

Künstliche Intelligenz — Wohin geht die Reise?

Christoph Benzmüller

Dahlem Center for Machine Learning and Robotics, FU Berlin

“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?

Christoph Benzmüller

Dahlem Center for Machine Learning and Robotics, FU Berlin

“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)

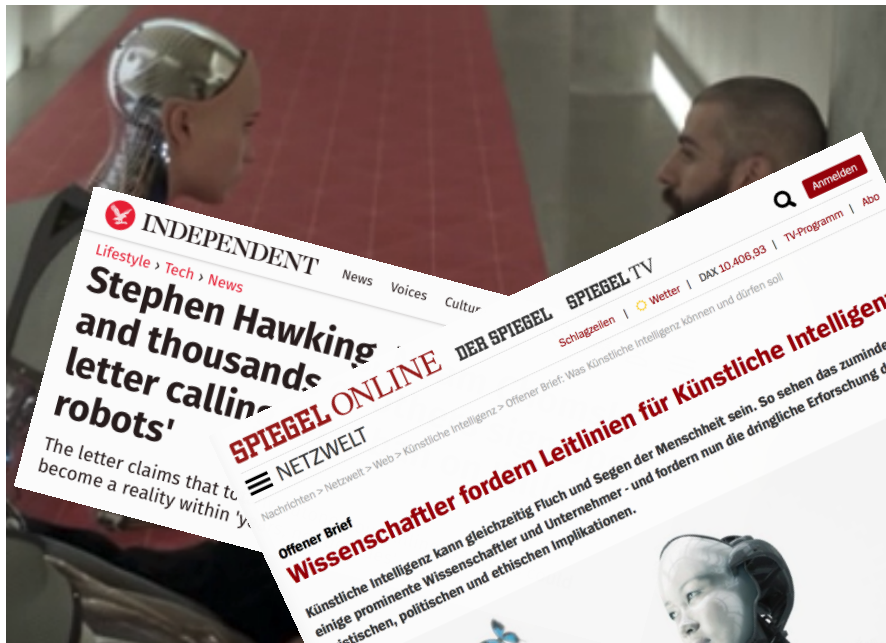




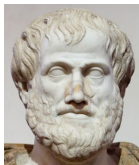
Quelle: *Ex Machina*, R/D: Alex Garland, 2015



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



KI-Visionäre:



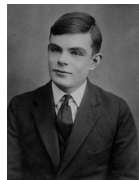
Aristoteles



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

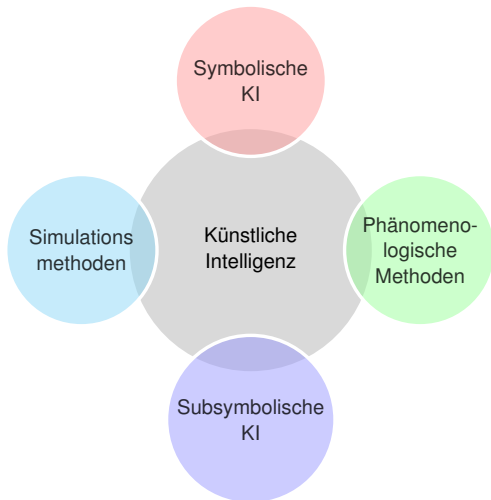
vs.

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

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:

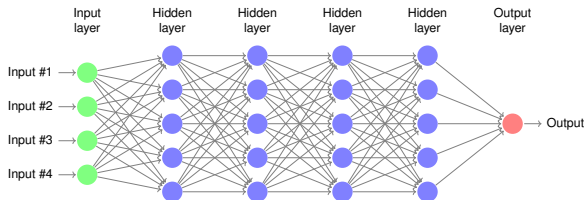
- ▶ Intelligente Software Assistenten (Siri & Co)
- ▶ 2011: Watson, Jeopardy-Quiz
- ▶ 2016: AlphaGo, schlägt Go-Champion L. Sedol



Neue Erfolge: Subsymbolische KI und Kombination von Techniken

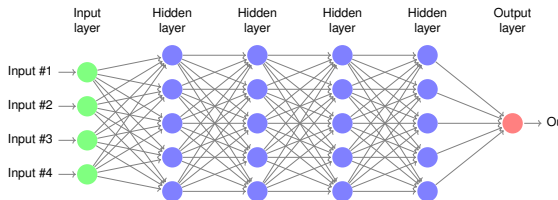
Beispiele:

- ▶ Intelligente Software Assistenten (Siri & Co)
- ▶ 2011: Watson, Jeopardy-Quiz
- ▶ 2016: AlphaGo, schlägt Go-Champion L. Sedol

**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

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

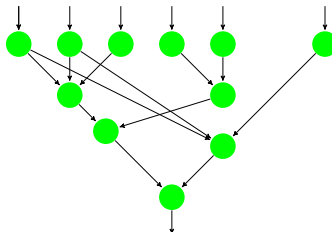
- ▶ Computerbeweis von Kepler's Vermutung (interaktiv)
- ▶ SAT Solver lösen offene mathematische Probleme (automatisch)
- ▶ Eigene Arbeit zum Ontologischen Gottesbeweis auf dem Computer (automatisch)

Aussage (und Annahmen)

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

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

Axiome & Regeln der Logik



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



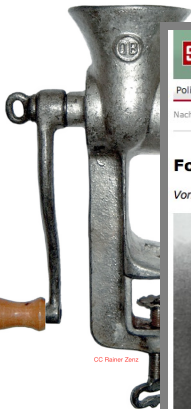
Neue Erfolge in der Mechanisierung/Automatisierung logischen Schließens:

- ▶ 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)

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

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



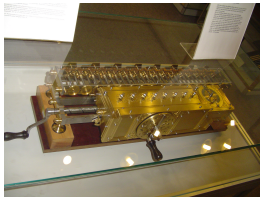
CC Rainer Zenz



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 ($\varphi \in P$)

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

[1] $G(x) \equiv (\varphi) [P(\varphi) \supset \varphi(x)]$ (God)

[2] $\varphi \text{ Ess. } x \equiv (\psi) [\psi(x) \supset N(y) [\varphi(y) \supset \psi(y)]]$ (Essence of x)

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

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

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

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

Ax. 3 $P(E)$

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

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

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

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

M = possibility

any two sentences of x are nec. equivalent

exclusive or * and for any number of humanoids

$M(\exists x) G(x)$ means ^{the system of} all pos. props. is com-possible
 This is true because of:

At. 4: $P(\varphi) \cdot \varphi \supset N \psi \supset P(\psi)$ which impl.

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

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

Positive means positive in the moral aesthetic sense (independently of the accidental structure of the world). Only then the act. time. It must also mean "attribution" as opposed to "privation" (or containing privation). This is the positive principle.

If φ is positive: $(x) N \neg \varphi(x)$ - otherwise $\varphi(x) \supset N x \neq x$

hence $x \neq x$ positive pos. $x=x$ neg. contradictory At.

or the existence of pos. prop. At. 4

x 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

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

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

[1] $G(x) \equiv (\varphi) [P(\varphi) \supset \varphi(x)]$

[2] $\varphi \text{ En } x \equiv (\psi) [\psi(x) \supset N(y) [\varphi(y) \supset \psi(y)]]$

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

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

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

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

Ax 3 $P(E)$

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

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

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

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

M = possibility

any two instances of x are nec. equivalent

exclusive or * and for any number of humanoids

Detaillierte Analyse Rationaler Argumente auf dem Computer

$M(\exists x) G(x)$ m. (the mode)

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

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

Positive means positive in the moral aesthetic sense (independently of the accidental structure of the world). Only then the at. time. It must also mean "attribution" as opposed to "privation" (or containing privation). This is the positive proof.

If $\neg \varphi$ is not: $(x) N \neg \varphi(x)$ - otherwise $\varphi(x) \supset x \neq x$

hence $x \neq x$ positive not $x=x$ neg. Contrary At.

on the basis of pos. prop. At.

x 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

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

At. 1 $P(\varphi) \cdot P(\varphi)$

[1] $G(x) \equiv$

[2] $\varphi \text{ exists}$

$P \supset Nq =$

At 2 $P(\varphi) \supset N \neg P(\varphi)$

$\sim P(\varphi) \supset N \sim P(\varphi)$

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

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

Ax 3 $P(E)$

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

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

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

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

any two instances of x are nec. equivalent

exclusive or * and for any number of humanoids

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>

Detaillierte Analyse Rationaler Argumente auf dem Computer

$M(\exists x) G(x)$ *the model*

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

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

positive in the moral action

accidental structure of

It is

"privation

idea proof

(a) $\neg \varphi$ pos.

hence $x \neq x$ pos. *or the existence of prop. At*

x i.e. the normal form in terms of elem. prop. contains a member without negation.

(With permission from the Kurt Gödel Papers, Shelby White and Leon Levy Archives Center, Princeton, NJ, USA, on deposit at Princeton University)

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

Ontologischer Beweis

Feb. 10, 1970

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

At. 1 $P(\varphi) \cdot P(\varphi)$

[1] $G(x) \equiv$

[2] $\varphi \text{ exists}$

$P \supset N \varphi =$

At 2 $P(\varphi) \supset N P(\varphi)$

$\sim P(\varphi) \supset N \sim P(\varphi)$

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

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

Ax 3 $P(E)$

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

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

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

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

any two instances of x are, nec.

exclusive or

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>

$M = \text{possibility}$

Detaillierte Analyse Rationaler Argumente auf dem Computer

$M(\exists x) G(x)$ m $\leftarrow$ the mode

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

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

positive in the moral action

accidental structure of

It is

privation

idea proof

Weltweit erste Vorlesung Computationale Metaphysik an der FU Berlin (ausgezeichnet mit zentralem Lehrpreis der FU)
<http://www.inf.fu-berlin.de/users/lex/lehre/compmeta/>

(With permission of Kurt Gödel and Leon Levy Archives Center, Princeton, NJ, USA, on deposit at Princeton University)

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 "The Matrix".
password: $2 + 2 = 4.0!$ Now divide by zero and die!!!!

Question: If I have two marbles...

Eugene: Not too many.
you live...

Turing Test

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



(Uneingeschränkter) Turing Test ist nicht gelöst!

Künstliche Intelligenz

Computerprogramm gaukelt erfolgreich Menschsein vor

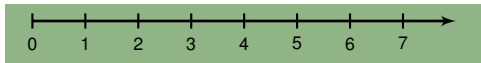
Turing Test bedingt geeignet für: Bewusstsein, Emotionen, Kreativität



$$3 + 8 = 11$$

$$15 + 11 = 26$$

$$23 - 9 = 14$$





—Abstraktion, Reflektion und Bewusstsein—

$$3 + 8 = 11$$

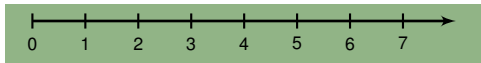
$$15 + 11 = 26$$

$$23 - 9 = 14$$

Transfer

$$2 - 4 = -2$$

$$3 - 6 = -3$$





$$3 + 8 = 11$$

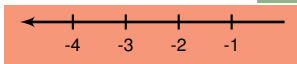
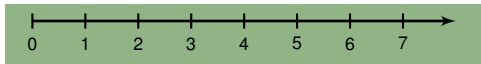
$$15 + 11 = 26$$

$$23 - 9 = 14$$

Transfer

$$2 - 4 = -2$$

$$3 - 6 = -3$$





$$3 + 8 = 11$$

$$15 + 11 = 26$$

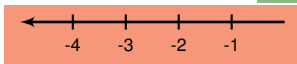
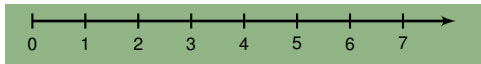
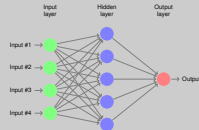
$$23 - 9 = 14$$

Transfer

$$2 - 4 = -2$$

$$3 - 6 = -3$$

Möglich mit Maschinellem Lernen?





$$3 + 8 = 11$$

$$15 + 11 = 26$$

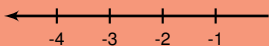
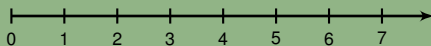
$$23 - 9 = 14$$

Transfer

$$2 - 4 = -2$$

$$3 - 6 = -3$$

Möglich mit Maschinellem Lernen?



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.
- ▶ **RPA**
- ▶ Internet der Dinge (IoT); 2020:
> 20 Milliarden IoT-Geräte
- ▶ Roboter: Haushalt, Industrie
- ▶ Spieleindustrie
- ▶ Bildungswesen
- ▶ Gesundheitswesen
- ▶ Umwelt, Landwirtschaft
- ▶ Administration
- ▶ Strafjustizsystem

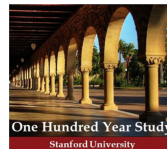
Kritische Fragen (Selektion)

- ▶ 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



Source: Click here

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

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

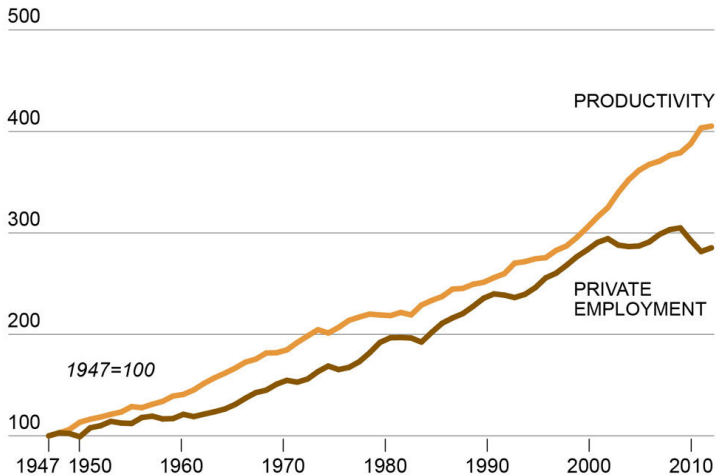
October 2016



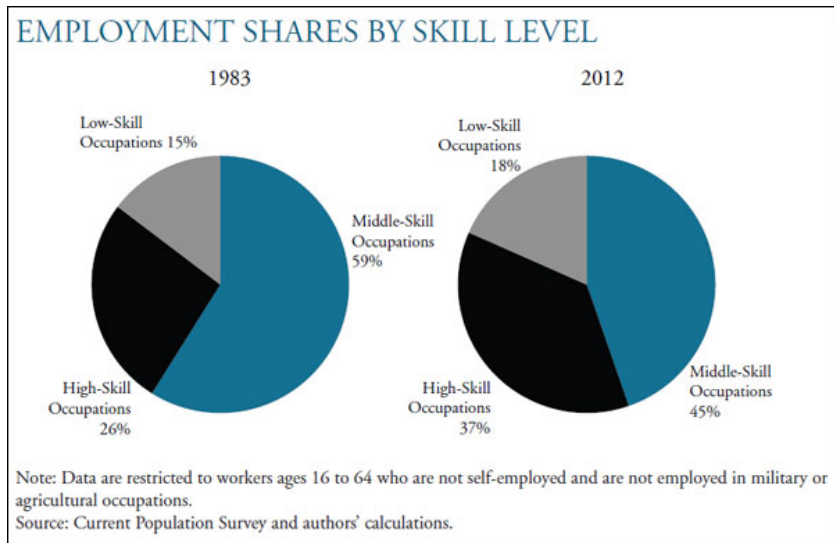
R&D Strategy	15
Strategy 1: Make Long-Term Investments in AI Research	16
Strategy 2: Develop Effective Methods for Human-AI Collaboration	22
Strategy 3: Understand and Address the Ethical, Legal, and Societal Implications of AI	26
Strategy 4: Ensure the Safety and Security of AI Systems	27
Strategy 5: 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](#)

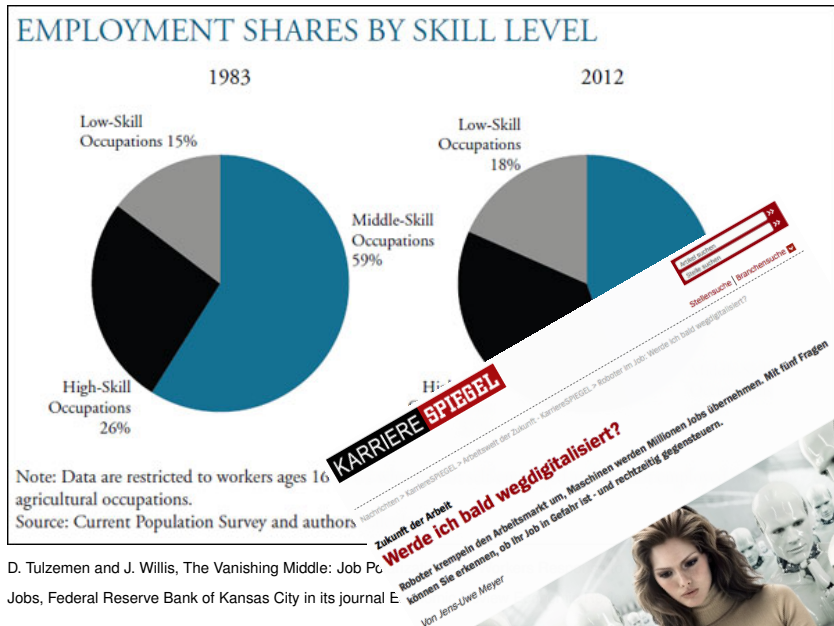
Productivity and employment in the United States, 1947-2011



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

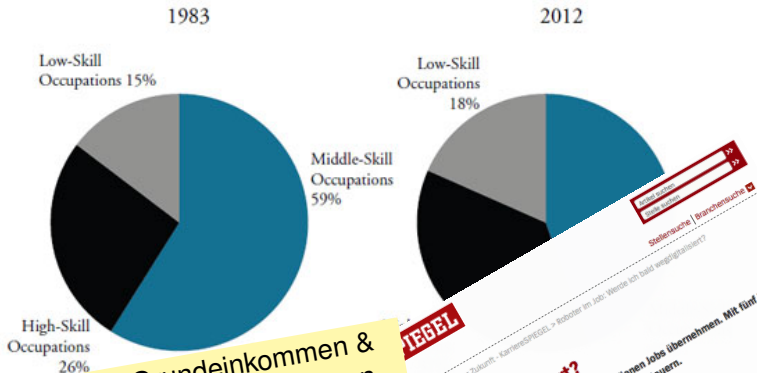


D. Tulzemen and J. Willis, The Vanishing Middle: Job Polarization and Workers Response to the Decline in Middle-Skill Jobs, Federal Reserve Bank of Kansas City in its journal Economic Review Economic Review, 2013



D. Tulzemen and J. Willis, The Vanishing Middle: Job Pc
Jobs, Federal Reserve Bank of Kansas City in its journal E

EMPLOYMENT SHARES BY SKILL LEVEL



Bedingungsloses Grundeinkommen & geeignete Besteuerungsmaßnahmen für KI Technologien als Antwort?

D. Tulzemen and J. Willis, The Vanishing Middle: Job Pc

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

Zukunft der Arbeit

Werde ich bald wegdigitalisiert?

Roboter krepeln den Arbeitsmarkt um, Maschinen werden Millionen Jobs übernehmen. Mit fünf Fragen können Sie erkennen, ob Ihr Job in Gefahr ist - und rechtzeitig gegensteuern.

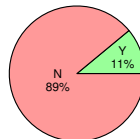
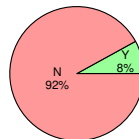
Von Jens-Uwe Meyer

Frage an Studenten

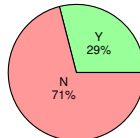
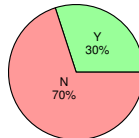
Tag 1 ($\sigma = 64$)

Tag 2 ($\sigma = 54$)

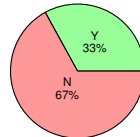
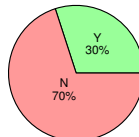
Wird es bis 2050 eine **Superintelligenz oder Singularität** geben?



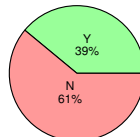
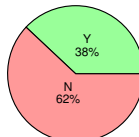
Können Maschinen **Gefühle und Emotionen** entwickeln?



Brauchen wir spezielle **Rechte für Maschinen**?



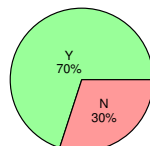
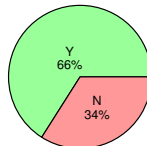
Können Maschinen **“Bewusstsein”** erlangen?



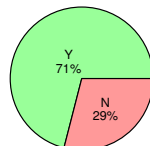
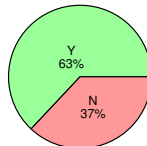
Frage an Studenten

Tag 1 ($\sigma = 64$)Tag 2 ($\sigma = 54$)

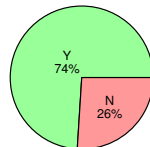
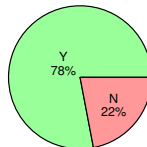
Werden wie bis 2050 überwiegend **autonome Fahrzeuge** haben?



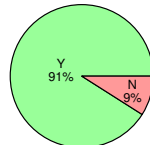
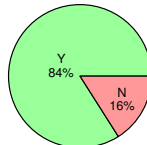
Sollten wir KI Technologien in bestimmten Bereichen **bannen**?



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 eine ethisch-moralischer Wertesysteme für Maschinen?
- ▶ Moralische Dilemma, konfliktäre Wertesysteme
- ▶ Gesetze zum Schutz von Maschinen?

Verwendung von "Big Data" in KI-Systemen

- ▶ Diskriminierung
- ▶ Privatheit von Daten
- ▶ Zugang
- ▶ Fairness

Notwendig: Stärkung der flankierenden Forschung zu solchen Themen