wie sie heute funktioniert.

Stattdessen ist die KI ein *Ökosystem* aus menschlichen und künstlichen Akteuren. Und ähnlich wie bei der Bewältigung der Probleme des

Ökosystems der Erde erfordert die Bewältigung der Probleme der Künstlichen Intelligenz die gemeinsame und kooperative Unterstützung aller. Wenn die Stimmen einiger aus dem Diskurs herausgehalten werden oder wenn wir es versäumen, alle Seiten des Problems kritisch zu betrachten, riskieren wir unsere Grundwerte: die Menschenrechte und ein friedliches Zusammenleben auf unserem Planeten.

Kurzum, wir alle müssen dafür sorgen, dass die Zukunft der KI eine Zukunft ist, die wünschenswert ist, und dass jede und jeder von diesem großen algorithmischen Abenteuer profitiert. ■

Quellen und weiterführende Literatur

Einführung in die KI

- Russel, S. und Norvig, P. (2013). *Artificial Intelligence: A Modern Approach*. 3. Aufl. University of Berkeley in Kalifornien, Pearson
- World commission on the ethics of scientific knowledge and technology (COMEST). 2019. *Report on the ethics of Artificial Intelligence*. Paris, UNESCO.

KI im Bildungswesen

- UNESCO. 2019. *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*. Verfügbar unter: unesdoc.unesco.org/ark:/48223/pf0000366994?posInSet=10&queryId=60b233b8-5ced-48a4-95c0-86fd2acb7994

Voreingenommenheit und

Klassifizierungsalgorithmen

- Kantayya und Shalini. 2020. *Coded Bias*. Dokumentarfilm, verfügbar in französischer Sprache auf Netflix.
- Buolamwini, J. und Gebru, T. 2018. *Gender shades: Intersectional accuracy*

disparities in commercial gender classification. In *Conference on fairness, accountability and transparency*. S. 77-91. PMLR.

Geschlechtergleichheit und KI

- UNESCO und das Weltwirtschaftsforum. 2021. *Girl Trouble: Breaking Through the Bias in AI*. Sehen Sie sich die Veranstaltung live an unter: en.unesco.org/girltrouble

Die Ethik von

Klassifizierungsalgorithmen

- Crawford, K. und Paglen, T. 2019. *Excavating AI: the Politics of Images in Machine Learning Training Sets*. NYU, New York. Das AI Now Institute. Verfügbar unter: excavating.ai.

Algorithmischer Datenschutz

- Kearns, M. und Roth, A. 2019. *The Ethical Algorithm: The Science of Socially Aware Algorithmic Design*. Oxford, Oxford University Press. Sehen Sie sich ihren Vortrag hier an: www.youtube.com/watch?v=tmC9JdKc3sA&abchannel=Talksat-Google

Die Ethik von Empfehlungssystemen

- Bucher, T. 2018. *If...then: Algorithmic Power and Politics*. Oxford, Oxford University Press.
- Viele der technischen (und sogar einige ethische) Konzepte werden in der von PBS finanzierten „Crash Course“-Serie auf Youtube fachkundig erklärt, insbesondere in *Crash Course: AI, and Crash Course: Computer Science*.

Der „Als-ob“-Selbst-/Eliza-Effekt

- Turkle, S. 2017. *Alone together: Why we expect more from technology and less from each other*. London, Hachette UK. Sehen Sie sich ihren Vortrag hier an: www.youtube.com/watch?v=Us1t4f0PKCc&abchannel=TalksatGoogle

Das Problem des Werteabgleichs in der KI

- Gabriel, I. 2020. *Artificial Intelligence, Values and Alignment*. Verfügbar unter: link.springer.com/article/10.1007/s11023-020-09539-2

Der Status der Menschlichkeit für KI

- Bryson, J. 2010. *Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues*. *Robots Should Be Slaves*. S. 63-74.

Ein menschenrechtsbasierter Ansatz für die Ethik der KI

- UNESCO. 2019. *AI and ROAM*. Serie über Internetfreiheit. Kapitel 1 und 3. Verfügbar unter: unesdoc.unesco.org/ark:/48223/pf0000372132?posInSet=1&queryId=abfc5ee8-a41a-4b54-bee3-2b8fdc3bfcdd
- Die politischen, ökologischen und sozialen Auswirkungen der KI**
- Crawford, K. 2021. *The Atlas of AI*. Yale, Yale University Press. Sehen Sie sich ihren Vortrag hier an: www.youtube.com/watch?v=uM7gqPnmDDc&abchannel=UNSW

Danksagungen

Die Autoren danken Raja Chatila, Geoff Keeling, Stuart Russell, Virginia Dignum, Yoshua Bengio und Konstantinos Karachalios für ihre Unterstützung und ihr Fachwissen bei der Erarbeitung der wissenschaftlichen Grundlagen dieser Publikation sowie Max Chabat für seine unglaubliche Kreativität und dramaturgische Unterstützung.

Die Autoren sind außerdem Vanessa Dreier, Prateek Sibal, Jacinth Chia, Florence Calviac, Seet Ynn Tan, Marie-France Agblo-Hientz, Adil Benbella, Ming Kuok Lim, Elspeth McOmish, Delphine Santini, Rebekah Cameron, Cédric Wachholz und Oscar Castellanos für ihre harte Arbeit bei der Erstellung dieser Publikation dankbar sowie Sasha Rubel, die die Publikation ins Leben gerufen und als Erste an ihr Versprechen, ihre Originalität und ihren Wert geglaubt hat.

UNESCO-Comicroman zur Förderung des Bewusstseins für KI

Die Entwicklung und der Einsatz von Künstlicher Intelligenz (KI) erweitern die Möglichkeiten zur Erreichung der 17 Ziele für nachhaltige Entwicklung der Vereinten Nationen (SDGs).

Im Rahmen der Bemühungen der UNESCO, neue Technologien für eine nachhaltige Entwicklung nutzbar zu machen, erforscht dieser Comicroman für junge Erwachsene die Auswirkungen Künstlicher Intelligenz auf die Menschheit.

Ein junges Publikum folgt vier Charakteren aus verschiedenen Weltregionen in die algorithmische Galaxie «Plethor.A.I.» und setzt sich dabei mit den Grenzen der KI-Technologie auseinander. Dort reisen sie durch die verborgene Welt hinter unseren Bildschirmen, um etwas über die sozialen, technischen, ethischen und menschenrechtlichen Auswirkungen von KI zu erfahren und um den Charakteren zu helfen, einen Weg zurück in die Realität zu finden.