

## VORWISSENSCHAFTLICHE ARBEIT

---

# Maschinelle Werteanpassung bei einer hypothetischen allgemeinen künstlichen Intelligenz

---

Autor:

Franz Srambical

Betreuungslehrerin:

Prof. Mag. Kurt Rauch

Klasse:

8C

Entwurf:

10. August 2019

# Abstract

Der Zusammenfassungstext kommt hier her. Abstract ist kein Vorwort und keine Einleitung! Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

# Vorwort

Das Vorwort ist optional: d. h. man muss kein Vorwort schreiben! Wer will, kann das in dieser Form tun. Am Ende sollten Ort, Datum und der Name des Autors des Vorworts angegeben werden.<sup>1</sup>

Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

Wien am 10. August 2019

Franz Srambical

---

<sup>1</sup> Vgl. WEIGL, Huberta. *Vorwort*. URL: [http://www.ahs-vwa.at/pluginfile.php/31/mod\\_data/content/1315/02-VWA-Vorwort.pdf](http://www.ahs-vwa.at/pluginfile.php/31/mod_data/content/1315/02-VWA-Vorwort.pdf) (besucht am 3. 2. 2017).

# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>6</b>
<b>2. Allgemeine künstliche Intelligenz</b>	<b>7</b>
2.1. Definition von Intelligenz . . . . .	7
2.2. Definition von künstlicher Intelligenz . . . . .	7
2.3. Definition von allgemeiner künstlicher Intelligenz . . . . .	8
2.4. Wann wird es sie geben? . . . . .	8
2.5. Die These der Intelligenzexplosion . . . . .	8
<b>3. Mein allererstes Kapitel</b>	<b>9</b>
3.1. Mein allererstes Unterkapitel . . . . .	9
3.1.1. Ein Unterunterkapitel . . . . .	9
3.1.2. Zitierbeispiele . . . . .	9
3.2. Einheiten . . . . .	9
3.3. Text . . . . .	10
<b>4. Probleme einer allgemeinen künstlichen Intelligenz</b>	<b>14</b>
4.1. Fehlerhafte Vorstellungen einer KI-Katastrophe . . . . .	14
4.1.1. Bösertige KI . . . . .	14
4.1.2. KI, die ein Bewusstsein erlangt . . . . .	14
4.1.3. Roboter als Auslöser einer Katastrophe . . . . .	14
4.2. Gesamtmenschheitlicher Konsens über gemeinsame Werte . . . . .	14
4.3. Wertekodierung in einer Programmiersprache . . . . .	14
4.3.1. Statische Wertekodierung . . . . .	14
4.3.2. Dynamische maschinelle Werteanpassung . . . . .	14
4.3.3. Bestärkendes Lernen . . . . .	14
4.3.4. Reziprok-bestärkendes Lernen . . . . .	14
4.4. Sichere und vertrauenswürdige KI . . . . .	14
4.5. KI-Ethik . . . . .	14
<b>5. Lösungsansätze</b>	<b>15</b>
5.1. Bestärkendes Lernen . . . . .	15
5.1.1. Reziprok-bestärkendes Lernen . . . . .	15
5.2. Mensch-Maschinen-Interface . . . . .	15
5.3. Hirnemulation . . . . .	15
<b>Literaturverzeichnis</b>	<b>16</b>
Print-Quellen . . . . .	16
Internet-Quellen . . . . .	16
Interviews . . . . .	16

Abbildungsverzeichnis	17
Tabellenverzeichnis	17
A. Interview mit Ex Perte	18
B. Hier könnte Ihr Anhang stehen	20
Erklärungen	21

# 1. Einleitung

Ich möchte diese Arbeit mit einer Analogie beginnen.

Nehmen wir an, Sie hätten eine Pille. Sie wissen, dass diese Pille entweder alle Krankheiten, Ungleichheiten, Schmerzen - kurz gesagt alles "Schlechte" - auf der Welt beseitigen, oder aber die gesamte Menschheit auslöschen könnte. Würden Sie die Pille nehmen? Wahrscheinlich nicht. Nun stellen wir uns vor, nicht Sie hätten die Pille, sondern sie befände sich irgendwo auf unserem Planeten und *jeder* könnte sie finden. Wären Sie besorgt? Wahrscheinlich schon.

Dann sollten Sie dies schon heute sein, denn eine solche Pille wird es geben - ihr Name: allgemeine künstliche Intelligenz.

Was rechtfertigt eine so technovolatile Haltung wie diese?

*„There are all sorts of extreme forces coming onto the game board that were not there before. To expect them to all fail or exactly cancel out for the purpose of making the outcome normal would be one heck of a coincidence.“<sup>1</sup>*

Jede technologische Neuentdeckung bedeutet in erster Linie Veränderung. Die Erfindungen der letzten Jahrhunderte hatten mehrheitlich positive Auswirkungen zur Folge, sonst wäre unser Lebensstandard heute nicht der höchste in der Menschheitsgeschichte.<sup>2</sup> So ermutigend das auch klingt, so dürfen wir nicht einfach nach dem Trend der Vergangenheit in die Zukunft extrapolieren, sondern müssen - so Richard A. Easterlin - versuchen, die Kräfte zu verstehen, die für den Anstieg der Lebensqualität verantwortlich sind.<sup>3</sup> Was eine allgemeine künstliche Intelligenz betrifft, müssen wir sie nicht nur verstehen, sondern auch lenken können, um das Wohlbefinden der Spezies Mensch nicht zu gefährden, sondern zu bestärken.

---

<sup>1</sup> Eliezer Yudkowsky on Intelligence Explosion - YouTube. URL: <https://www.youtube.com/watch?v=D6peN9LiTWA> (besucht am 7. 8. 2019), 30:51–31:07.

<sup>2</sup> Vgl. EASTERLIN, Richard A. „The Worldwide Standard of Living since 1800“. In: *The Journal of Economic Perspectives* 14.1 (2000), S. 7–26. ISSN: 0895-3309. URL: <https://www.jstor.org/stable/2647048> (besucht am 9. 8. 2019), S. 22–23.

<sup>3</sup> Vgl. ebd., S. 23.

## 2. Allgemeine künstliche Intelligenz

### 2.1. Definition von Intelligenz

Seit Jahrhunderten versuchen Wissenschaftler und Laien gleichermaßen eine Definition für den Intelligenzbegriff zu finden. Da bis heute keine Definition ihre Vollständig- oder Richtigkeit beweisen konnte, wird in dieser Arbeit der Einfachheit halber versucht, den Begriff durch Beobachtungen zu erklären, wie Eliezer Yudkowsky in dem Podcast “AI: Racing Toward the Brink” vorschlägt.

1. Menschen waren auf dem Mond.
2. Mäuse waren nicht auf dem Mond.

Yudkowsky wählt dieses Beispiel um zu demonstrieren, dass Menschen auch Orte erreichen, wofür die natürliche Selektion sie nicht vorbereitet hat. Daraus könne geschlossen werden, dass Menschen *intelligenter* als Mäuse sind, weil sie *domänenübergreifend* arbeiten können. Deshalb sei das *domänenübergreifende* Erlernen neuer Fähigkeiten ein zentraler Teil des Intelligenzbegriffs.<sup>1</sup>

### 2.2. Definition von künstlicher Intelligenz

*„Artificial intelligence (AI)—defined as a system’s ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation“<sup>2</sup>*

Laut angeführter Definition muss eine künstliche Intelligenz nicht nur Daten richtig interpretieren, sondern auch die dadurch gewonnen Erkenntnisse mittels *dynamischer Anpassung* zur Erreichung bestimmter Ziele benutzen können.

Diese Definition enthält die Idee des *domänenübergreifenden* Lernens im Gegensatz zum oben beschriebenen Ansatz zur Intelligenzerklärung nicht, was laut Experten jedoch nicht an einer unvollständigen Definition liegt, sondern vielmehr daran, dass

---

<sup>1</sup> Vgl. YUDKOWSKY, Eliezer. *AI: Racing Toward the Brink*. Sam Harris. Feb. 2018. URL: <https://saharris.org/podcasts/116-ai-racing-toward-brink/>, 06:01–09:49.

<sup>2</sup> KAPLAN, Andreas und HAENLEIN, Michael. „Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence“. In: *Business Horizons* 62.1 (1. Jän. 2019), S. 15–25. ISSN: 0007-6813. DOI: 10.1016/j.bushor.2018.08.004. URL: <http://www.sciencedirect.com/science/article/pii/S0007681318301393> (besucht am 6.8.2019).

wir den Begriff der KI in einer Art gebrauchen, wofür er nicht vorgesehen war. Um Missverständnisse zu vermeiden, wird für KI wie sie heutzutage bereits in Benutzung ist der Begriff schwache KI (engl. *weak AI* oder *narrow AI*) verwendet.<sup>3</sup> Dieser beschreibt eine *domänenspezifische* KI.

## 2.3. Definition von allgemeiner künstlicher Intelligenz

## 2.4. Wann wird es sie geben?

## 2.5. Die These der Intelligenzexplosion

---

<sup>3</sup> Vgl. BOSTROM, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press, 3. Juli 2014. 328 S. ISBN: 978-0-19-967811-2, S. 18–19.



# 3. Mein allererstes Kapitel

## 3.1. Mein allererstes Unterkapitel

### 3.1.1. Ein Unterunterkapitel

Hier ist ein Bisschen **Text**. „Hier direkte Rede.“

Neuer Absatz. **Hervorgehobener Text**. Grotesk-Schrift KAPITÄLCHEN, *Kursiv-schrift*.

Formeln lassen sich leicht setzen.  $y = x^2$  entweder in der Zeile oder als eigene Display-Umgebung.

$$f(x) = x^2 \quad (3.1)$$

### 3.1.2. Zitierbeispiele

Neulich habe ich mit Ex Perte darüber gesprochen. Er hat meine Ansichten bestätigt. „*Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen.*“<sup>1</sup>

„*Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen.*“<sup>2</sup>

Ein weiteres Zitat: „*Niemand hat das je gesagt.*“<sup>3</sup>

## 3.2. Einheiten

Heute hatte es nur 7 °C. Der Bisamberg ist ca. 150 m hoch. Die Lichtgeschwindigkeit ist ca.  $3,0 \times 10^8 \text{ m s}^{-1}$ .<sup>4</sup>

„*Das sage ich nicht.*“<sup>5</sup> „*Das sage ich nicht.*“<sup>6</sup>

---

<sup>1</sup> PERTE, Ex. persönliches Gespräch mit dem Verfasser, Transkript in Anhang A auf Seite 18. 25. Juli 2017, S. 18.

<sup>2</sup> Ebd., S. 18.

<sup>3</sup> NIEMAND. *Nichts*. 1983.

<sup>4</sup> Vgl. LESSING, Gotthold Ephraim. *Gedichte*. Hrsg. von Holzinger, Michael. 3. Aufl. CreateSpace Independent Publishing Platform, 27. Feb. 2013. URL: [http://static.zeno.org/ebooks/Lessing\\_Gotthold\\_Ephraim\\_-\\_Gedichte\\_-\\_Zeno.org\\_ISBN\\_9783843047753.epub](http://static.zeno.org/ebooks/Lessing_Gotthold_Ephraim_-_Gedichte_-_Zeno.org_ISBN_9783843047753.epub) (besucht am 2.4.2017), S. 31; FEYNMAN, Richard P. *Sie belieben wohl zu scherzen, Mr. Feynman! Abenteuer eines neugierigen Physikers*. 16. Aufl. Piper, 1996, S. 42.

<sup>5</sup> FEYNMAN, *Sie belieben wohl zu scherzen, Mr. Feynman!*, S. 42.

<sup>6</sup> Ebd., S. 42.



Abbildung 3.1.: Das Logo.

### 3.3. Text

Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

Siehe auch Abb. 3.1. Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich

zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschlief ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschlief ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschlief ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.



Abbildung 3.2.: Das Logo zum zweiten Mal.

Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

Hier ist ein Absatz voll sinnlosem Text. Bitte erst nach diesem Absatz weiterlesen. Hier kommt nichts mehr. Es folgen unterschiedlich lange Wörter. Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

## 4. Probleme einer allgemeinen künstlichen Intelligenz

### 4.1. Fehlerhafte Vorstellungen einer KI-Katastrophe

#### 4.1.1. Böartige KI

#### 4.1.2. KI, die ein Bewusstsein erlangt

#### 4.1.3. Roboter als Auslöser einer Katastrophe

### 4.2. Gesamtmenschheitlicher Konsens über gemeinsame Werte

### 4.3. Wertekodierung in einer Programmiersprache

#### 4.3.1. Statische Wertekodierung

#### 4.3.2. Dynamische maschinelle Werteanpassung

#### 4.3.3. Bestärkendes Lernen

#### 4.3.4. Reziprok-bestärkendes Lernen

### 4.4. Sichere und vertrauenswürdige KI

### 4.5. KI-Ethik

# **5. Lösungsansätze**

## **5.1. Bestärkendes Lernen**

### **5.1.1. Reziprok-bestärkendes Lernen**

## **5.2. Mensch-Maschinen-Interface**

## **5.3. Hirnemulation**

# Literaturverzeichnis

## Print-Quellen

- BOSTROM, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press, 3. Juli 2014. 328 S. ISBN: 978-0-19-967811-2.
- EASTERLIN, Richard A. „The Worldwide Standard of Living since 1800“. In: *The Journal of Economic Perspectives* 14.1 (2000), S. 7–26. ISSN: 0895-3309. URL: <https://www.jstor.org/stable/2647048> (besucht am 9.8.2019).
- FEYNMAN, Richard P. *Sie liebten wohl zu scherzen, Mr. Feynman! Abenteuer eines neugierigen Physikers*. 16. Aufl. Piper, 1996.
- KAPLAN, Andreas und HAENLEIN, Michael. „Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence“. In: *Business Horizons* 62.1 (1. Jän. 2019), S. 15–25. ISSN: 0007-6813. DOI: 10.1016/j.bushor.2018.08.004. URL: <http://www.sciencedirect.com/science/article/pii/S0007681318301393> (besucht am 6.8.2019).
- LESSING, Gotthold Ephraim. *Gedichte*. Hrsg. von Holzinger, Michael. 3. Aufl. CreateSpace Independent Publishing Platform, 27. Feb. 2013. URL: [http://static.zeno.org/ebooks/Lessing,\\_Gotthold\\_Ephraim\\_-\\_Gedichte\\_-\\_Zeno.org\\_ISBN\\_9783843047753.epub](http://static.zeno.org/ebooks/Lessing,_Gotthold_Ephraim_-_Gedichte_-_Zeno.org_ISBN_9783843047753.epub) (besucht am 2.4.2017).
- NIEMAND. *Nichts*. 1983.

## Internet-Quellen

- Eliezer Yudkowsky on Intelligence Explosion - YouTube. URL: <https://www.youtube.com/watch?v=D6peN9LiTWA> (besucht am 7.8.2019).
- WEIGL, Huberta. *Vorwort*. URL: [http://www.ahs-vwa.at/pluginfile.php/31/mod\\_data/content/1315/02-VWA-Vorwort.pdf](http://www.ahs-vwa.at/pluginfile.php/31/mod_data/content/1315/02-VWA-Vorwort.pdf) (besucht am 3.2.2017).

## Interviews

- PERTE, Ex. persönliches Gespräch mit dem Verfasser, Transkript in Anhang A auf Seite 18. 25. Juli 2017.



# Abbildungsverzeichnis

3.1. Das Logo. Bildquelle: NIEMAND. <i>Nichts</i> . Jän.–Dez. 2018 . . . . .	10
3.2. Das Logo zum zweiten Mal. Bildquelle: NIEMAND, <i>Nichts</i> . . . . .	11

# Tabellenverzeichnis

# A. Interview mit Ex Perte

Interviewdatum: 25. Juli 2017

I Könnten Sie mir einen Absatz sinnlosen Texts formulieren?

B Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

I Könnten Sie mir einen Absatz sinnlosen Texts formulieren?

B Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

I Könnten Sie mir einen Absatz sinnlosen Texts formulieren?

B Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

I Könnten Sie mir einen Absatz sinnlosen Texts formulieren?

B Die Abruchbirne kringelte ihre Hürde in eine unbekannte Überschwänglichkeit, um so die Sitzordnung der Fensterscheiben in der unteren Waldkante zu verjubeln. Niemandem ist absichtlich zu kürzen, wessen Woligkeit hier in abermaligem Abgesang aufgeschlagen ist. Deswegen soll dieser aber nicht heimreisen, sondern abermals die Einigkeit des Urwalds in die aufgeregte Höhensonne schlagen. Wenn nicht der hiesige Erdball des aberwitzigen Ungemachs aufgedrungene Kröte wäre, entschieße ich mich zu unsachgemäßem Handlungsablauf. Wiegleich zudem ein weiterer Honigkuchen ausbricht.

**B. Hier könnte Ihr Anhang stehen**

# Erklärungen

## Selbstständigkeitserklärung

Ich erkläre, dass ich diese vorwissenschaftliche Arbeit eigenständig angefertigt und nur die im Literaturverzeichnis angeführten Quellen und Hilfsmittel benutzt habe.

Wien, 10. August 2019

Franz Srambical

## Informatikschwerpunkt

Die vorliegende Arbeit erfüllt die Kriterien zur Abbildung des Informatikschwerpunktes an der De La Salle Schule Strebersdorf, AHS.

**Begründung:** Die Arbeit wurde in L<sup>A</sup>T<sub>E</sub>X mit entscheidenden Kenntnissen zum Quelltext verfasst.

*Geprüft am ... durch Mag. Rainer Zufall und Mag. Ernst Haft*