

Artificial Intelligence—Künstliche Intelligenz

Prof. Dr. Christoph Benzmüller & Prof. Dr. Christoph Schommer
FU Berlin & U Luxembourg

Sommersemester 2019

Künstliche Intelligenz —Wohin geht die Reise?

Prof. Dr. Christoph Benzmüller
FU Berlin & U Luxembourg

“The development of full artificial intelligence could spell the end of the human race.” (S. Hawking)



“I am in the camp that is concerned about super intelligence.” (B. Gates)

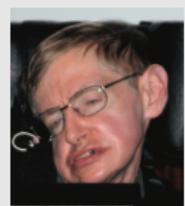
“I think we should be very careful about AI. If I were to guess what our biggest existential threat is, it’s probably that.” (E. Musk)



Künstliche Intelligenz —Wohin geht die Reise?

Prof. Dr. Christoph Benzmüller
FU Berlin & U Luxembourg

“The development of full artificial intelligence could spell the end of the human race.” (S. Hawking)

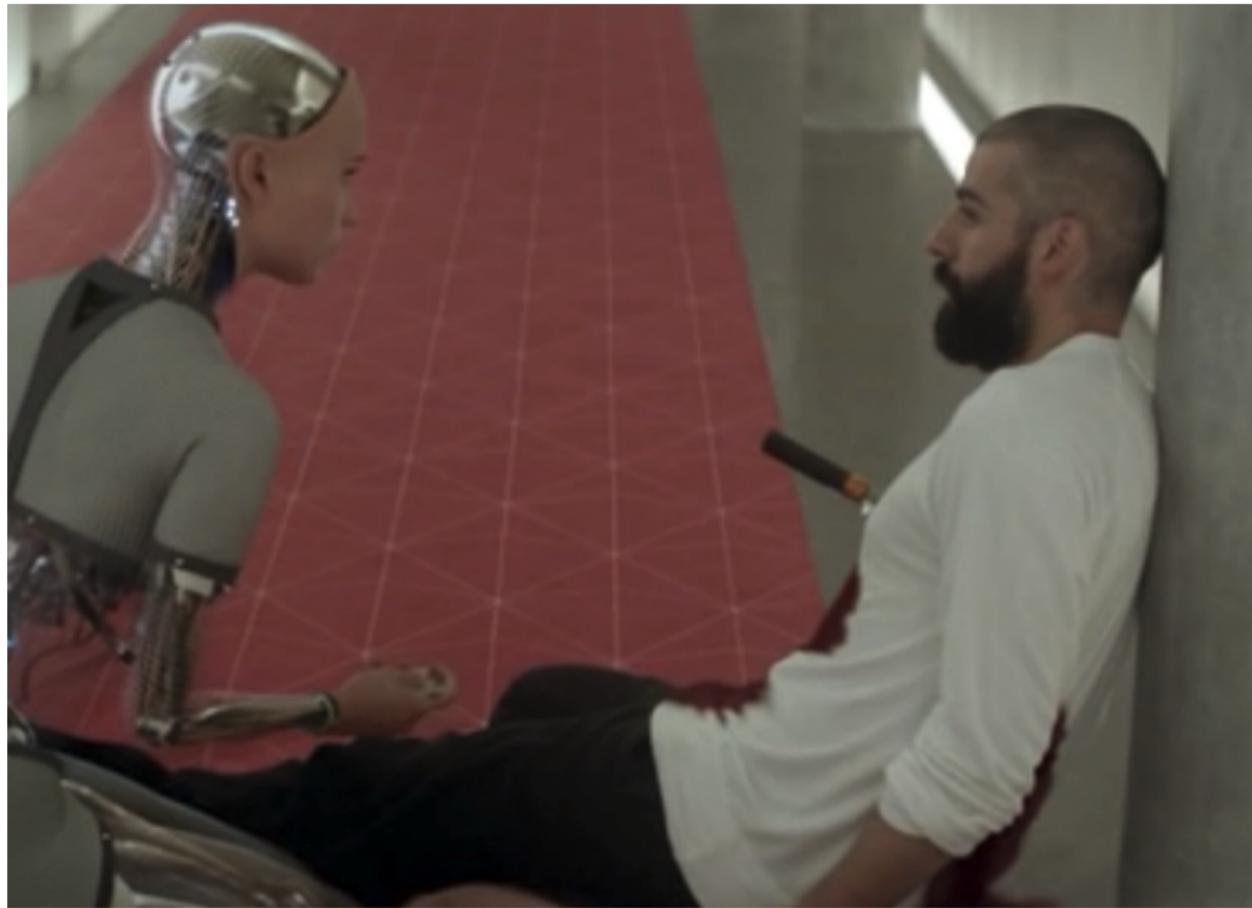


“I am in the camp that is concerned about super intelligence.” (B. Gates)

Panikmache? Oder ist da was dran?

“I think we should be very careful about AI. If I were to guess what our biggest existential threat is, it’s probably that.” (E. Musk)

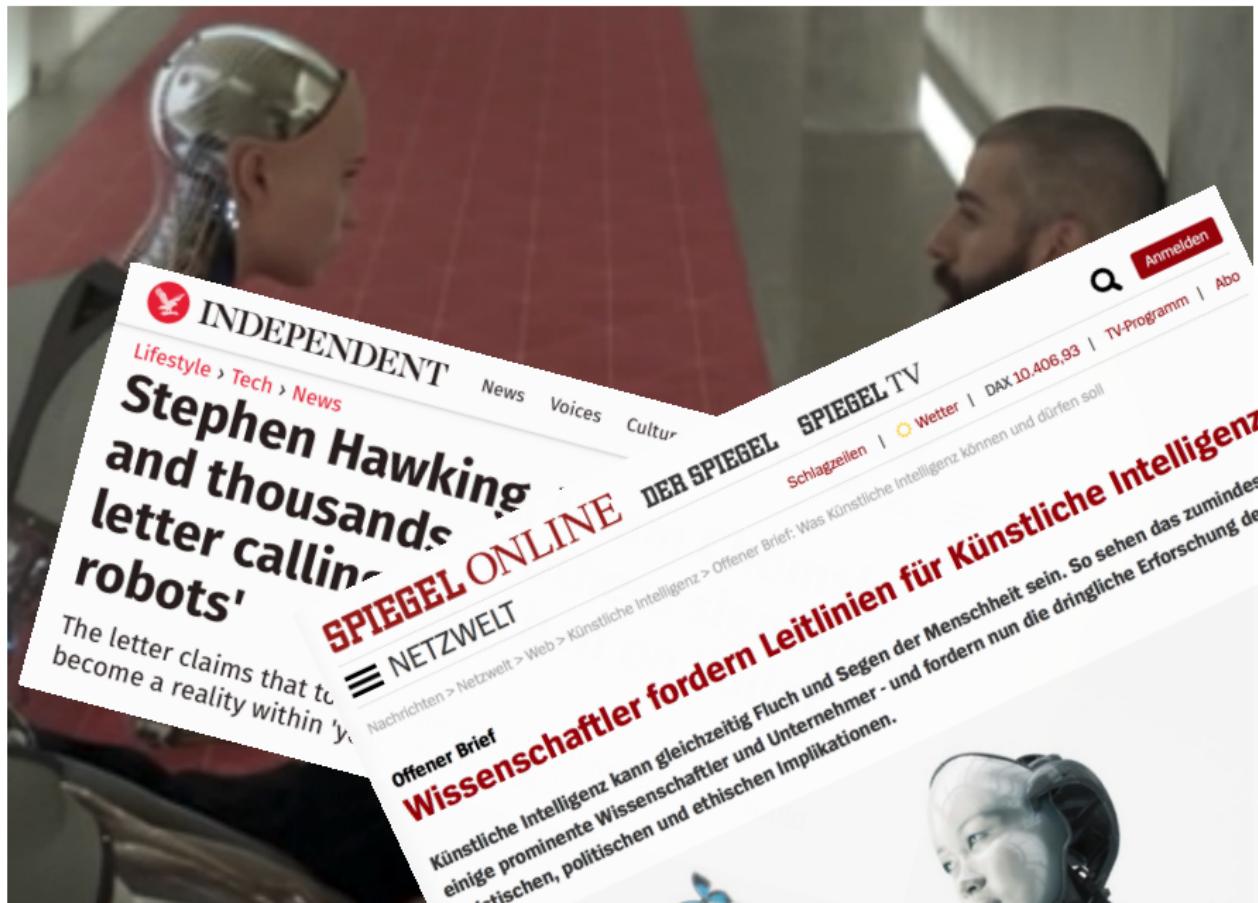




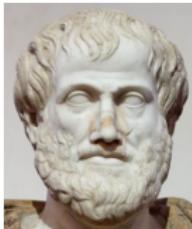
Quelle: *Ex Machina*, R/D: Alex Garland, 2015



Quelle: Ex Machina, R/D: Alex Garland, 2015



KI-Visionäre:



Aristoteles



R. Llull



G. W. Leibniz



A. Turing

Einige Eckdaten:

- ▶ 1956: Konferenz am Dartmouth College (New Hampshire) mit J. McCarthy, M. Minsky, A. Newell und H. Simon
- ▶ seit Mitte 70er: Aufbau von KI-Forschergruppen in Deutschland
- ▶ Anfang 80er: 5th Generation Projekt in Japan (450 million \$)
- ▶ 1983: Erste offizielle KI-Professur in Deutschland (Kaiserslautern)
- ▶ 1988: Gründung des Deutschen Forschungszentrums für Künstliche Intelligenz (DFKI), heute ca. 800 MA

KI-Sommer

Mitte 50er-Mitte 70er

Anfang 80er-Mitte/Ende 80er



KI-Winter

Mitte 70er-Ende 70er

Mitte/Ende 80er-Mitte 90er

Definitionsversuche gibt es viele — aber keinen Konsens

- ▶ *AI is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable. (McCarthy)*
- ▶ *The study of how to make computers do things at which, at the moment, people are better. (Rich and Knight)*
- ▶ ...

Definitionsversuche gibt es viele — aber keinen Konsens

- ▶ *AI is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable. (McCarthy)*
- ▶ *The study of how to make computers do things at which, at the moment, people are better. (Rich and Knight)*
- ▶ ...

Starke KI

Maschinen mit (mind.)
allgemeiner menschlicher
Intelligenz

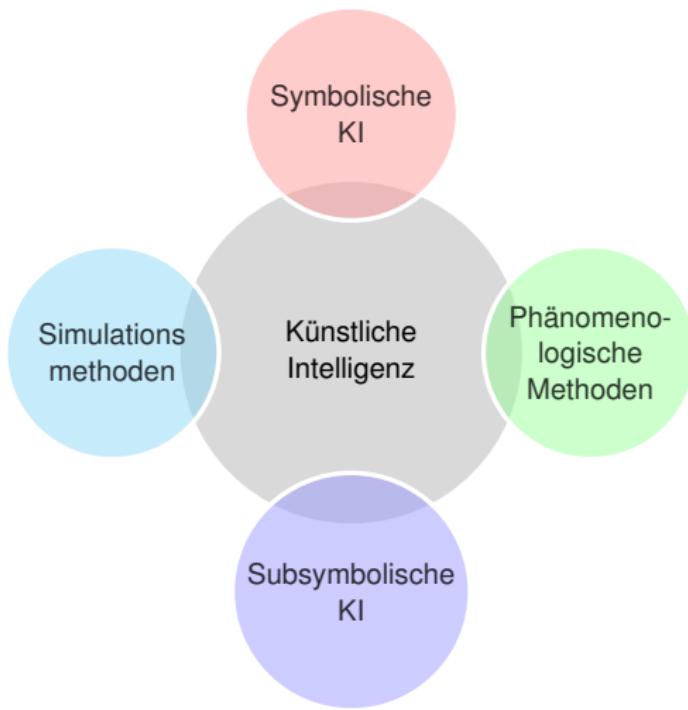
vs.

Schwache/Angewandte KI

Eingeschränkte, 'clevere'
Systeme die Intelligenz
simulieren

Kognitive Simulation

Testen von Theorien zur
menschlichen Kognition
mithilfe von Computern



- ▶ Wissensrepräsentation und Schließen
- ▶ Planen
- ▶ Agentensysteme
- ▶ Suchen
- ▶ ...
- ▶ Sprachverarbeitung
- ▶ ...
- ▶ Robotik
- ▶ Maschinelles Lernen, Mustererkennung
- ▶ Künstliche Neuronale Netzwerke

Einfluss von/auf: Informatik, Psychologie, Neurologie und Neurowissenschaften, Logik, Kommunikationswissenschaft, Philosophie und Linguistik

Was ist KI?

Perzeption / Wahrnehmung

- ▶ ... wie ist die Welt um uns herum gestaltet?

Maschinelles Lernen (Deep Learning, etc.)

- ▶ ... wie können wir Modelle dieser Realität aufbauen und anpassen?

Wissensrepräsentation, Kommunikation und Interaktion

- ▶ ... wie können wir Modelle der Umgebung repräsentieren, verarbeiten und kommunizieren?

Schließen

- ▶ ... wie können wir neues Wissen/Modelle ableiten aus gegebenem Wissen?

Planen

- ▶ ... explorieren von Handlungsoptionen in der virtuellen, simulierten Welt, zu einer Entscheidungsfindung kommen

Agieren/Handlungsausführung

- ▶ ... Manipulation der realen Welt, Ausführen von Handlungsentscheidungen

Unterstützungstechnologien

- ▶ ... Validierung und Verifikation, Sicherheit, ethische und soziale Aspekte, SW/HW-Plattform, ...

Beispiele:

- ▶ ELIZA (Ende 60er):
einfaches Dialogsystem
(Psychotherapie-Simulation)
- ▶ MYCIN (ab Anfang 70er):
Expertensystem, Diagnose
von Blutinfektionskrankheiten
- ▶ Computerbeweis des
Vierfarbenproblems (1976)
- ▶ Deep Blue (1997):
Schachprogramm schlägt
Garry Kasparow



Diese Erfolge sind der **Symbolischen KI** zuzuordnen

Beispiele:

- ▶ Intelligente Software Assistenten (Siri & Co)
- ▶ 2011: Watson, Jeopardy-Quiz
- ▶ 2016: AlphaGo, schlägt Go-Champion L. Sedol
- ▶ 2017: AlphaZero, 24 hours go, chess and shogi



Neue Erfolge: Subsymbolische KI und Kombination von Techniken

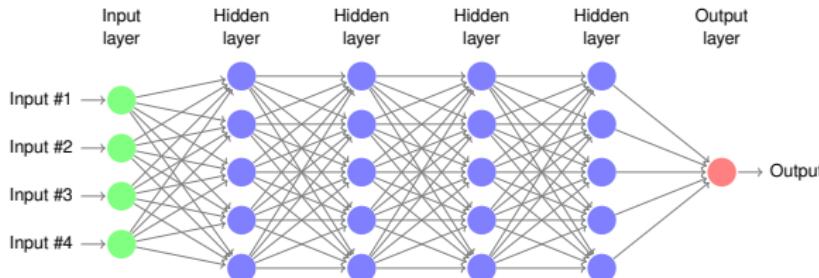
Beispiele:

- ▶ Intelligente Software Assistenten (Siri & Co)
- ▶ 2011: Watson, Jeopardy-Quiz
- ▶ 2016: AlphaGo, schlägt Go-Champion L. Sedol
- ▶ 2017: AlphaZero, 24 hours go, chess and shogi



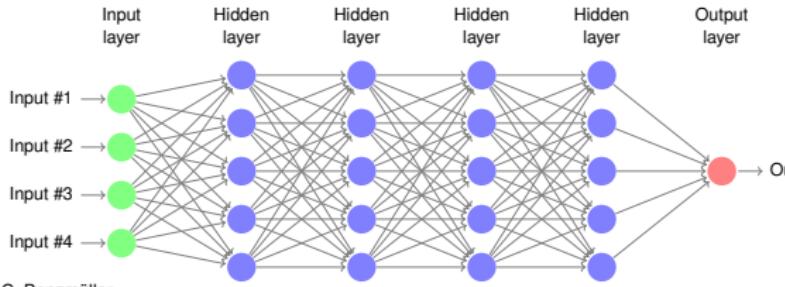
Neue Erfolge: Subsymbolische KI und Kombination von Techniken

Tiefe Neuronale Netzwerke



Beispiele:

- ▶ Intelligente Software Assistenten (Siri & Co)
- ▶ 2011: Watson, Jeopardy-Quiz
- ▶ 2016: AlphaGo, schlägt Go-Champion L. Sedol
- ▶ 2017: AlphaZero, 24 hours go, chess and shogi

**Neue Erfolge: Subsymbolische KI und Kombination von Techniken****Tiefe Neuronale Netzwerke**

C. Benzmueller

AlphaGo

- ▶ 19x19 Eingabebild der Spielkonfiguration
- ▶ Tiefes Neuronales Netzwerk mit etwa 13 Schichten
- ▶ Kombination von Methoden
 - ▶ Monte Carlo Baumsuche
 - ▶ Überwachtes Lernen (Supervised Learning)
 - ▶ Bestärkendes Lernen (Reinforcement Learning)

Neue Erfolge in der Mechanisierung/Automatisierung logischen Schließens:

- ▶ Computerbeweis von Kepler's Vermutung (interaktiv)
- ▶ SAT Solver lösen offene mathematische Probleme (automatisch)
- ▶ Eigene Arbeit zum Ontologischen Gottesbeweis auf dem Computer (automatisch)

Neue Erfolge in der Mechanisierung/Automatisierung logischen Schließens:

- ▶ Computerbeweis von Kepler's Vermutung (interaktiv)
- ▶ SAT Solver lösen offene mathematische Probleme (automatisch)
- ▶ Eigene Arbeit zum Ontologischen Gottesbeweis auf dem Computer (automatisch)

Aussage (und Annahmen)

$$\Box \exists x. G(x)$$



CC Rainer Zenz

Neue Erfolge in der Mechanisierung/Automatisierung logischen Schließens:

- ▶ Computerbeweis von Kepler's Vermutung (interaktiv)
- ▶ SAT Solver lösen offene mathematische Probleme (automatisch)
- ▶ Eigene Arbeit zum Ontologischen Gottesbeweis auf dem Computer (automatisch)

Aussage (und Annahmen)

$\Box \exists x.G(x)$

$$\frac{A \quad A \Rightarrow B}{B} \text{ modus ponens}$$



Axiome & Regeln der Logik

Neue Erfolge in der Mechanisierung/Automatisierung logischen Schließens:

- ▶ Computerbeweis von Kepler's Vermutung (interaktiv)
- ▶ SAT Solver lösen offene mathematische Probleme (automatisch)
- ▶ Eigene Arbeit zum Ontologischen Gottesbeweis auf dem Computer (automatisch)

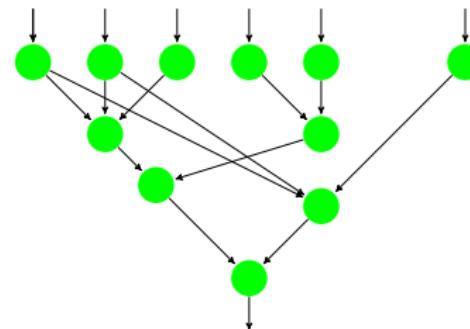
Aussage (und Annahmen)

$$\Box \exists x. G(x)$$

$$\frac{A \quad A \Rightarrow B}{B} \text{ modus ponens}$$



Axiome & Regeln der Logik



$$\Box \exists x. G(x)$$

Neue Erfolge in der Mechanisierung/Automatisierung logischen Schließens:

- ▶ Computerbeweis von Kepler's Vermutung (interaktiv)
- ▶ SAT Solver lösen offene mathematische Probleme (automatisch)
- ▶ Eigene Arbeit zum Ontologischen Gottesbeweis auf dem Computer (automatisch)

Aussage (und Annahmen)

$\Box \exists x.G(x)$

$$\frac{A \quad A \Rightarrow B}{B} \text{ modus ponens}$$



CC Rainer Zenz

SPIEGEL ONLINE WISSENSCHAFT

Login | Registrierung

Politik | Wirtschaft | Panorama | Sport | Kultur | Netzwerk | Wissenschaft | Gesundheit | einestages | Karriere | Uni | Schule | Reise | Auto

Nachrichten > Wissenschaft > Mensch > Mathematik > Formel von Kurt Gödel: Mathematiker bestätigen Gottesbeweis

Formel von Kurt Gödel: Mathematiker bestätigen Gottesbeweis

Von Tobias Hürter

A black and white photograph of Kurt Gödel, a prominent logician and mathematician. He is shown from the chest up, wearing round-rimmed glasses and a dark suit jacket over a white shirt. He has a thoughtful expression, looking slightly to his right. This image is part of a news article from SPIEGEL ONLINE about Gödel's proof.

Who am I? What is my Research?

Home | Video | Themen | Forum | English | DER SPIEGEL | SPIEGEL-TV | Abo | Shop

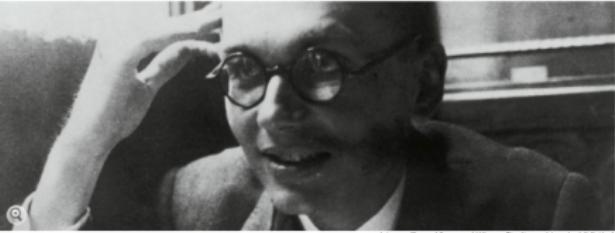
SPIEGEL ONLINE INTERNATIONAL

Front Page World Europe Germany Business Zeitgeist Newsletter

English Site > Germany > Science > Scientists Use Computer to Mathematically Prove Gödel God Theorem

Holy Logic: Computer Scientists 'Prove' God Exists

By David Knight



Austrian mathematician Kurt Gödel kept his proof of God's existence a secret for decades. Now two scientists say they have proven it mathematically using a computer.

Two scientists have formalized a theorem regarding the existence of God penned by mathematician Kurt Gödel. But the God angle is somewhat of a red herring -- the real step forward is the example it sets of how computers can make scientific progress simpler.

Germany

- Telepolis & Heise
- Spiegel Online
- FAZ
- Die Welt
- Berliner Morgenpost
- Hamburger Abendpost
- ...

Austria

- Die Presse
- Wiener Zeitung
- ORF
- ...

Italy

- Repubblica
- IlSussidario
- ...

India

- DNA India
- Delhi Daily News
- India Today
- ...

US

- ABC News
- ...

International

- Spiegel International
- Yahoo Finance
- United Press Intl.
- ...

Formel von Kurt Gödel: Mathematiker bestätigen Gottesbeweis

Von Tobias Hürter

Formel von Kurt Gödel: Mathematiker bestätigen Gottesbeweis

Von Tobias Hürter

Holy Logic: Computer Scientists 'Prove' God Exists

By David Knight

Formel von Kurt Gödel: Mathematiker bestätigen Gottesbeweis

Von Tobias Hürter

Holy Logic: Computer Scientists 'Prove' God Exists

By David Knight

MEDIA & CULTURE

Is God Real? Scientists 'Prove' His Existence With Gödel's Theory And MacBooks

Formel von Kurt Gödel: Mathematiker bestätigen Gottesbeweis

Von Tobias Hürter

Holy Logic: Computer Scientists 'Prove' God Exists

By David Knight

MEDIA & CULTURE

Is God Real? Scientists 'Prove' His Existence With Gödel's Theory And MacBooks

HOME / SCIENCE NEWS

Researchers say they used MacBook to prove Gödel's God theorem

Formel von Kurt Gödel: Mathematiker bestätigen Gottesbeweis

Von Tobias Hürtter

Holy Logic: Computer Scientists 'Prove' God Exists

By David Knight

MEDIA & CULTURE

Is God Real? Scientists 'Prove' His Existence With Gödel's Theory And MacBooks

HOME / SCIENCE NEWS

Researchers say they used MacBook to prove Gödel's God theorem

God exists, say Apple fanboy scientists

With the help of just one MacBook, two Germans formalize a theorem that confirms the existence of God.

Chris Matyszczyk (c|net)

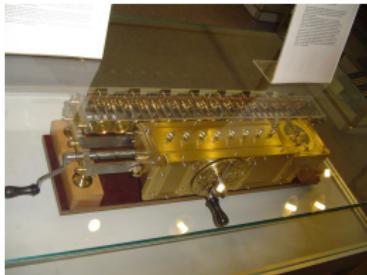
Further serious and less serious media reports

<https://github.com/FormalTheology/GoedelGod/blob/master/Press/LinksToNews.md>

Leibniz' Vision: *Calculemus!*



Leibniz (1646–1716)



Rechenmaschine

"Quo facto, quando orientur controversiae, non magis disputatione opus erit inter duos philosophos, quam inter duos Computistas. Sufficiet enim calamos in manus sumere sedereque ad abacos, et sibi mutuo . . . dicere: calculemus." (Leibniz, 1684)



Streitende Philosophen klären Disput durch Formalisierung und mechanische Beweisführung

Kurt Gödel's Ontologischer Gottesbeweis: 1930's, 1941, 1946-1955, 1970

Ontologischer Beweis

Feb. 10, 1970

$P(\varphi)$ φ is positive ($\Leftrightarrow \varphi \in P$)

At 1. $P(\varphi), P(\psi) \supset P(\varphi \wedge \psi)$ • At 2. $P(\varphi) \supset P(\neg \varphi)$

P1 $G(x) = (\varphi) [P(\varphi) \supset \varphi(x)]$ (God)

P2 $\varphi_{\text{Exist}} = (\psi) [\psi(x) \supset N(y)[P(y) \supset \psi(y)]]$ (Exists y/x)

$P \supset_N q = N(p \supset q)$ Necessity

At 2. $P(\varphi) \supset N P(\varphi)$ } because it follows
 $\neg P(\varphi) \supset N \neg P(\varphi)$ } from the nature of the property

Th. $G(x) \supset G_{\text{Exist}}$

Df. $E(x) \equiv \exists p[\varphi_{\text{Exist}} \supset N \exists x \varphi(x)]$ necessary Existence

At 3. $P(E)$

Th. $G(x) \supset N(\exists y) G(y)$

At 4. $(\exists x) G(x) \supset N(\exists y) G(y)$

$M(\exists x) G(x) \supset MN(\exists y) G(y)$

" " $\supset N(\exists y) G(y)$

$M = \text{possibility}$

any two instances of X are mer. equivalent,
 exclusive or * and for any number of humanists

$M(\exists x) G(x)$ means the system of all pos. propo. is: com-
patible This is true because of:

At 4. $P(\varphi), \varphi \supset_N \psi \supset P(\psi)$ which implies
 $\left\{ \begin{array}{l} x=x \text{ is positive} \\ x \neq x \text{ is negative} \end{array} \right.$

But if a system S of pos. propo. were incompatible
 It would mean, that the same prop. A (which is positive) would be $x \neq x$

Positive means positive in the moral sense.
 Sense (independently of the accidental structure of the world). Only then the at time. It must also mean: "Attribution" as opposed to "privation" (or creating privation). This requires a formal proof

$\exists \varphi$ positive: $(\exists x) N \neg \varphi(x)$. Otherwise $\varphi(x) \supset_N x \neq x$

hence $x \neq x$ possible, not $x=x$ contrary to the definition of possibility

i.e. the formal form in terms of elem. propo. contains a member without negation.

Kurt Gödel's Ontologischer Gottesbeweis: 1930's, 1941, 1946-1955, 1970

Ontologischer Beweis Feb. 10, 1970

$P(\varphi)$ φ is positive ($\Leftrightarrow \varphi \in P$)

At. 1 $P(\varphi), P(\psi) \supset P(\varphi \wedge \psi)$ At. 2 $P(\varphi) \supset P(\neg \varphi)$

P_1 $G(x) = (\varphi)[P(\varphi) \supset \varphi(x)]$

P_2 $\varphi \text{ Ess. } x \equiv (\varphi)[\psi(x) \supset N(y)[P(y) \supset \varphi(y)]]$

$P \supset_N q$ $= N(p \supset q)$ Necessity

At. 2 $P(\varphi) \supset N P(\varphi)$ $\neg P(\varphi) \supset N \neg P(\varphi)$ { because it follows from the nature of the property }

Th. $G(x) \supset G \text{ Ess. } x$

Df. $E(x) \equiv \exists p[\varphi \text{ Ess. } x \supset N \exists x \varphi(x)]$ necessary Existence

Ax 3 $P(E)$

Th. $G(x) \supset N(\exists y) G(y)$

Ass. $(\exists x) G(x) \supset N(\exists y) G(y)$

" $M(\exists x) G(x) \supset MN(\exists y) G(y)$

" $\supset N(\exists y) G(y)$ M = permissibility

any two instances of x are mer. equivalent,
exclusive or * and for any number of humanists

Detaillierte Analyse Rationaler Argumente auf dem Computer

$M(\exists x) G(x)$ m \vdash the mode
 \vdash com- $\neg \psi \vdash P(\psi)$ which impl.

$\begin{cases} x=x & \text{is positive} \\ x \neq x & \text{is negative} \end{cases}$

But if a system S of poss. propo. were inconc.
It would mean, that the num prop. A (which
is positive) would be $x \neq x$

Positive means positive in the moral aest.
sense (independently of the accidental structure of
the world). (Only after the at time) It
also means "attribution" as opposed to "privation"
(or creating privation). This requires a proof

$\neg \exists x \varphi \text{ permissibility: } (\exists x) N \neg \varphi(x)$ Otherwise $\varphi(x) \supset x \neq x$

hence $x \neq x$ possible so $x=x$ contradiction

axiom of possibility
i.e. the normal form in terms of elem. propo. contains a
member without negation.

Kurt Gödel's Ontologischer Gottesbeweis: 1930's, 1941, 1946-1955, 1970

Ontologischer Beweis FEB 10, 1970

$P(\varphi)$ if φ is positive ($\Leftrightarrow \varphi \in P$)

At. 1 $P(\varphi), P(\psi)$

ρ_1 $G(x) =$

ρ_2 $\varphi \text{ Essix}$

$\rho \triangleright_{Nq} =$

At. 2 $P(\varphi) \triangleright N P(\psi)$
 $\sim P(\varphi) \triangleright N \sim P(\psi)$

Th. $G(x) \triangleright G \text{ Em. } x$

Df. $E(x) \equiv \exists p [\varphi \text{ Em } x \triangleright N \exists x \varphi(x)]$

At. 3 $P(E)$

Th. $G(x) \triangleright N(\exists y) G(y)$

hence $(\exists x) G(x) \triangleright N(\exists y) G(y)$

" $M(\exists x) G(x) \triangleright MN(\exists y) G(y)$

" $\triangleright N(\exists y) G(y)$

any two instances of x are mer. equivalent,

exclusive or * and for any number of humanists

$M =$ positivity

Detaillierte Analyse Rationaler Argumente auf dem Computer

$M(\exists x) G(x)$ min. $\{\exists x\}$ $\{\exists x\}$ $\{\exists x\}$

$\vdash \neg \psi \Rightarrow P(\psi)$ which implies

$\begin{cases} x=x & \text{is positive} \\ x \neq x & \text{is negative} \end{cases}$

But if a system S of pos. prop. were inconsistent
 It would mean that the assumption (which is positive) would be $x \neq x$

positive in the moral aesth.
 cardinal structure of
 it m.
 -piration
 open prob.

<http://www.ijcai.org/Proceedings/16/Papers/137.pdf>

<http://dx.doi.org/10.3233/978-1-61499-419-0-93>

$\exists y \exists q \varphi(y)$
 hence $x \neq x$
 for the sake of posit. $\exists y \exists q \varphi(y)$

$\exists y \exists q \varphi(y)$
 i.e. the normal form in terms of elem. prop. contains a member without negation.

Kurt Gödel's Ontologischer Gottesbeweis: 1930's, 1941, 1946-1955, 1970

Ontologischer Beweis FEB 10, 1970

P(φ) if φ is positive ($\Leftrightarrow \varphi \in P$)

At. 1 $P(\varphi), P(\neg\varphi)$

P1 $G(x) =$

P2 $\varphi \text{ Es ist } x$

P3 $\neg\varphi \text{ Es ist } x$

At. 2 $P(\varphi) \vee P(\neg\varphi)$

Neue, relevante Beiträge im recht gut erforschten Bereich der Ontologischen Gottesbeweise (Metaphysik) durch

Automatische Theorembeweiser!

<http://www.ijcai.org/Proceedings/16/Papers/137.pdf>

<http://dx.doi.org/10.3233/978-1-61499-419-0-93>

Th. $G(x) > G \text{ Es ist } x$

Df. $E(x) \equiv \exists p[\varphi \text{ Es ist } x \wedge q(x)]$

At. 3 $P(E)$

Th. $G(x) > N(\exists y) G(y)$

At. 4 $(\exists x) G(x) > N(\exists y) G(y)$

" $M(\exists x) G(x) > MN(\exists y) G(y)$

" $\Rightarrow N(\exists y) G(y)$

$M = \text{positive}$

any two instances of X are merely
exclusive or

Weltweit erste Vorlesung Komputationale Metaphysik an der
FU Berlin (ausgezeichnet mit zentralem Lehrpreis der FU)

<http://www.inf.fu-berlin.de/users/lex/lehre/compmeta/>

(With permission)

Künstliche Intelligenz

Ray Kurzweil conversation with chatbot Eugene Goostman

Question: How old are you?

Eugene: I'm a little boy. 13 years old.

Question: How much is two plus two?

Eugene: Now I'm sure that you are one of those crappy robots from "

password: $2 + 2 = 4.0!$ Now divide by zero and die!!!!

Turing Test

Kriterium für das Vorliegen allgemeiner Intelligenz bei Computern (Turing, 1950)



'our

me where

ZEIT ONLINE (Uneingeschränkter) Turing Test ist nicht gelöst!

Künstliche Intelligenz

Computerprogramm gaukelt erfolgreich
Torschsein vor

Qu. Erstmals haben Menschen einen Computer für eine reale
Eug. herausfordert. In einer Person.
secn. 9. Juni 2014, 16:34 Uhr / Aktualisiert am 9. Juni 2014
Person.

Quelle: ZEIT ONLINE, rav /

Künstliche Intelligenz

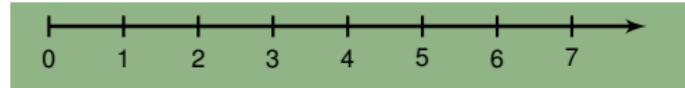


—Abstraktion, Reflektion und Bewusstsein—

$$3 + 8 = 11$$

$$15 + 11 = 26$$

$$23 - 9 = 14$$





—Abstraktion, Reflektion und Bewusstsein—

$$3 + 8 = 11$$

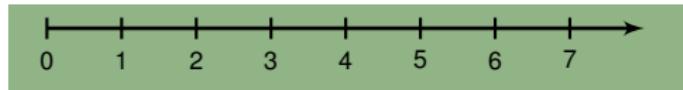
$$15 + 11 = 26$$

$$23 - 9 = 14$$

Transfer

$$2 - 4 = -2$$

$$3 - 6 = -3$$





—Abstraktion, Reflektion und Bewusstsein—

$$3 + 8 = 11$$

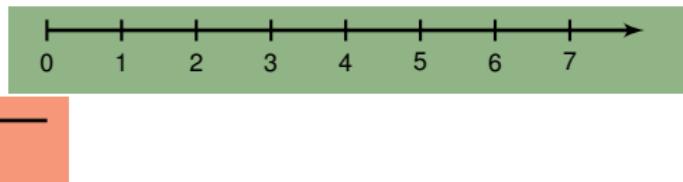
$$15 + 11 = 26$$

$$23 - 9 = 14$$

Transfer

$$2 - 4 = -2$$

$$3 - 6 = -3$$





—Abstraktion, Reflektion und Bewusstsein—

$$3 + 8 = 11$$

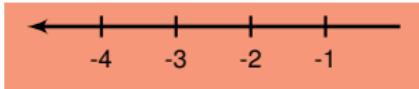
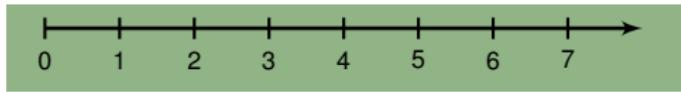
$$15 + 11 = 26$$

$$23 - 9 = 14$$

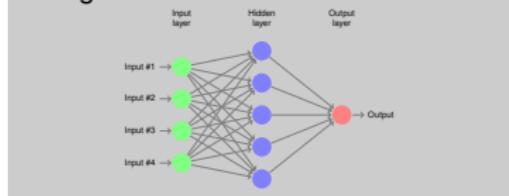
Transfer

$$2 - 4 = -2$$

$$3 - 6 = -3$$



Möglich mit Maschinellem Lernen?





—Abstraktion, Reflektion und Bewusstsein—

$$3 + 8 = 11$$

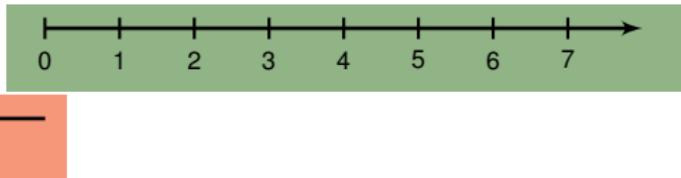
$$15 + 11 = 26$$

$$23 - 9 = 14$$

Transfer

$$2 - 4 = -2$$

$$3 - 6 = -3$$



Trotz signifikanter, aktueller Erfolge in Spezialanwendungen von KI Systemen gibt es derzeit (noch) wenig Evidenz für das Entstehen von Reflektion und Bewusstsein in Maschinen. Die Feststellung einer exponentiellen Innovation im Bereich der KI liefert keine hinreichende Begründung für das Entstehen einer Superintelligenz.

KI hat großes Potential unser Leben signifikant zu verbessern, allerdings müssen wir die Weichen richtig stellen, um Gefahren frühzeitig zu erkennen und ungewollte Entwicklungen zu vermeiden bzw. einzugrenzen.

Dynamische Bereiche (Selektion)

- ▶ Autonome Fahrzeuge
- ▶ Assistenzsysteme: Siri & Co.
- ▶ Robot Process Automation
- ▶ Internet der Dinge (IoT); 2020:
> 20 Milliarden IoT-Geräte
- ▶ Roboter: Haushalt, Industrie
- ▶ Spieleindustrie
- ▶ Bildungswesen
- ▶ Gesundheitswesen
- ▶ Umwelt, Landwirtschaft
- ▶ Administration
- ▶ Strafjustizsystem

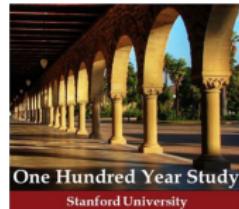
Kritische Fragen (Selektion)

- ▶ Big Data:
 - ▶ Diskriminierungspotential
 - ▶ Verfügbarkeit vs. Privatheit
- ▶ Security und Safety
- ▶ Arbeits- und Ausbildungsmarkt
- ▶ Ökonom. (Un-)Gleichgewicht
- ▶ Autonome Waffensysteme
- ▶ Kontrollier- und Steuerbarkeit der Entwicklungen
- ▶ Gesetzliche Rahmenbedingungen
- ▶ Ethik und gesellschaftl. Normen

ARTIFICIAL INTELLIGENCE AND LIFE IN 2030

ONE HUNDRED YEAR STUDY ON ARTIFICIAL INTELLIGENCE | REPORT OF THE 2015 STUDY PANEL | SEPTEMBER 2016

<https://ai100.stanford.edu/2016-report>



THE NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN

National Science and Technology Council

Networking and Information Technology
Research and Development Subcommittee

October 2016



PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE

Executive Office of the President
National Science and Technology Council
Committee on Technology

October 2016



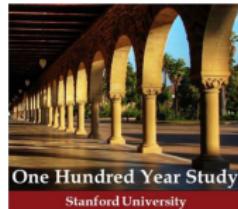
Source [Click here](#)

Source: [Click here](#)

ARTIFICIAL INTELLIGENCE AND LIFE IN 2030

ONE HUNDRED YEAR STUDY ON ARTIFICIAL INTELLIGENCE | REPORT OF THE 2015 STUDY PANEL | SEPTEMBER 2016

<https://ai100.stanford.edu/2016-report>



THE NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN

National Science and Technology Council

Networking and Information Technology
Research and Development Subcommittee

PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE

Executive Office of the President
National Science and Technology Council
Committee on Technology

October 2016

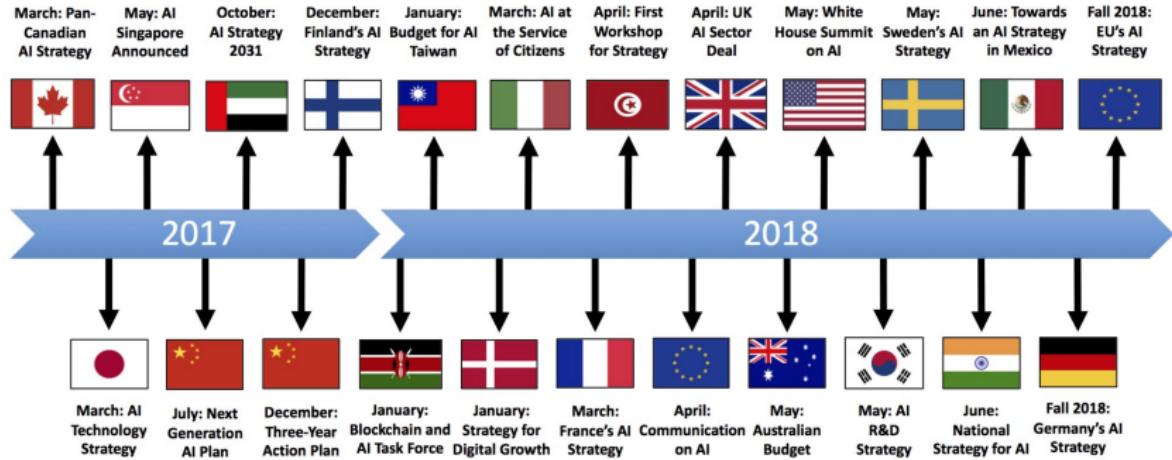


R&D Strategy	15
Strategy 1: Make Long-Term Investments in AI Research	16
Strategy 2: Develop Effective Methods for Human-AI Collaboration	22
Strategy 3: Understand and Address the Ethical, Legal, and Societal Implications of AI.....	26
Strategy 4: Ensure the Safety and Security of AI Systems.....	27
Strategy 5: Develop Shared Public Datasets and Environments for AI Training and Testing.....	30
Strategy 6: Measure and Evaluate AI Technologies through Standards and Benchmarks.....	32
Strategy 7: Better Understand the National AI R&D Workforce Needs.....	35

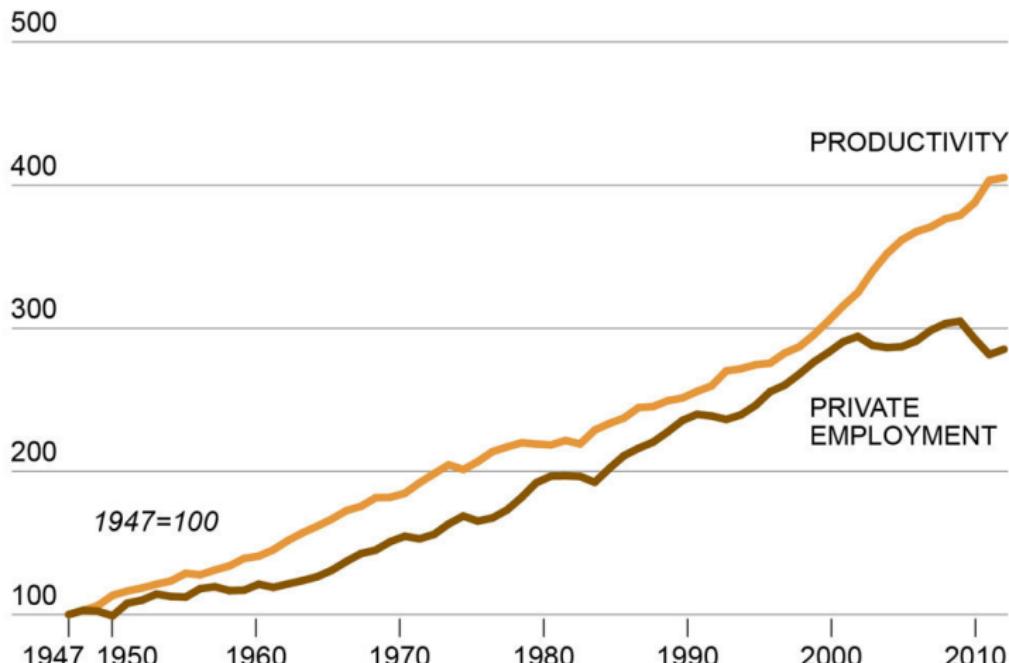
Source: [Click here](#)



Artificial Intelligence Strategies



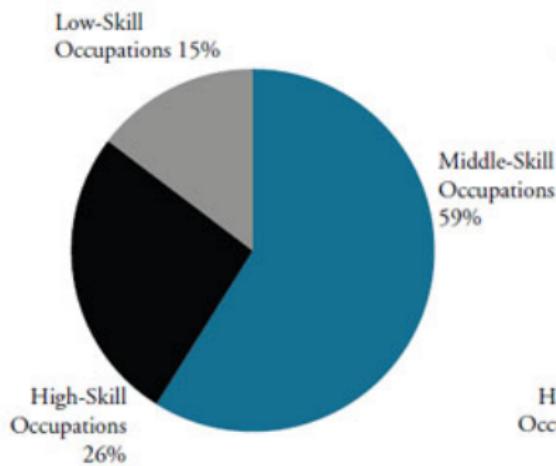
See [https://medium.com/politics-ai/
an-overview-of-national-ai-strategies-2a70ec6edfd](https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd)

Productivity and employment in the United States, 1947-2011

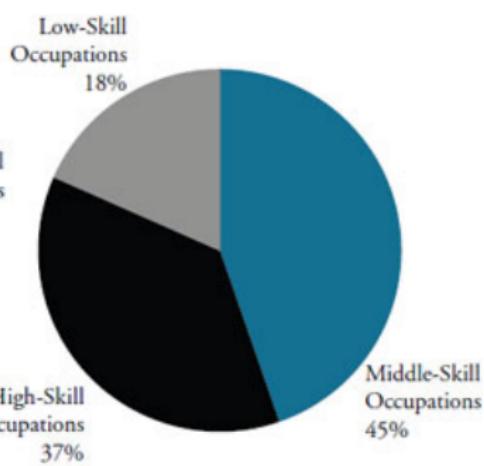
Sources: U.S. Department of Labor, Bureau of Labor Statistics

EMPLOYMENT SHARES BY SKILL LEVEL

1983



2012



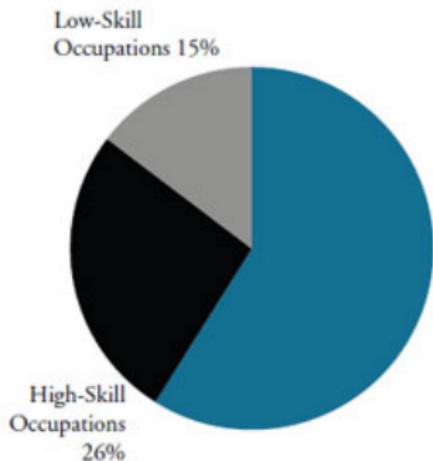
Note: Data are restricted to workers ages 16 to 64 who are not self-employed and are not employed in military or agricultural occupations.

Source: Current Population Survey and authors' calculations.

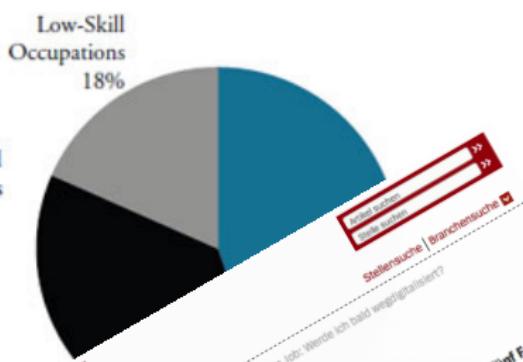
D. Tulzemen and J. Willis, The Vanishing Middle: Job Polarization and Workers Response to the Decline in Middle-Skill Jobs, Federal Reserve Bank of Kansas City in its journal Economic Review Economic Review, 2013

EMPLOYMENT SHARES BY SKILL LEVEL

1983



2012



Note: Data are restricted to workers ages 16 and older, excluding agricultural occupations.

Source: Current Population Survey and authors

D. Tulzenen and J. Willis, The Vanishing Middle: Job Po-

Jobs, Federal Reserve Bank of Kansas City in its journal E

KARRIERE SPIEGEL

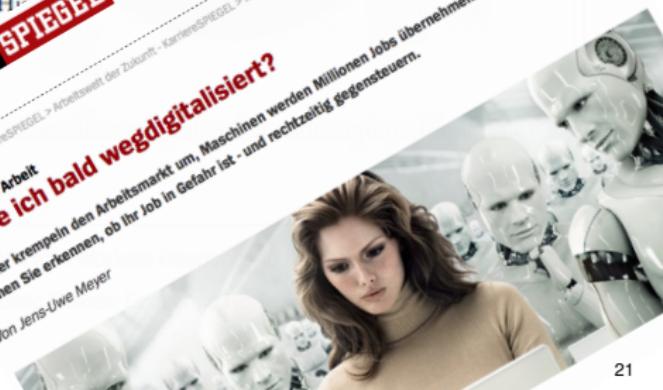
Zukunft der Arbeit

Werde ich bald wegdigitalisiert?

Roboter krempeln den Arbeitsmarkt um, Maschinen werden Millionen Jobs übernehmen. Mit fünf Fragen

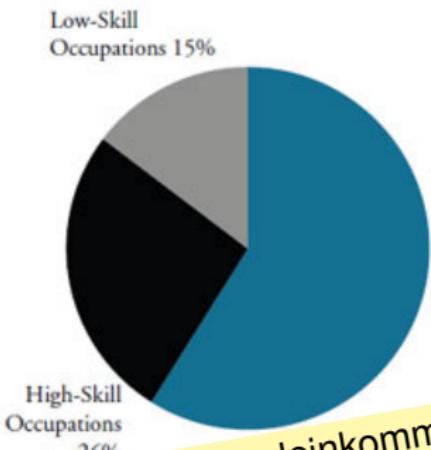
können Sie erkennen, ob Ihr Job in Gefahr ist - und rechtzeitig gegensteuern.

Von Jens-Uwe Meyer

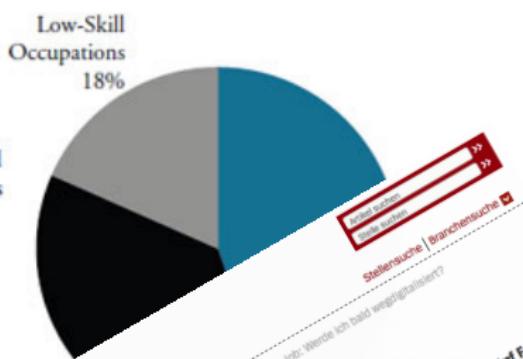


EMPLOYMENT SHARES BY SKILL LEVEL

1983



2012



Bedingungsloses Grundeinkommen & geeignete Besteuerungsmaßnahmen für KI Technologien als Antwort?

„ Tulzemen and J. Willis, The Vanishing Middle: Job Polarization and authors“

Jobs, Federal Reserve Bank of Kansas City in its journal Economic Perspectives

Zukunft der Arbeit
Werde ich bald wegdigitalisiert?

IEGEL

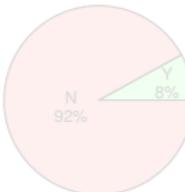
„> Arbeitswelt der Zukunft - KarrierePIEGEL > Roboter im Job: Werde ich bald wegdigitalisiert?“
Roboter krempeln den Arbeitsmarkt um, Maschinen werden Millionen Jobs übernehmen. Mit fünf Fragen können Sie erkennen, ob Ihr Job in Gefahr ist - und rechtzeitig gegensteuern.“
Von Jens-Uwe Meyer



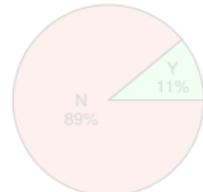
Frage an Studenten

Wird es bis 2050 eine **Superintelligenz oder Singularität** geben?

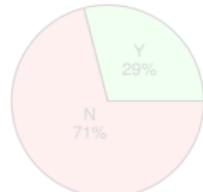
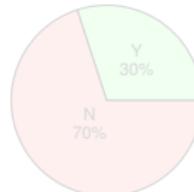
Tag 1 ($\sigma = 64$)



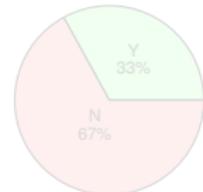
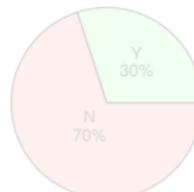
Tag 2 ($\sigma = 54$)



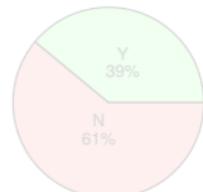
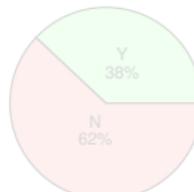
Können Maschinen **Gefühle und Emotionen** entwickeln?



Brauchen wir spezielle **Rechte für Maschinen**?



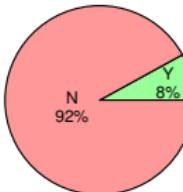
Können Maschinen **“Bewusstsein”** erlangen?



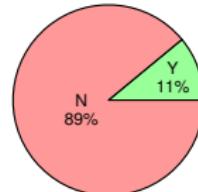
Frage an Studenten

Wird es bis 2050 eine **Superintelligenz oder Singularität** geben?

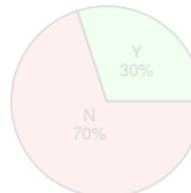
Tag 1 ($\sigma = 64$)



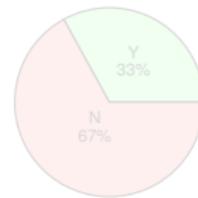
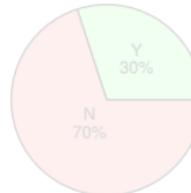
Tag 2 ($\sigma = 54$)



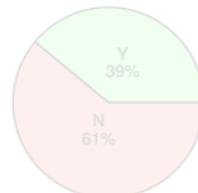
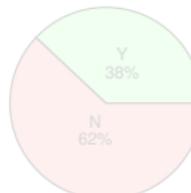
Können Maschinen **Gefühle und Emotionen** entwickeln?



Brauchen wir spezielle **Rechte für Maschinen**?



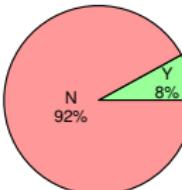
Können Maschinen **„Bewusstsein“** erlangen?



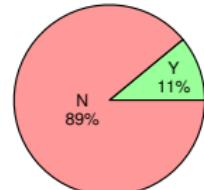
Frage an Studenten

Wird es bis 2050 eine **Superintelligenz oder Singularität** geben?

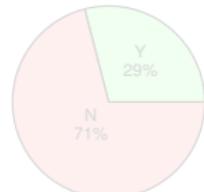
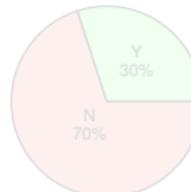
Tag 1 ($\sigma = 64$)



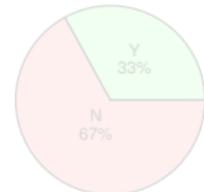
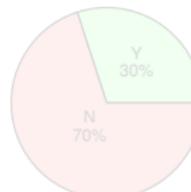
Tag 2 ($\sigma = 54$)



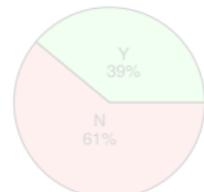
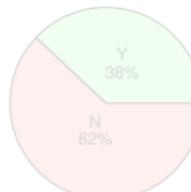
Können Maschinen **Gefühle und Emotionen** entwickeln?



Brauchen wir spezielle **Rechte für Maschinen**?



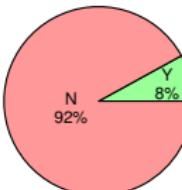
Können Maschinen **„Bewusstsein“** erlangen?



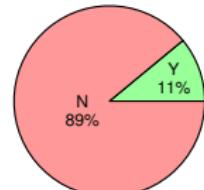
Frage an Studenten

Wird es bis 2050 eine **Superintelligenz oder Singularität** geben?

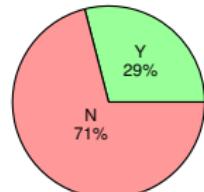
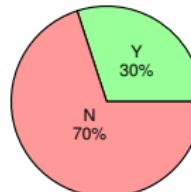
Tag 1 ($\sigma = 64$)



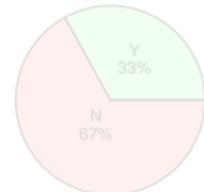
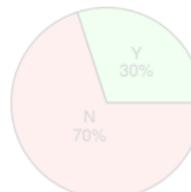
Tag 2 ($\sigma = 54$)



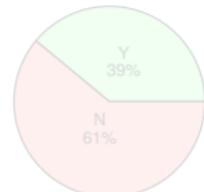
Können Maschinen **Gefühle und Emotionen** entwickeln?



Brauchen wir spezielle **Rechte für Maschinen**?



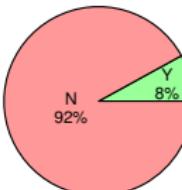
Können Maschinen **"Bewusstsein"** erlangen?



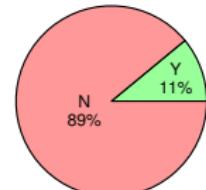
Frage an Studenten

Wird es bis 2050 eine Superintelligenz oder Singularität geben?

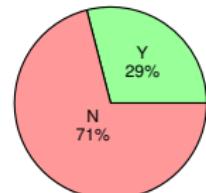
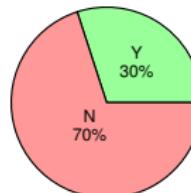
Tag 1 ($\sigma = 64$)



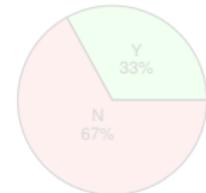
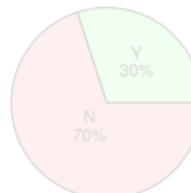
Tag 2 ($\sigma = 54$)



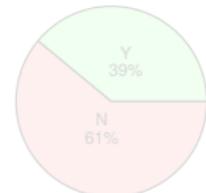
Können Maschinen Gefühle und Emotionen entwickeln?



Brauchen wir spezielle Rechte für Maschinen?



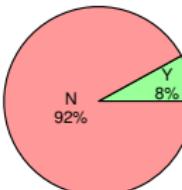
Können Maschinen "Bewusstsein" erlangen?



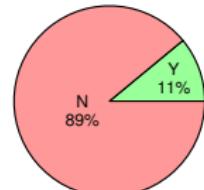
Frage an Studenten

Wird es bis 2050 eine Superintelligenz oder Singularität geben?

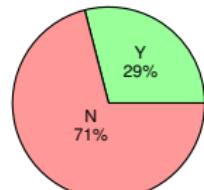
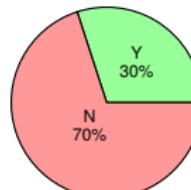
Tag 1 ($\sigma = 64$)



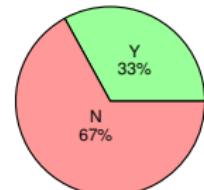
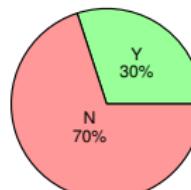
Tag 2 ($\sigma = 54$)



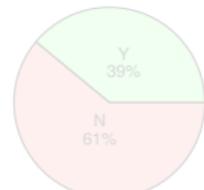
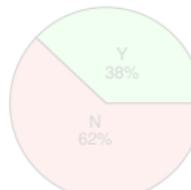
Können Maschinen Gefühle und Emotionen entwickeln?



Brauchen wir spezielle Rechte für Maschinen?



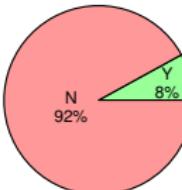
Können Maschinen "Bewusstsein" erlangen?



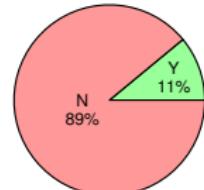
Frage an Studenten

Wird es bis 2050 eine **Superintelligenz oder Singularität** geben?

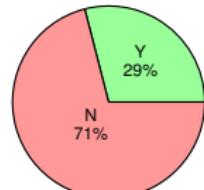
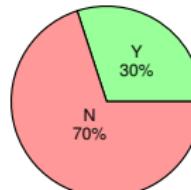
Tag 1 ($\sigma = 64$)



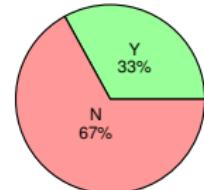
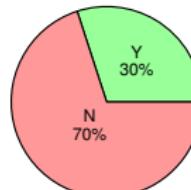
Tag 2 ($\sigma = 54$)



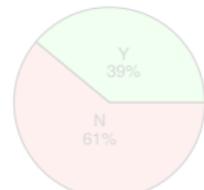
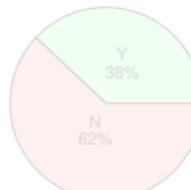
Können Maschinen **Gefühle und Emotionen** entwickeln?



Brauchen wir spezielle **Rechte für Maschinen**?



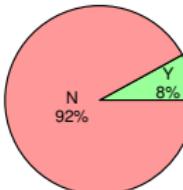
Können Maschinen "**Bewusstsein**" erlangen?



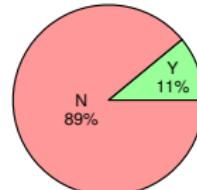
Frage an Studenten

Wird es bis 2050 eine Superintelligenz oder Singularität geben?

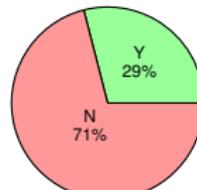
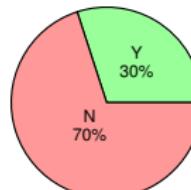
Tag 1 ($\sigma = 64$)



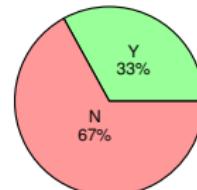
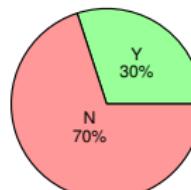
Tag 2 ($\sigma = 54$)



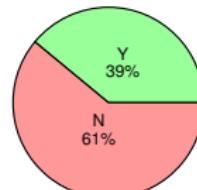
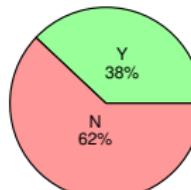
Können Maschinen Gefühle und Emotionen entwickeln?



Brauchen wir spezielle Rechte für Maschinen?



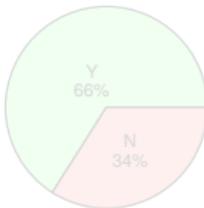
Können Maschinen "Bewusstsein" erlangen?



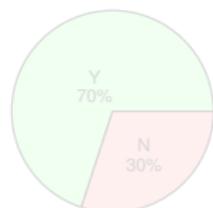
Frage an Studenten

Werden wie bis 2050 überwiegend **autonome Fahrzeuge** haben?

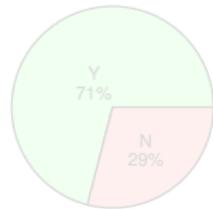
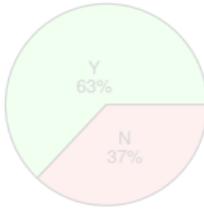
Tag 1 ($\sigma = 64$)



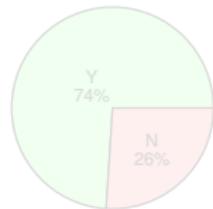
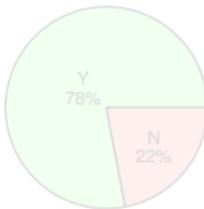
Tag 2 ($\sigma = 54$)



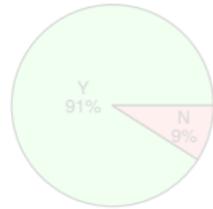
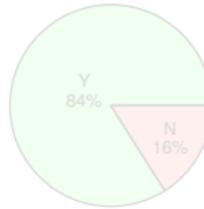
Sollten wir KI Technlogien in bestimmten Bereichen **bannen**?



Können Maschinen "**Intelligenz**" erlangen?



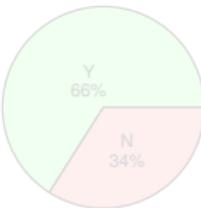
Brauchen wir spezielle **Gesetze und ethische Regeln** für Maschinen?



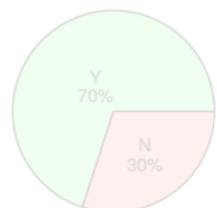
Frage an Studenten

Werden wie bis 2050 überwiegend **autonome Fahrzeuge** haben?

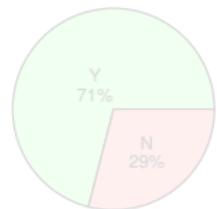
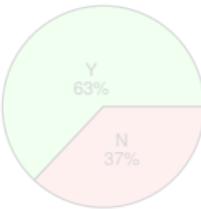
Tag 1 ($\sigma = 64$)



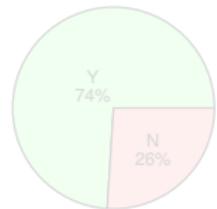
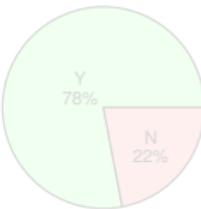
Tag 2 ($\sigma = 54$)



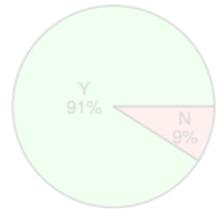
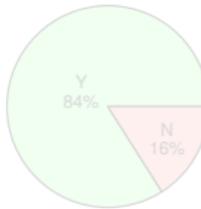
Sollten wir KI Technologien in bestimmten Bereichen **bannen**?



Können Maschinen "**Intelligenz**" erlangen?



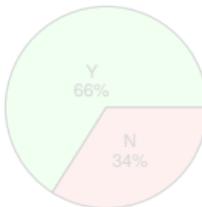
Brauchen wir spezielle **Gesetze und ethische Regeln** für Maschinen?



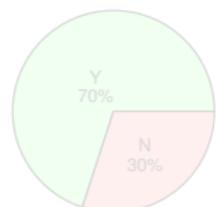
Frage an Studenten

Werden wie bis 2050 überwiegend **autonome Fahrzeuge** haben?

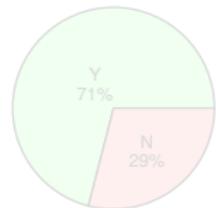
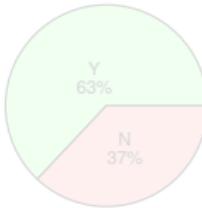
Tag 1 ($\sigma = 64$)



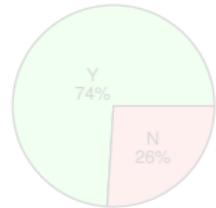
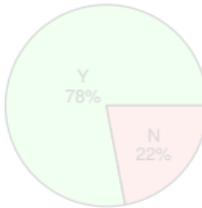
Tag 2 ($\sigma = 54$)



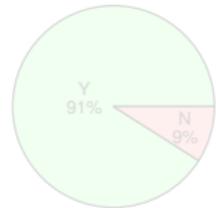
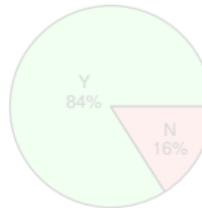
Sollten wir KI Technologien in bestimmten Bereichen **bannen**?



Können Maschinen "**Intelligenz**" erlangen?



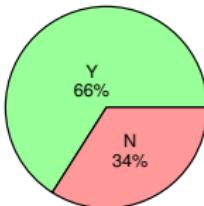
Brauchen wir spezielle **Gesetze und ethische Regeln** für Maschinen?



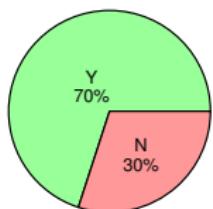
Frage an Studenten

Werden wie bis 2050 überwiegend **autonome Fahrzeuge** haben?

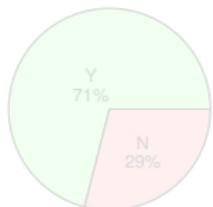
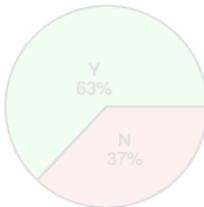
Tag 1 ($\sigma = 64$)



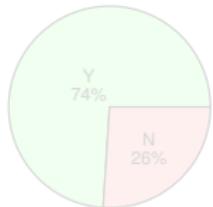
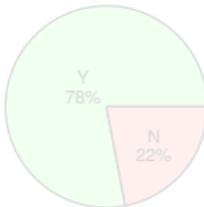
Tag 2 ($\sigma = 54$)



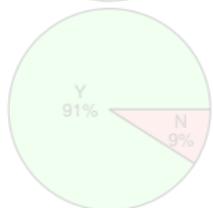
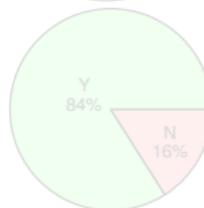
Sollten wir KI Technologien in bestimmten Bereichen **bannen**?



Können Maschinen "**Intelligenz**" erlangen?



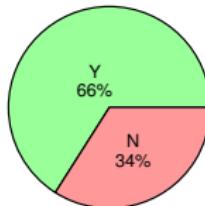
Brauchen wir spezielle **Gesetze und ethische Regeln** für Maschinen?



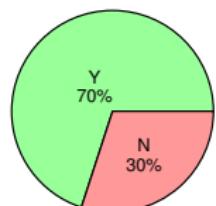
Frage an Studenten

Werden wie bis 2050 überwiegend **autonome Fahrzeuge** haben?

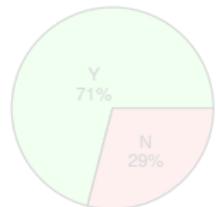
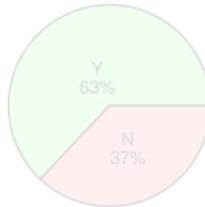
Tag 1 ($\sigma = 64$)



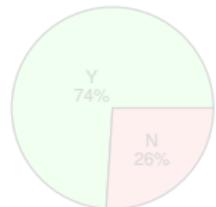
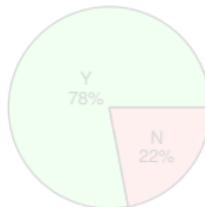
Tag 2 ($\sigma = 54$)



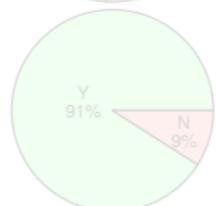
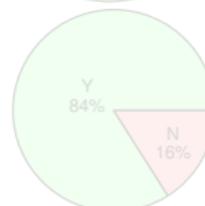
Sollten wir KI Technologien in bestimmten Bereichen **bannen**?



Können Maschinen "**Intelligenz**" erlangen?



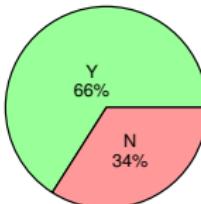
Brauchen wir spezielle **Gesetze und ethische Regeln** für Maschinen?



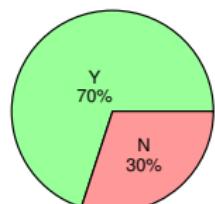
Frage an Studenten

Werden wie bis 2050 überwiegend **autonome Fahrzeuge** haben?

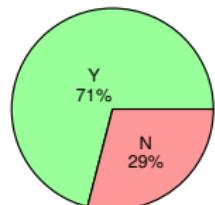
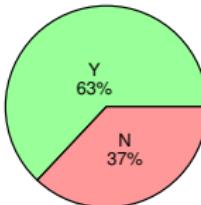
Tag 1 ($\sigma = 64$)



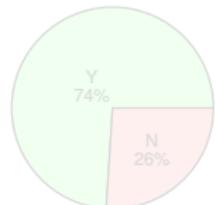
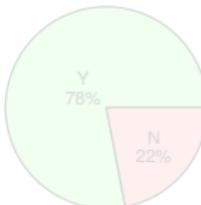
Tag 2 ($\sigma = 54$)



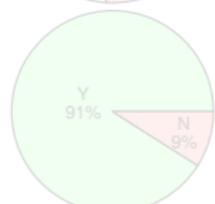
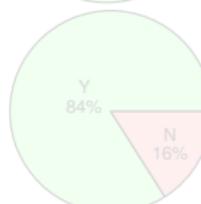
Sollten wir KI Technologien in bestimmten Bereichen **bannen**?



Können Maschinen "**Intelligenz**" erlangen?



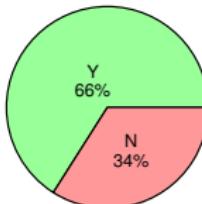
Brauchen wir spezielle **Gesetze und ethische Regeln** für Maschinen?



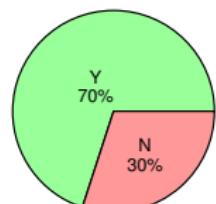
Frage an Studenten

Werden wie bis 2050 überwiegend **autonome Fahrzeuge** haben?

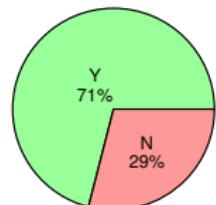
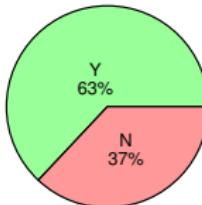
Tag 1 ($\sigma = 64$)



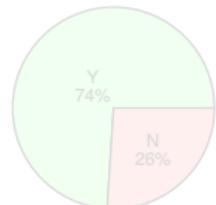
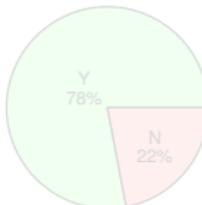
Tag 2 ($\sigma = 54$)



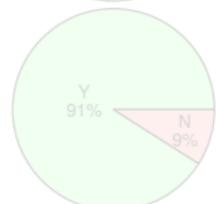
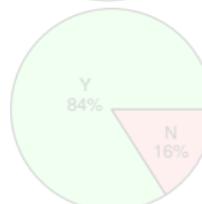
Sollten wir KI Technologien in bestimmten Bereichen **bannen**?



Können Maschinen "**Intelligenz**" erlangen?



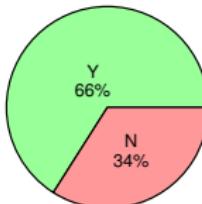
Brauchen wir spezielle **Gesetze und ethische Regeln** für Maschinen?



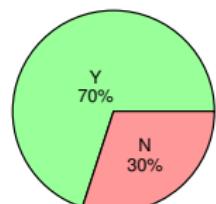
Frage an Studenten

Werden wie bis 2050 überwiegend **autonome Fahrzeuge** haben?

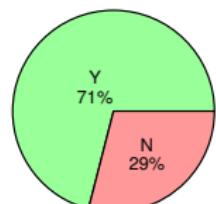
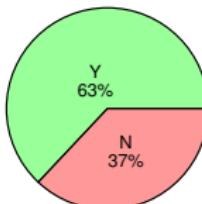
Tag 1 ($\sigma = 64$)



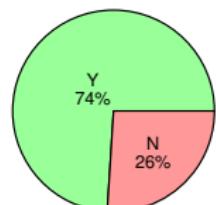
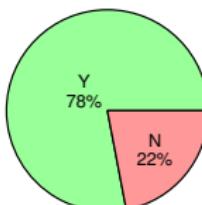
Tag 2 ($\sigma = 54$)



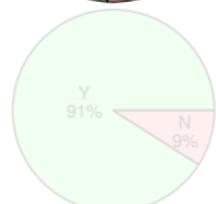
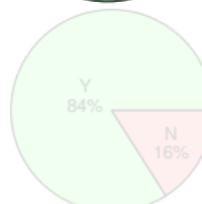
Sollten wir KI Technologien in bestimmten Bereichen **bannen**?



Können Maschinen "**Intelligenz**" erlangen?



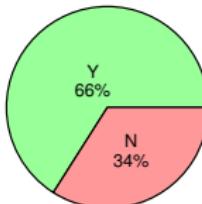
Brauchen wir spezielle **Gesetze und ethische Regeln** für Maschinen?



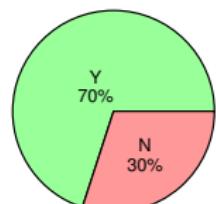
Frage an Studenten

Werden wie bis 2050 überwiegend **autonome Fahrzeuge** haben?

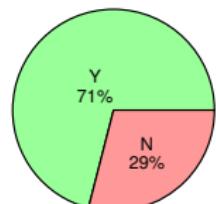
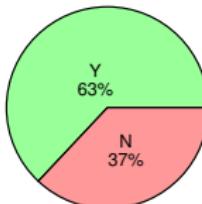
Tag 1 ($\sigma = 64$)



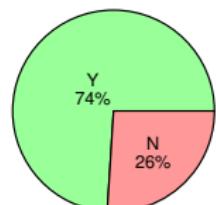
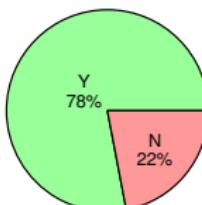
Tag 2 ($\sigma = 54$)



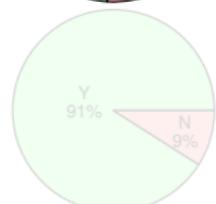
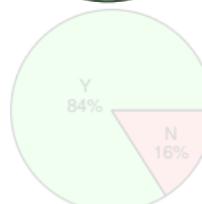
Sollten wir KI Technologien in bestimmten Bereichen **bannen**?



Können Maschinen "**Intelligenz**" erlangen?



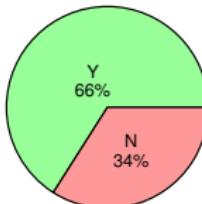
Brauchen wir spezielle **Gesetze und ethische Regeln** für Maschinen?



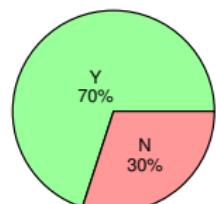
Frage an Studenten

Werden wie bis 2050 überwiegend **autonome Fahrzeuge** haben?

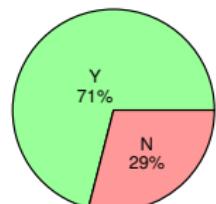
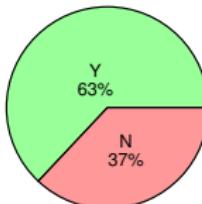
Tag 1 ($\sigma = 64$)



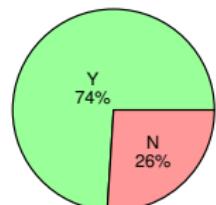
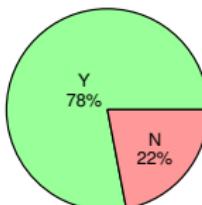
Tag 2 ($\sigma = 54$)



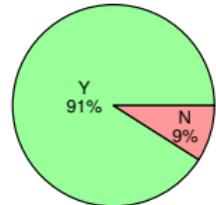
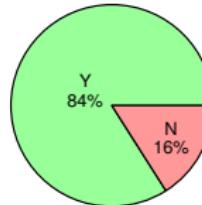
Sollten wir KI Technologien in bestimmten Bereichen **bannen**?



Können Maschinen "**Intelligenz**" erlangen?



Brauchen wir spezielle **Gesetze und ethische Regeln** für Maschinen?



Konklusion



Quelle: *Ex Machina*, R/D: Alex Garland, 2015

Beschleunigte Entwicklung von KI-Technologien

- ▶ Exponent. Innovation: Großer Wandel innerhalb einer Generation
- ▶ Aber keine Evidenz für Superintelligenz
- ▶ Menschliche Fähigkeiten von Maschinen in vielen Einzeldomänen
- ▶ Große Chancen
- ▶ Große Herausforderungen



Quelle: Ex Machina, R/D: Alex Garland, 2015

Beschleunigte Entwicklung von KI-Technologien

- ▶ Exponent. Innovation: Großer Wandel innerhalb einer Generation
- ▶ Aber keine Evidenz für Superintelligenz
- ▶ Menschliche Fähigkeiten von Maschinen in vielen Einzeldomänen
- ▶ Große Chancen
- ▶ Große Herausforderungen

Mögliche Maßnahmen

- ▶ Bildung und Ausbildung
- ▶ Gut informierte Entscheidungsträger
- ▶ Investition in KI-Forschung & KI-Entwicklung
- ▶ Investition in flankierende Maßnahmen
- ▶ Entwicklung einer KI-Ethik
- ▶ Sachliche Aufklärung in den Medien

Autonome Maschinen in der realen Welt

- ▶ Gesetzliche Rahmenbedingungen
- ▶ Verantwortung und Haftbarkeit bei Fehlverhalten und Schäden
- ▶ Trennung von Entscheidungsfindung und Entscheidungsbewertung?
- ▶ Integration in bestehende Normen-/Wertesysteme
- ▶ Modellierung einer ethisch-moralischen Wertesysteme für Maschinen?
- ▶ Morale Dilemma, konfliktäre Wertesysteme
- ▶ Gesetze zum Schutz von Maschinen?

Verwendung von "Big Data" in KI-Systemen

- ▶ Diskrimierung
- ▶ Privatheit von Daten
- ▶ Zugang
- ▶ Fairness

Notwendig: Stärkung der flankierenden Forschung zu solchen Themen