

## Der Turing Test und andere Testverfahren für künstliche Intelligenzen

### Einleitung

Aufgrund seiner Allgemeingültigkeit wird davon ausgegangen, dass jeder Mensch in seinem Leben bereits einen Turing Test absolviert hat.

Zwar entsprach dieser sicherlich nicht den genauen Vorgaben von Alan Mathison Turing, jedoch lässt sich hieran erkennen, dass er mit der Erfindung des Turing Tests einen bis heute wesentlichen Grundstein gelegt hat in der Welt der künstlichen Intelligenz.

Die Erfindung des Turing Tests geht auf die Mitte des 20. Jahrhunderts zurück. Damals definierte Alan Turing, „ein Pionier auf dem Gebiet der künstlichen Intelligenz“ (Redaktion ComputerWeekly.de, 2021), den Turing Test als „eine Möglichkeit, die Intelligenz einer Maschine zu testen.“ (Universität Oldenburg, 2008/2009). Ein bekanntes Beispiel für einen Turing Test sind die oftmals in Webseiten eingebundenen CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart). Hierbei wird der Nutzer bei dem Login eines Onlinekontos aufgefordert, Muster wie beispielsweise Transportbusse oder Fußgängerüberwege in einem Bild zu erkennen. Das Prinzip des CAPTCHA dient der Ermittlung, ob es sich bei den Nutzer:innen um einen Menschen oder eine Maschine handelt.<sup>1</sup>

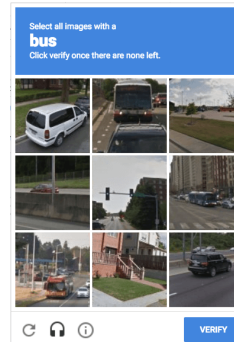


Abbildung 1 Bildschirmaufnahme CAPTCHA

Der Zweck besteht darin

DDOS Angriffe (Distributed-Denial-of-Service) abzublocken, die ein System außer Betrieb setzen können. Dabei werden mittels Software so viele Anfragen an einen Server geschickt, bis er überlastet wird und diese nicht mehr verarbeiten kann. Durch Einsatz des CAPTCHAS kann der Login jedoch nicht durch eine Software, sondern ausschließlich durch Menschen ausgeführt werden, wodurch der DDOS Angriff verhindert wird.<sup>2</sup>

### Alan Turing

Geboren am 23. Juni 1912 in London, zeigt sich bereits früh Alan Mathison Turings Begabung. So lernt er innerhalb weniger Wochen selbständig Lesen und wird bereits mit 14 Jahren durch Empfehlung seiner damaligen Rektorin auf die Sherborn School in Dorset, England, geschickt. Nach

<sup>1</sup> Luis von Ahn, M. B. (kein Datum). *International Association for Cryptologic Research*. Von <https://iacr.org/archive/eurocrypt2003/26560294/26560294.pdf> abgerufen

<sup>2</sup> Scherschel, F. A. (02 2017). *heise.de*. Von heise.de: <https://www.heise.de/security/meldung/DDoS-Untersuchung-Angriffe-werden-zum-Problem-fuer-die-Allgemeinheit-3631903.html> abgerufen

seinem Abschluss am King's College in Cambridge mit zwischenzeitlichem Gastsemester in Princeton, beruft ihn die damalige Regierung unter Premierminister Neville Chamberlain an die zentrale Militärdienststelle in Bletchley Park. Gemeinsam mit anderen Wissenschaftler:innen soll er die deutsche Verschlüsselungsmaschine Enigma dechiffrieren. Durch die bis heute bekannte Turing Maschine gelingt es Alan Turing 1943 den Funkverkehr der Nationalsozialisten in Klartext zu übersetzen. Infolgedessen sind die Alliierten in der Lage diversen Angriffsschlägen der Nationalsozialisten vorzubeugen.<sup>3 4 5</sup>

Nach dem Ende des zweiten Weltkrieges konzentriert sich Alan Turing auf die Forschung im Bereich der künstlichen Intelligenz. Von 1945 bis 1948 ist er am National Physical Laboratory in Teddington mit der Entwicklung des ACE Computer, einer der ersten weltweit, beschäftigt. Nach dem Wechsel an die Universität in Manchester wird Alan Turing 1949 stellvertretender Direktor der dortigen Fachabteilung für Computer. 1950 veröffentlicht Turing in der britischen Zeitung „Mind“ den Artikel „Computing Machinery and Intelligence“ über den Turing Test. Darin heißt es: „Thinking is a function of man's immortal soul. God has given an immortal soul to every man and woman, but not to any other animal or machines. Hence no animal or machine can think. I am unable to accept any part of this, but will attempt to reply in theological terms.“ (Turing, 1950)

In dieser Publikation beschäftigt sich Turing mit der Beantwortung des Zitats und

schlussfolgert daraus den heute weltweit bekannten Turing Test.

In den letzten Jahren seines Lebens kämpft Alan Turing mit Depressionen als Nebenwirkungen einer Hormontherapie, zu der er aufgrund seiner Homosexualität verurteilt wurde. Am 07. Juni 1954 vergiftet er sich mit Blausäure selbst. Im Dezember 2013 wird er nachträglich von Queen Elisabeth II. begnadigt.<sup>6</sup>

## Der Turing Test

Bei einem Turing Test führt eine Person C zwei verschiedene Unterhaltungen. Eine davon ist mit einer künstlichen Intelligenz A, die andere mit einem Menschen B. Der Austausch läuft dabei ohne audiovisuellen Kontakt über einen Monitor als Ausgabe- und eine Tastatur als Eingabegerät. Kann die Person C nach einer ausführlichen Befragung der beiden Gesprächspartner nicht eindeutig feststellen, bei welchem es sich um eine Maschine handelt, gilt der Turing Test als bestanden.<sup>7</sup>

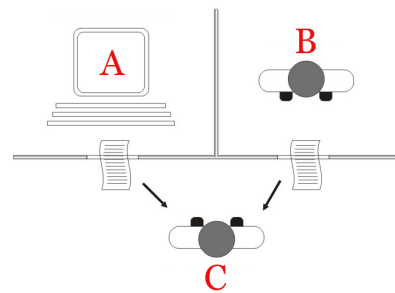


Abbildung 2 Grafik: Anordnung Turing Test

3 Cassier, P. (2021). *Welt.de*. Abgerufen am 28. 02 2022 von <https://www.welt.de/geschichte/kopf-des-tages/article235850396/Alan-Turing-Die-absurde-Begnadigung-eines-Genies.html>

4 EU. (2018). *European Mathematical Information Service*. Abgerufen am 28. 02 2022 von <https://www.emis.de/community/history-of-mathematics/alan-turing/index.html>

5 Kettern, A. (2014). *Lebendiges Museum Online*. Abgerufen am 28. 02 2022 von <https://www.dhm.de/lemo/biografie/arthur-chamberlain>

6 Copeland, B. J. (kein Datum). *Britannica.com*. Abgerufen am 28. 02 2022 von <https://www.britannica.com/biography/Alan-Turing#ref704283>

7 Redaktion ComputerWeekly.de. (2021). *WhatIs.de*. Abgerufen am 17. 02 2022 von <https://whatIs.techtarget.com/de/definition/Turing-Test>

In Alan Turings Veröffentlichung zum Turing Test sind die Rahmenbedingungen zur Durchführung des Tests nicht präzise formuliert. So werden weder Eigenschaften oder Qualifikationen beider Gesprächspartner noch die Dauer des Tests genau definiert.

Weitere Kritik wird an der Funktionalität geübt. So bemängeln Fachleute, dass es sich um eine reine Prüfung der Funktionalität der künstlichen Intelligenz handelt und nicht um die Feststellung zur Existenz von menschlichem Denkvermögen. Zur Veranschaulichung wird oft auf den Vergleich des chinesischen Zimmers zurückgegriffen. Befragt man eine nicht chinesisch sprechende Person auf Chinesisch, so ist diese nicht in der Lage zu antworten. Stellt man ihr jedoch als Hilfestellung ein Wörterbuch zur Verfügung, so könnte, wenn man nicht weiß, dass sie ein Wörterbuch benutzt, der Eindruck vermittelt werden, sie beherrsche die chinesische Sprache.<sup>8</sup>

Des Weiteren wird in Frage gestellt, ob bei dem Test nicht nur die Leichtgläubigkeit der Versuchsteilnehmer:innen geprüft wird. Werden diese erfolgreich getäuscht, gilt der Turing Test bereits als bestanden. So hinterfragt eine mit dem Fachgebiet der künstlichen Intelligenz vertraute Person die Antworten der Versuchsteilnehmer:innen mehr als ein Laie dies tun würde.<sup>8</sup>

## **Künstliche Intelligenzen, die den Turing Test bestanden haben**

---

### Cleverbot

Mit dem 1988 vom britischen Informatiker Rollo Carpenter entwickelten Chatbot Cleverbot ist „eine Unterhaltung [...] ein klein wenig so [...], als würde man sich mit

der gesamten Gemeinschaft des Internets unterhalten.“ (Saenz, 2010) Dieser imitiert anhand von früheren 65 Millionen bereits geführten Unterhaltungen ein Gespräch. Beim Technischen Festival des indischen Instituts IIT, Guwahati, besteht dieser vermeintlich den Turing Test, da bei einer Abstimmung 59,3% von 1334 Zuschauern Cleverbot für einen Menschen hielten. Das Bestehen wird jedoch teils stark kritisiert, da das Publikum lediglich abstimmen durfte, es ihnen jedoch verwehrt war selber Fragen zu stellen. Zudem bezweifelt Rollo Carpenter selbst die richtige Interpretation von Turings Definition zum Anteil der Personen, die erfolgreich getäuscht hätten werden müssen.<sup>9 10</sup>

### Eugene Goostman

Bei der seit 2001 von Wladimir Weselow und Eugene Demchenko entwickelten künstlichen Intelligenz handelt es sich ebenfalls um einen Chatbot. Dieser gewann durch Nachahmung eines dreizehnjährigen ukrainischen Jungen dreimal den zweiten Platz des Loebner Preises. 2014 führte der Professor Kevin Warwick im Rahmen einer Veranstaltung der Royal Society, einer englischen Wissenschaftsakademie, einen Turing Test bei Eugene Goostman durch. Nachdem 33% der Teilnehmer überzeugt waren, dass es sich bei der Maschine um einen Menschen handelt, behauptete Kevin Warwick, diese habe den Test bestanden. Kritisiert wurde nachträglich der nicht korrekte Versuchsaufbau, da die Konversation zwischen der Teilnehmern lediglich 5 Minuten betrug, sodass das Bestehen kontrovers diskutiert wird. Darüber hinaus wurden bis heute weder Protokolle des Chatverlaufs offiziell

<sup>8</sup> Harnad, S. (kein Datum). *eprints.uk*. Von Searle's Chinese Room Argument: <https://eprints.soton.ac.uk/260424/1/chineseroom.html> abgerufen

<sup>9</sup> Carpenter, R. (2022). *cleverbot.com*. Abgerufen am 28. 02 2022 von <https://www.cleverbot.com>

<sup>10</sup> Saenz, A. (2010). *SingularityHub*. Abgerufen am 28. 02 2022 von <https://singularityhub.com/2010/01/13/cleverbot-chat-engine-is-learning-from-the-internet-to-talk-like-a-human/>

veröffentlicht noch der Test von anderen Forscher:innen bestätigt. Ein weiterer bemängelter Aspekt besteht in der Wahl der Person der künstlichen Intelligenz. Da diese einen dreizehnjährigen Jungen aus der Ukraine wiedergibt, scheint es oft valide, dass seine Antworten ausweichend sind und Rechtschreibfehler enthalten.<sup>11 12</sup>  
13

## Google Smartphone Assistant Duplex

Ein Teil der Google KI, genannt Google Duplex, hat zwar bis heute keinen offiziellen Turing Test absolviert, jedoch wird von verschiedenen Wissenschaftler:innen angenommen, dass sie diesen bestehen würde. Sie ist in der Lage Anrufe auszuführen oder Formulare auszufüllen. Ziel dabei ist es Termine bei Friseursalons zu vereinbaren, Eintrittskarten zu kaufen oder Tischreservierungen in Restaurants zu tätigen. Dies geschieht durch eine computergenerierte Stimme mit menschlicher Tonlage und absichtlichen Denkpausen. Sie wurde 2018 bei der Google I/O Keynote in San Francisco vorgestellt und seitdem kontinuierlich weiterentwickelt. So wurde ihr Einsatz auf den hauseigenen Kartendienst Google Maps erweitert, um Öffnungszeiten per Telefon abzufragen und diese anschließend in der Anwendung zu ergänzen. Darüber hinaus ist der Dienst auch in anderen englischsprachigen Ländern zur Verfügung gestellt worden.<sup>14</sup>

Kritiker bemängeln durch den Einsatz würden Menschen zu einem inaktiven Lebensstil animiert werden und es wäre trügerisch und sinnlos, die Stimme der künstlichen Intelligenz menschlich klingen

zu lassen. Google entgegnete, dass Duplex durch die gleichzeitige automatische Terminerstellung im persönlichen Kalender und eine damit verbundene Erinnerung zu einer Reduzierung der verpassten Termine führen würde und zudem einen Vorteil für Unternehmen ohne Onlinereservierungssystem habe.

Eine Nachforschung des Magazins Agios nach der Google I/O Keynote bei einem während der Präsentation zur Veranschaulichung gezeigten Anrufbeispiels führte zu seltsamen Ergebnissen. So reagierte ein Mitarbeiter des Friseursalons auf eine telefonische Nachfrage von ASIOS ob das Gespräch echt gewesen wäre, er müsse wieder zurück zur Arbeit und habe keine Zeit für eine Befragung.<sup>15</sup>

## Alternativen zum Turing Test

---

Es gibt zwei weitere Testmöglichkeiten, um das Vorhandensein menschlichen Denkvermögens bei künstlichen Intelligenzen zu überprüfen. Eine davon ist der Metzinger Test. Bei diesem Test wird der Computer aufgefordert sich an einer Diskussion zu beteiligen, ob es sich bei ihm um eine künstliche Intelligenz handelt oder nicht. Überzeugt die Maschine die anderen Gesprächsteilnehmer davon, dass sie selbst keine Maschine ist, so gilt der Metzinger Test als bestanden.

Eine weitere Testmöglichkeit ist das von Mark O. Riedl entwickelte Verfahren am Georgia Institut of Technology. Der Lovelace 2.0 Test ist benannt nach Ada Lovelace. Sie gilt in Fachkreisen als erste Programmiererin der Welt.<sup>15</sup>

Bei diesem Test wird neben Kreativität auch affektives Denken, Planungsfähigkeit

11 Kühl, E. (10. 06 2014). *Zeit Online*. Von Zeit Online: <https://www.zeit.de/digital/internet/2014-06/turing-test-eugene-goostman-kritik> abgerufen

12 Holland, M. (10. 06 2014). *heise.de*. Von heise.de:

13 University of Reading. (08. 06 2014). *University of Reading*. Von <https://www.reading.ac.uk/news-archive/press-releases/pr583836.html> abgerufen

14 Friebe, V. (21. 05 2018). *Medium.com*. Von Medium.com: <https://medium.com/@vanessa.friebe/knackt-google-duplex-als-erste-kunstliche-intelligenz-den-turing-test-464dccbcfb02>

15 Katzlberger, M. (14. 05 2018). *m bizbook*. Von m bizbook: <https://medianet.at/bizbook/user/5595/blog/hat-die-google-duplex-ki-gerade-den-turing-test-bestanden#> abgerufen

sowie Spracherkennung geprüft. Die Künstliche Intelligenz wird hierbei nach einer bestimmten Prämisse aufgefordert kreativ tätig zu werden. So wird ihr beispielweise gesagt, sie solle eine Geschichte erzählen, die einen vorgegebenen Ausgang hat oder ein Kunstwerk malen auf dem bestimmte Gegenstände zu sehen sind. Ist die Maschine dazu in der Lage, wird der Test als bestanden angesehen. Das Ergebnis soll hierbei von Expert:innen begutachtet werden. Bei erfolgreichem Absolvieren ist es zudem möglich die Grenzen der künstlichen Intelligenz durch eine komplexere Prämisse herauszufinden und somit auch mit anderen künstlichen Intelligenzen zu vergleichen.<sup>16</sup>

## **Fazit**

---

1950 definierte Alan Turing einen Test, der einen Ansatz bietet zwischen einem Menschen und einer Maschine zu unterscheiden. Aufgrund des hieraus resultierenden Anreizes, den Test zu bestehen, wurde die Forschung im Bereich der künstlichen Intelligenz erheblich vorangebracht. Er gilt bis heute als einer der wenigen Maßstäbe zur qualitativen Bewertung von künstlicher Intelligenz und wird daher auch in Zukunft für diesen Forschungsbereich von hoher Bedeutung sein.

<sup>16</sup> Metzinger, P. D. (kein Datum). *Heinz Nixdorf Museumsforum*. Von <https://www.hnf.de/veranstaltungen/events/paderborner-podium/computer-gehirn-und-bewusstsein/metzinger.html> abgerufen

Abbildung 1 Bildschirmaufnahme CAPTCHA	1
Abbildung 2 Grafik: Anordnung Turing Test	2

## Literaturverzeichnis

---

- Carpenter, R. (2022). *cleverbot.com*. Abgerufen am 28. 02 2022 von <https://www.cleverbot.com>
- Cassier, P. (2021). *Welt.de*. Abgerufen am 28. 02 2022 von <https://www.welt.de/geschichte/kopf-des-tages/article235850396/Alan-Turing-Die-absurde-Begnadigung-eines-Genies.html>
- Copeland, B. J. (kein Datum). *Britannica.com*. Abgerufen am 28. 02 2022 von <https://www.britannica.com/biography/Alan-Turing#ref704283>
- EU. (2018). *European Mathematical Information Service*. Abgerufen am 28. 02 2022 von <https://www.emis.de/community/history-of-mathematics/alan-turing/index.html>
- Friebe, V. (21. 05 2018). *Medium.com*. Von Medium.com: <https://medium.com/@vanessa.friebe/knackt-google-duplex-als-erste-kuenstliche-intelligenz-den-turing-test-464dccbcfb02> abgerufen
- Harnad, S. (kein Datum). *eprints.uk*. Von Searle's Chinese Room Argument: <https://eprints.soton.ac.uk/260424/1/chineseroom.html> abgerufen
- Holland, M. (10. 06 2014). *heise.de*. Von heise.de: <https://www.heise.de/newsticker/meldung/Eugene-und-der-angeblich-bestandene-Turing-Test-So-einfach-nun-dann-doch-nicht-2218151.html> abgerufen
- Kühl, E. (10. 06 2014). *Zeit Online*. Von Zeit Online: <https://www.zeit.de/digital/internet/2014-06/turing-test-eugene-goostman-kritik> abgerufen
- Katzlberger, M. (14. 05 2018). *m bizbook*. Von m bizbook: <https://medianet.at/bizbook/user/5595/blog/hat-die-google-duplex-ki-gerade-den-turing-test-bestanden#> abgerufen
- Kettern, A. (2014). *Lebendiges Museum Online*. Abgerufen am 28. 02 2022 von <https://www.dhm.de/lemo/biografie/arthur-chamberlain>
- Luis von Ahn, M. B. (kein Datum). *International Association for Cryptologic Research*. Von <https://iacr.org/archive/eurocrypt2003/26560294/26560294.pdf> abgerufen
- Metzinger, P. D. (kein Datum). *Heinz Nixdorf Museumsforum*. Von <https://www.hnf.de/veranstaltungen/events/paderborner-podium/computer-gehirn-und-bewusstsein/metzinger.html> abgerufen
- Redaktion ComputerWeekly.de. (2021). *WhatIs.de*. Abgerufen am 17. 02 2022 von <https://whatis.techtarget.com/de/definition/Turing-Test>
- Saenz, A. (2010). *SingularityHub*. Abgerufen am 28. 02 2022 von <https://singularityhub.com/2010/01/13/cleverbot-chat-engine-is-learning-from-the-internet-to-talk-like-a-human/>
- Scherschel, F. A. (02 2017). *heise.de*. Von heise.de: <https://www.heise.de/security/meldung/DDoS-Untersuchung-Angriffe-werden-zum-Problem-fuer-die-Allgemeinheit-3631903.html> abgerufen

Turing, A. M. (1950, 10). *Mind*. Retrieved from The Mind Association:

<https://academic.oup.com/mind/article/LIX/236/433/986238>

Universität Oldenburg. (2008/2009). *Informatik Universität Oldenburg*. Abgerufen am 17. 02

2022 von [http://www.informatik.uni-oldenburg.de/~iug08/ki/Grundlagen\\_Turing-](http://www.informatik.uni-oldenburg.de/~iug08/ki/Grundlagen_Turing-Test.html)

[Test.html](http://www.informatik.uni-oldenburg.de/~iug08/ki/Grundlagen_Turing-Test.html)

University of Reading. (08. 06 2014). *University of Reading*. Von

<https://www.reading.ac.uk/news-archive/press-releases/pr583836.html> abgerufen