

# Maschinen lernen lassen: Technologien für die Zukunft

Making machine learning, AI and emerging technologies work for us.

Ein Bericht des Center for the Cultivation of Technology  
verfasst von Zara Rahman und Julia Kloiber, September 2018

Zusammenfassung	2
1. Einleitung	2
2. Künstliche Intelligenz und Maschinelles Lernen	3
2.1 Begriffsdefinition	3
2.2 Warum sind diese Technologien von Relevanz?	4
2.3 Herausforderung	5
2.3.2 Investitionen	7
2.3.3 Verständnis	7
Sprache	8
2.3.4 Mangel an ExpertInnen	8
2.3.5 Weitere Herausforderungen	9
2.4 Chancen	10
2.5 Kontroversen	12
2.6 Fragestellungen, die sich aus den Herausforderungen und Chancen ergeben	12
3. Vorschlag des Themenschwerpunkts	13
3.1 Beschreibung	13
3.2 Themenvorschlag	13
3.3 Chancen des Schwerpunktes	14
3.4 Trends	15
4. Empfehlungen für die Umsetzung	15
5. Ziele	16
5.1 Interne Ziele	16
5.2 Externe Ziele	16

# Zusammenfassung

Der vorliegende Bericht basiert auf umfangreicher Sekundärforschung und informellen Gesprächen mit ExpertInnen. Er zeigt Chancen und Herausforderungen für einen verantwortungsvollen Einsatz von Intelligenzen Systemen auf.

Für den Themenschwerpunkt der 5. Runde des Prototype Funds werden im Folgenden neue/emergente Technologien untersucht. Viele Unternehmen und wissenschaftliche Einrichtungen erforschen und entwickeln so z.B: aktuell Systeme im Bereich Künstliche Intelligenz. Maschinelles Lernen mit seinen Unterkategorien ist bereits vielerorts im Einsatz: Die schnelle Entwicklung, die Größenordnung, die vielseitigen Anwendungsszenarien und der teilweise rudimentäre Entwicklungsstand machen diese emergenten Technologien besonders interessant für eine Analyse. Künstliche Intelligenz vereint viele Herausforderungen und Chancen, die man auch bei anderen digitalen Technologien beobachten kann. Erkenntnisse aus dem vorliegenden Bericht lassen sich daher auch auf andere Bereiche übertragen.

## 1. Einleitung

Die Entwicklung neuer Technologien verläuft rasant. Sie nehmen eine immer größere Rolle im Leben von Menschen ein und haben Auswirkungen auf viele Aspekte des Zusammenlebens. Nur mit Hilfe von Technologie, die verantwortungsvoll entwickelt wird, kann eine Zukunft gestaltet werden, die inklusiv und offen für alle ist. Neue Entwicklungen durchlaufen zumeist eine Phase des Hypes, eine Phase, in der sich alle auf bestimmte Ansätze stürzen und versuchen diese auf allerlei Arten von Problemen anzuwenden. Die Zyklen, in denen Technologie entwickelt und emergent wird, werden immer kürzer – vieles wird veröffentlicht, ohne dass man einen genauen Überblick darüber hat, welche Risiken und Auswirkungen damit einhergehen können. Große Unternehmen und Forschungseinrichtungen sind Treiber in Sachen Technologieentwicklung, sie können auf die nötigen finanziellen, personellen und materiellen Ressourcen zurückgreifen. Im Bereich von datengetriebenen Technologien wie Künstlicher Intelligenz und Maschinellem Lernen sitzen die Unternehmen auf den für die Entwicklung essentiellen Datensätzen. Dazu, wie emergente Technologien für das Gemeinwohl eingesetzt werden können, gibt es nur eine überschaubare Anzahl an Beispielen. Dabei ist es wichtig, und gesellschaftlich unabdingbar, neue Entwicklungen auf ihre Risiken und Chancen hin zu untersuchen, sie zu erklären und zugänglich zu machen, sodass möglichst viele Menschen davon profitieren können. Open Source und Transparenz sind dabei wichtige Bausteine, wenn es um Rechenschaftspflicht und faire Nutzung geht.

## 2. Künstliche Intelligenz und Maschinelles Lernen

### 2.1 Begriffsdefinition

In diesem Papier werden die Bereiche Künstliche Intelligenz und Maschinelles Lernen untersucht. Da es an Begriffsdefinitionen mangelt, starten wir mit einer Einordnung der Begriffe.

**Künstliche Intelligenz** (*KI, artifizielle Intelligenz, artificial Intelligence*) ist ein Teilgebiet der Informatik, welches zum Ziel hat, Systeme zu entwickeln, die wie Menschen denken und sich „wie Menschen“ verhalten – Systeme, die rational denken und sich rational verhalten.<sup>1</sup> Der Begriff, insbesondere Rationalität, ist dabei nicht eindeutig abgrenzbar, da es an einer genauen Definition von „Intelligenz“ mangelt.

Grundsätzlich lässt sich KI-Technologie in zwei Typen aufteilen: Schwache Künstliche Intelligenz (*Narrow AI*) und Starke Künstliche Intelligenz (*General AI*). General AI ist, was häufig unter dem Begriff Künstliche Intelligenz verstanden wird: Technologie, die dieselben Eigenschaften aufweist wie menschliche Intelligenz, die „denken“ kann wie Menschen. Die Idee von General AI wurde durch Science-Fiction-Erzählungen populär, in der Realität ist die Wissenschaft jedoch noch weit von dieser Art von Entwicklung entfernt.

Narrow AI ist KI-Technologie, die ein bestimmtes Set an Aufgaben ausführen kann. Der Bereich ist oft auf einen eng definierten Kontext begrenzt. Zum Beispiel: Schach und Go spielen oder Vorhersagen im Bereich Verkauf und Wetter treffen. Narrow AI spiegelt den aktuellen Entwicklungsstand im Bereich Künstliche Intelligenz wider. Auch Bilderkennungsverfahren und Sprachanalyse fallen unter den Begriff der Narrow AI.

**Maschinen lernen (machine learning)** ist eine der häufigsten Techniken in der Künstlichen Intelligenz. Es ist ein Verfahren, bei dem durch die systematische Auswertung von Daten ein Modell erstellt wird, mit dem sich Vorhersagen treffen oder Muster erkennen lassen.<sup>2</sup> Maschinelles Lernen ist ein selbstadaptiver Algorithmus.

Während Künstliche Intelligenz sich auf die Idee einer holistischen intelligenten Entität bezieht, kann Maschinelles Lernen als untergeordneter Prozess verstanden werden, der sich auf die Lösung spezifischer Aufgaben fokussiert.

---

<sup>1</sup> As computer scientists Stuart Russell and Peter Norvig observe in their widely used AI textbook, the history of artificial intelligence has not produced a clear definition of AI but rather can be seen as variously emphasizing four possible goals: „systems that think like humans, systems that act like humans, systems that think rationally, systems that act rationally.“ Stuart J. Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, Englewood Cliffs, NJ: Prentice Hall, 1995: 27.

<sup>2</sup> AI Pattern Language by Data and Society, <https://autonomy.datasociety.net/patternlanguage/>

## 2.2 Warum sind diese Technologien von Relevanz?

Künstliche Intelligenz verspricht großes Potenzial in einer Reihe von Gebieten: in der Medizin, der Bildung, im Verkehr, der Wissenschaft und vielen weiteren, allgemein: im Data Mining. In den letzten Jahren gab es große Fortschritte in der Entwicklung dieser Systeme, was zur Entstehung nützlicher Anwendungen geführt hat. Intelligente Systeme sind eine wichtige Komponente von Anwendungen zur Spracherkennung, in der Übersetzung, in Suchmaschinen und bei Spam-Filtern. Auch im Bereich von autonomen Verkehrsmitteln, sowie im Bereich der digitalen Assistenz für ÄrztInnen und Pflegekräfte, wird Künstliche Intelligenz in der Zukunft wahrscheinlich eine wichtige Rolle spielen.<sup>3</sup>

Während in manchen Bereichen noch einiges an Entwicklung und Forschung passieren muss, bevor die Technologien einsatzbereit sind, gibt es Felder in denen intelligente Systeme bereits seit Jahren im Einsatz sind.

Automatisierte Entscheidungsfindung wird beispielsweise schon seit Jahren im Recruiting von Angestellten und der Vergabe von Krediten eingesetzt.<sup>4</sup> Auch öffentliche Einrichtungen experimentieren zunehmend mit intelligenten Systemen, die ihnen dabei helfen, Infrastruktur zu planen und zu managen, nach Steuerhinterziehung zu fahnden und die Auslastung und Ressourcen von Schulen und Krankenhäuser zu steuern. Im Bereich der Strafverfolgung werden neue Technologien aus dem Bereich der Gesichtserkennung eingesetzt, um StraftäterInnen ausfindig zu machen.

In der öffentlichen Wahrnehmung von Künstlicher Intelligenz lässt sich beobachten, was Science-Fiction Schriftsteller Arthur C. Clarke mit folgendem Zitat beschreibt: „Any sufficiently advanced technology is indistinguishable from magic.“<sup>5</sup> (Jede ausreichend fortgeschrittene Technologie ist von Magie nicht zu unterscheiden) Der Hype um eine Technologie, kombiniert mit mangelndem Wissen und Forschung zu deren genauer Funktionsweise und Implementierung kann Nachteile mit sich bringen.

Kurz: Künstliche Intelligenz und andere emergente Technologien bergen gleichzeitig großes Potenzial und sehr reale Risiken. Im vorliegenden Bericht werden die Chancen und Herausforderungen dieser Systeme erörtert und daraus geschlussfolgert, in welchen Bereichen es an noch mehr Forschung und Entwicklung bedarf.

---

<sup>3</sup> The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation, <https://arxiv.org/pdf/1802.07228.pdf>

<sup>4</sup> Artificial Intelligence: Real Public Engagement; RSA's Forum for Ethical AI: Brhmie Balaram, Tony Greenham and Jasmine Leonard. <https://www.thersa.org/discover/publications-and-articles/reports/artificial-intelligence-real-public-engagement>

<sup>5</sup> Profiles of the Future: An Inquiry into the Limits of the Possible. Clarke, Arthur C., 1973

## 2.3 Herausforderung

Die Automatisierung von Prozessen und das Auslagern von Entscheidungen an Maschinen birgt Risiken, die in der Forschung bereits skizzenhaft beschrieben wurden. WissenschaftlerInnen wie Virginia Eubanks haben gezeigt, dass der Einsatz von Intelligenten Systemen von Regierungen dazu geführt hat, dass Menschen zu unrecht verfolgt und verurteilt wurden. In ihrem Buch „Automating Inequality: How High-Tech Tools Profile, Punish, and Police the Poor“<sup>6</sup> (Die Automatisierung der Ungleichheit: Wie High-Tech-Werkzeuge die Armen bewerten, bestrafen und überwachen), beschreibt Eubanks die teilweise schwerwiegenden Konsequenzen für sozial schlechter gestellte Menschen.

Eine der zugrunde liegenden Ursachen ist laut Eubanks die Tatsache, dass Technologien wie Künstliche Intelligenz Daten aus der Vergangenheit heranziehen, um Schlüsse auf die Zukunft zu ziehen. Erlernete Muster werden so auf die Zukunft projiziert.

Viele der beschriebenen Fälle in Eubanks' Buch beziehen sich auf Entscheidungen, die über Individuen getroffen werden. Rückschlüsse auf Verhalten und Kontext werden dabei aus dem Verhalten Dritter gezogen. In den Anwendungsfällen wird häufig das Schlimmstmögliche angenommen, ohne auf spezielle Umstände einzugehen. Verschärft wird die Situation oft durch einen Mangel an Verantwortlichkeit und die Möglichkeit Entscheidungen anzufechten oder nachzuvollziehen. Der Mangel an Rechenschaft bei automatisierten Entscheidungen, ist nicht neu und bezieht sich auch nicht ausschließlich auf KI-Technologien. Neu sind die Geschwindigkeit und die Größenordnung in der Entscheidungen getroffen werden können. Diese Faktoren verschärfen die Problematik.

### 2.3.1 Fairness und Verzerrung

Wie eben beschrieben spielen algorithmische Entscheidungsverfahren eine wichtige Rolle beim Einsatz und der Entwicklung von KI-Technologien.

Von wem werden Modelle und Technologien entwickelt, und zu welchem Zweck? Auf welche Daten wird dabei zurückgegriffen und wie werden Ergebnisse interpretiert? All diese Fragen sind wichtig, um Neigungen und Verzerrungen in den Ergebnissen erkennen und beurteilen zu können.

Viele Daten, mit denen eine Künstliche Intelligenz trainiert wird, werden von Menschen ausgewählt und klassifiziert. Das bedeutet, dass alle Datensammlungen höchst subjektiv sind. Die Frage, was gemessen wird, wie und wann; wie unterschiedliche Daten kategorisiert und eingesetzt werden; was nicht erhoben wird – all das sind Fragen und Entscheidungen, die von

---

<sup>6</sup> <https://us.macmillan.com/books/9781250074317>

einzelnen Menschen oder Unternehmen getroffen werden. Diese Entscheidungen unterliegen bestimmten Wertungen, Neigungen und Prämissen und sind damit nie neutral.

Künstliche Intelligenz und Maschinelles Lernen beruhen auf Trainingsdaten. Das sind Datensätze, auf deren Basis Muster und Modelle trainiert werden. Die Modelle bzw. die aus ihnen generierte Muster werden im Folgenden wiederum auf weitere Datensätze angewandt – der Algorithmus lernt aus seinen vorherigen Entscheidungen. Die Tatsache, dass Maschinelles Lernen und Künstliche Intelligenz auf bereits existierenden Datensammlungen basieren, erhöht das Risiko, dass Vorurteile und Tendenzen aus der Vergangenheit in die Ergebnisse übertragen werden.

Um KI-Systeme zu trainieren, benötigt man große Datenmengen. Die besten Datensammlungen und Modelle besitzen große Unternehmen wie Google, Amazon und Facebook. Neue KI-Modelle werden aus diesem Grund aktuell hauptsächlich von großen Unternehmen entwickelt und trainiert. Damit kleinere Player Schritt halten können, empfiehlt das Strategiepapier des französischen Abgeordneten Cédric Villani mit dem Titel „For a meaningful artificial intelligence“<sup>7</sup>, dass die wichtigsten Interessengruppen ihre Daten „zusammenlegen“ sollten, um gemeinsam davon zu profitieren. Das Papier argumentiert, dass der Vorsprung von US-Technologie-Riesen (Google, Amazon, Microsoft, Apple und Facebook) als auch der von chinesischen Unternehmen (Baidu, Alibaba, Tencent und Xiaomi) auf ihre Datenerfassungspraktiken zurückzuführen ist – und damit auch auf ihre Fähigkeit, diese Daten nach Belieben zu nutzen. Damit andere Unternehmen und Stakeholder an KI-Entwicklung arbeiten und diese vorantreiben können, sind verantwortungsvolle Methoden des Datenaustausches notwendig. Diese Art des Datenaustausches wird derzeit jedoch nicht praktiziert, sowohl aus Datenschutzgründen als auch aus technischen Gründen. Abgesehen von wenigen Ausnahmen, teilen sich öffentliche und private Akteure in der Regel keine großen Datenmengen, da der Zugriff auf Daten bisher als Wettbewerbsvorteil angesehen wird. Um dies zu ändern, so argumentiert der Bericht, müssen die Beteiligten (öffentliche und private) Allianzen eingehen und Möglichkeiten für einen datenschutzkonformen Datenaustausch entwickeln.

Für einen verantwortungsvollen Einsatz von Künstlicher Intelligenz ist es wichtig zu verstehen, wie Modelle entstanden sind und mit welchen Daten sie trainiert wurden. Ein Mangel an Verständnis darüber, wie diese Technologie funktioniert, führt häufig zu Missinterpretationen und dazu, dass Ergebnisse nicht kritisch genug hinterfragt werden. Ohne das Wissen von EntwicklerInnen und ExpertInnen, die den Kontext kennen und ihn navigieren können, ist es schwierig festzustellen, an welchen Stellen es zu Verzerrungen und falschen Darstellungen kommt. Zu bestimmen, wie Fairness im Kontext von Algorithmen und Künstlicher Intelligenz aussehen kann, ist eine komplexe Herausforderung, der sich momentan einige AkademikerInnen und Organisationen widmen. Unter ihnen die Konferenz „Fairness, Accountability and Transparency in Machine Learning“<sup>8</sup> und der gemeinnützige Verein Algorithm Watch<sup>9</sup>.

---

<sup>7</sup> [https://www.aiforhumanity.fr/pdfs/MissionVillani\\_Report\\_ENG-VF.pdf](https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf)

<sup>8</sup> <https://www.fatml.org/>

<sup>9</sup> <https://algorithmwatch.org/en/>

Einige der hier beschriebenen Herausforderungen treffen auch auf andere Technologien zu – so kann der Einsatz zu manipulativen und schädlichen Zwecken nie gänzlich ausgeschlossen werden, egal wie nobel die Vorsätze bei der Entwicklung sind. Ein großer Unterschied bei KI-Systemen sind die Geschwindigkeit und die Größenordnung, mit der sich negative Auswirkungen verbreiten können, zum Beispiel wenn Konzerne oder Regierungen Gesichtserkennungstechnologie einsetzen.

### 2.3.2 Investitionen

Der größte Teil an Investitionen in Künstliche Intelligenz kommt derzeit aus dem privaten Sektor. Motivation und Anreize für diese Investitionen stehen häufig im Gegensatz zu gesellschaftlichen Interessen. Für den öffentlichen Sektor und sozial orientierte Initiativen ist es daher von entscheidender Bedeutung, bei der Entwicklung der Systeme mit am Tisch zu sitzen und sich einzubringen. Wie das Strategiepapier „For a meaningful artificial intelligence“ feststellt, werden führende Länder, auf dem Gebiet der künstlichen Intelligenz, nicht nur einen großen Teil der Profite einstreichen, sie werden die Systeme und Technologien auch kontrollieren.<sup>10</sup>

### 2.3.3 Verständnis

Unter dem Begriff Künstliche Intelligenz wird wie bereits angerissen besonders im Laiengebrauch vieles verstanden, von *General AI* – Technologie die gleichgesetzt wird mit menschlicher Intelligenz – bis zu *Narrow AI* – Technologie, die bestimmte Aufgaben ausführt. Zweitere ist die Form von Künstlicher Intelligenz, mit der wir es heute häufig zu tun haben. General AI ist hingegen vorrangig aus Science-Fiction-Erzählungen bekannt – und in diesen nicht besonders akkurat dargestellt.

Der Historiker Noah Yuval Harari vertritt die These, dass unsere Vorstellung davon, wie Roboter und Künstliche Intelligenz unsere Gesellschaft verändern werden, völlig fehlgeleitet ist. Die populäre Darstellung von Künstlicher Intelligenz in Science Fiction Filmen und Literatur, in denen Maschinen neben Intelligenz auch ein Bewusstsein zu haben scheinen, führt zu einem großen Missverständnis. *„Intelligenz und Bewusstsein gleichzusetzen, verzerrt die gesamte Diskussion [...] Intelligenz ist die Fähigkeit, Probleme zu lösen; Bewusstsein die Fähigkeit, etwas zu fühlen. Bei Menschen gehört beides zusammen. Maschinen lösen Probleme auf eine ganz andere Art.“*<sup>11</sup>

Die amerikanische Wissenschaftlerin und Schriftstellerin Zeynep Tufekci spricht davon, dass wir in Hinblick auf Künstliche Intelligenz „realistische Alpträume“ brauchen.<sup>12</sup> Sie bezieht sich damit darauf, dass viele der klassischen KI-Szenarien auf dystopische Erzählungen wie 1984 oder Terminator beziehen – kurz: Narrativen über KI, die auf unrealistischen und überzogenen

---

<sup>10</sup> [https://www.aiforhumanity.fr/pdfs/MissionVillani\\_Report\\_ENG-VF.pdf](https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf)

<sup>11</sup> <https://join-ada.com/20180701-noah-yuval-harari-ki-revolution>

<sup>12</sup> <https://twitter.com/zeynep/status/1012357341981888512>

Szenarien fußen. Diese Szenarien werden unhinterfragt übernommen und auf die heutige Welt übertragen, was dazu führt, dass man realitätsferne Risiken und Auswirkungen der Technologien diskutiert und Unsicherheit schürt.

Grundsätzlich ist zu beobachten, dass es im nicht-akademischen Bereich wenig allgemein verständliche Visualisierungen und Erläuterungen dazu gibt, wie Künstliche Intelligenz funktioniert. Ein Beispiel für eine gelungene Visualisierung und Aufbereitung des Themas Maschinelles Lernen ist das Projekt r2d3<sup>13</sup> von Stephanie Yee und Tony Chu.

## Sprache

Eine weitere Auffälligkeit im Bereich von KI ist, dass ein Großteil der akademischen Aufsätze und Forschung zum Thema Künstlicher Intelligenz in englischer Sprache verfasst ist. Ein eher geringer Prozentsatz der wissenschaftlichen Aufsätze und Artikel zu den Themen ist in deutscher Sprache zugänglich. Das führt dazu, dass englische Begriffe dominieren und sich auf einen unterschiedlichen Kontext beziehen bzw. sich aus diesem speisen. Im Deutschen wird häufig auf Anglizismen zurückgegriffen, weil die richtigen Begriffe und Worte fehlen oder zu wenig geläufig sind. Ein Beispiel: Ein im Englischen häufig verwendetes Wort im Zusammenhang mit maschinellem Lernen ist das Wort „bias“ – häufig wird dieses auch im Deutschen verwendet. Eine Übersetzung können die Begriffe „Voreingenommenheit, technische Befangenheit, Verzerrung“ sein. Ein Mangel an Sprache, um neue technologische Entwicklungen zu beschreiben, kann sich negativ auf das Verständnis eben dieser Technologien auswirken. Auch die Datensätze, die verwendet werden um Systeme zu trainieren, sind zu vorrangig in englischer Sprache vorhanden. Das führt dazu, dass alle Schlüsse und Muster, die aus diesen Datensets gezogen werden, nur in einem ebenfalls englischen Kontext „funktionieren“. Für eine erfolgreiche Positionierung im Markt ist es also notwendig, dass ebenfalls kontextspezifische deutsche Datensets zum Training zur Verfügung stehen.

### 2.3.4 Mangel an ExpertInnen

Große Technologieunternehmen und Risikokapitalgeber haben Dutzende Milliarden Dollar in den Aufbau von KI-Technologien investiert. Die Technologie befindet sich dennoch in einem sehr frühen Stadium der Entwicklung. Das Talent um Intelligente Systeme zu bauen ist nach wie vor knapp und teuer. Eine Analyse von Tencent ergab, dass es weltweit nur 300.000 IngenieurInnen mit den für die KI-Entwicklung erforderlichen Fähigkeiten gibt. Das Kompilieren und Bereinigen der riesigen Datenmengen, sowie der Aufbau und die Wartung der Infrastruktur für KI-Systeme, ist zudem kostspielig und zeitaufwendig.<sup>14</sup>

---

<sup>13</sup> <http://www.r2d3.us/visual-intro-to-machine-learning-part-1/>

<sup>14</sup> <https://www.acumos.org/blog/2018/03/07/what-is-acumos/>



## 2.3.5 Weitere Herausforderungen

### **Unsichtbare Arbeit**

Einige Systeme wirken auf den ersten Blick vollautomatisch und autonom, im Hintergrund sind aber oftmals Menschen involviert. Sie assistieren und helfen dabei, Fehler auszugleichen und Entscheidungen zu treffen. Services wie Facebook, Youtube und weitere setzen Menschen aus Niedriglohnländern wie Indien oder den Philippinen ein, um Inhalte zu moderieren und zu kuratieren.<sup>15</sup> Mithilfe deren Arbeit werden Algorithmen trainiert. Für den/die NutzerIn sind diese Arbeitskräfte unsichtbar. Von außen scheint alles voll-automatisiert und reibungslos. In Wahrheit können viele KI-Systeme ohne menschliche Arbeit und Moderation im Hintergrund gar nicht funktionieren. Trotz des häufigen Einsatzes von Menschen zur Unterstützung von KI-Systemen ist das Thema wenig bekannt und wird von den Konzernen so gut wie nicht adressiert. Oftmals arbeiten die sogenannten Content-ModeratorInnen unter prekären Bedingungen, es gibt kaum Trainings und psychologische Unterstützung für die häufig psychisch sehr fordernde Arbeit. Es gibt wenig Transparenz zu Vorgaben und Richtlinien für die Moderation sowie den Arbeitsbedingungen. Häufig sind Content-ModeratorInnen nicht angestellt, sondern AuftragnehmerInnen, die über Firmen wie Amazon Mechanical Turk oder Clickworker vermittelt werden.<sup>16</sup>

### **Verantwortung und Rechenschaftspflicht**

Eine große Herausforderung im Outsourcing von Entscheidungen betrifft Fragen rund um Verantwortung und Rechenschaft – sprich: Wer ist verantwortlich für Entscheidungen der Technologie und deren (negative) Auswirkungen? Es gibt Ansätze, in denen erforscht wird, wie Maschinen zur Verantwortung gezogen werden können, diese stecken allerdings noch in den Kinderschuhen.

Im Fall von Deep Learning, einer speziellen Art des Maschinellen Lernens, in dem es mehrere Ebenen von einwirkenden Algorithmen gibt, ist es noch schwieriger nachzuverfolgen, an welchem Punkt der Verarbeitungskette Fehler passieren und wer dafür zur Rechenschaft gezogen werden kann.

## 2.4 Chancen

Wie bei allen Technologien liegen Chancen und Risiken für Künstliche Intelligenz nah beieinander. So können neue digitale Werkzeuge beispielsweise dafür eingesetzt werden, um soziale

---

<sup>15</sup> <https://variety.com/2018/film/reviews/the-cleaners-review-1202673932/>

<sup>16</sup> <https://hbr.org/2017/01/the-humans-working-behind-the-ai-curtain>

Gerechtigkeit zu befördern und auf statistischer Ebene Gleichheit herzustellen, oder im Gegenteil, um soziale Gegensätze weiter verschärfen. Gerade wenn sich eine Technologie noch im Entwicklungsstadium befindet, lassen sich viele Einsatzfelder und Szenarien noch nicht genau vorhersehen. Bei der Betrachtung der Chancen sollte daher unbedingt auch immer hinterfragt werden, welche Risiken und Nebenwirkungen mit der Nutzung der Technologie einhergehen können und an welchen Stellen auf den Einsatz verzichtet werden kann und sollte. Der französische Strategiepapier „For a meaningful artificial intelligence“<sup>17</sup> hebt hervor, dass neue Wege formuliert werden müssen, wie Menschen und intelligente Systeme zusammenarbeiten können. Anstatt individuellen Lebenswege und Wohlfahrtssysteme zu untergraben, sollte die KI in erster Linie dazu beitragen, unsere grundlegenden Menschenrechte zu fördern, die sozialen Beziehungen zu verbessern und die Solidarität zu stärken.

Im Moment gibt es viel politisches Interesse und Investitionen in KI-Systeme, es gibt aber auch einige Bereich mit denen Regierungen und private Akteure zu kämpfen haben: die Ethik dieser so genannten „intelligenten“ Systeme. In den letzten Monaten haben verschiedene Regierungen und große Institutionen Anstrengungen unternommen, um die Ethik von KI-Systemen genauer zu untersuchen, wie zum Beispiel das britische Centre for Data Ethics and Innovation.<sup>18</sup> In den geplanten Maßnahmen der EU Kommission liest man neben der Förderungen bei der Entwicklung der Technologien, auch über sozioökonomische Herausforderungen und darüber, dass die Kommission plant Ende 2018 ethische Richtlinien zu veröffentlichen.<sup>19</sup>

Ein besseres Verständnis dafür, was ethische, verantwortungsvolle und sinnvolle Entwicklung von künstlicher Intelligenz in der Praxis bedeutet, kann dabei helfen sicherzustellen, dass Investitionen in neue Technologien der Gesellschaft als Ganzes, einschließlich historisch marginalisierter Gruppen, zugute kommt. Mit der Gründung der Ethik-Kommission Automatisiertes und Vernetztes Fahren und ihrem im Juni 2017 veröffentlichten Bericht<sup>20</sup>, der verschiedene Empfehlungen für die ethische Umsetzung selbstfahrender Fahrzeuge enthält, wie z.B. die Förderung der Datensouveränität, hat die Bundesregierung bei Ethikfragen bereits eine Vorreiterrolle übernommen.

### **Bereiche in denen positive Entwicklungen beobachtet werden können:**

**Barrierefreiheit:** Für Menschen mit Einschränkungen bietet Software zur Sprach- und Bilderkennung große Vorteile. Als Beispiel kann hier Echtzeit-Untertitelung von Medien für gehörlose Menschen gelten.<sup>21</sup>

---

<sup>17</sup> [https://www.aiforhumanity.fr/pdfs/MissionVillani\\_Report\\_ENG-VF.pdf](https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf)

<sup>18</sup> <https://www.gov.uk/government/consultations/consultation-on-the-centre-for-data-ethics-and-innovation>

<sup>19</sup> [http://europa.eu/rapid/press-release\\_IP-18-3362\\_en.htm](http://europa.eu/rapid/press-release_IP-18-3362_en.htm)

<sup>20</sup> <https://www.bmvi.de/SharedDocs/DE/Pressemitteilungen/2017/084-dobrindt-bericht-der-ethik-kommission.html>

<sup>21</sup> <https://blogs.microsoft.com/ai/ai-powered-captioning/>

**Öffentliche Verwaltung und Dienstleistungen:** Intelligente Systeme können als Frühwarnsysteme für Naturkatastrophen dienen. Mit Hilfe von Künstlicher Intelligenz können beispielsweise Epidemien simuliert, und damit Menschenleben gerettet werden.<sup>22</sup>

**Verkehr:** Fortschritt durch den Einsatz intelligenter Systeme im Bereich Transport und Verkehr bezieht sich häufig auf die Themen Sicherheit, Effizienz und Komfort. Sensoren und Systeme sollen dabei helfen potenzielle Gefahren zu erkennen und gefährliche Situationen und Unfälle zu vermeiden. Verkehrsstaus sollen durch die Koordination von Fahrzeugen vermieden werden.<sup>23 24</sup>

**Medizin:** Im Bereich der Medizin werden KI-Systeme bei in der Früherkennung von Krankheiten und als Unterstützung bei Diagnosen eingesetzt. Bilderkennungsverfahren können beispielsweise dabei helfen Krankheiten wie Alzheimer frühzeitig zu erkennen.<sup>25</sup>

**Landwirtschaft:** Durch den Einsatz von KI-Technologien kann die Nutzung von Pestiziden und Düngern reduziert und die Produktivität, sowie die Qualität der Erzeugnisse gesteigert werden.<sup>26</sup>

**Sprache und Linguistik:** Künstliche Intelligenz kann die sprachübergreifende Zusammenarbeit erleichtern, wie zum Beispiel bei der Übersetzung von Texten oder der Spracherkennung. Für das Training der Systeme werden große Mengen übersetzter Dokumente und Texte benötigt.

**Verarbeitung großer Datenmengen** – von Satellitenbildern bis hin zu Daten über die Umwelt, wie sie beispielsweise im Rahmen des Projekts Tara Oceans<sup>27</sup> gesammelt werden, kann KI dabei helfen große Datenmengen zu analysieren und zu kategorisieren.

**Ressourcenintensive Vorgänge:** Momentan wird KI-Technologie häufig in Bereichen eingesetzt, die viele menschliche Ressourcen binden, wie zum Beispiel der Digitalisierung und Auswertung von Text.

## 2.5 Kontroversen

Ein durchaus kontroverses Einsatzfeld von KI ist die Strafverfolgungsarbeit. Polizeieinheiten setzen sogenannte „Predictive-Policing“-Algorithmen ein, um bessere Vorhersagen über das Auftreten von Straftaten treffen zu können, z.B. um besonders „gefährliche“ Gegenden identifizieren zu können.

---

<sup>22</sup> [ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=51610](http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51610)

<sup>23</sup> <http://www.sueddeutsche.de/auto/autonomes-fahren-roboter-an-bord-1.2323179>

<sup>24</sup> 3TH1CS, Philipp Otto und Eike Gräf, 2017

<sup>25</sup> <https://arxiv.org/abs/1709.02369v1>

<sup>26</sup> [ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=51610](http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51610)

<sup>27</sup> <https://oceans.taraexpeditions.org/en/>

Mittels Künstlicher Intelligenz werden zudem Bilder von Überwachungskameras ausgewertet, um kriminelle Handlungen auf den Bildern zu erkennen.<sup>28</sup> Maschinelles Lernen wird ebenfalls in der Risikobewertung von Straftätern angewandt. Eine von ProPublica veröffentlichte Studie zeigt, dass die in den USA verwendeten Systeme hierbei schwarze Amerikaner benachteiligen und diskriminieren.<sup>29</sup> Oft werden vorhandene Vorurteile von den Algorithmen weiter verstärkt. Es ist häufig intransparent, mit welchen Daten die Algorithmen trainiert werden.

## 2.6 Fragestellungen, die sich aus den Herausforderungen und Chancen ergeben

- Wie lassen sich Tendenzen, Vorurteile und Verzerrungen in Datensammlungen und Modellen identifizieren und eliminieren bzw. abschwächen?
- Was sind potenzielle Auswirkungen – Chancen und Risiken – von Künstlicher Intelligenz auf die Gesellschaft?
  - Beispiel: AI Pattern Language<sup>30</sup>
- Welche Stakeholder müssen an der Gestaltung von neuen Technologien beteiligt sein, um sicherzustellen, dass sie vielen Menschen zugutekommen?
  - Wie kann in diesem Kontext inhaltliche Expertise zu bestimmten Sachverhalten mit technischer Expertise vereint werden?
- Ethik in der Künstlichen Intelligenz: Wie können ethische Überlegungen in intelligenten Systemen verankert werden?
- Wie kann die Entwicklung neuer Open-Source-KI-Modelle vorangetrieben werden?
  - Beispiel: Model Zoo<sup>31</sup> Modelle für Maschinelles Lernen zur freien Nutzung
- Wie kommt man an unter einer offenen/freien Lizenz stehende Trainingsdaten für Maschinelles Lernen?
  - Beispiel: Das Projekt Common Voice<sup>32</sup> von Mozilla ist ein Projekt, das Daten für Sprachsysteme sammelt und zur freien Nutzung zur Verfügung stellt.
- Welche Rolle spielt Privatsphäre im Kontext von Künstlicher Intelligenz? Wie kann man Datensätze anonymisieren?
  - Beispiel: KI-Protect<sup>33</sup> und AirCloak<sup>34</sup>

---

<sup>28</sup> <http://www.spiegel.de/netzwelt/netzpolitik/mannheimer-weg-2-0-pilotprojekt-mit-intelligenten-kameras-startet-bald-a-1193622.html>

<sup>29</sup> <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

<sup>30</sup> <https://autonomy.datasociety.net/patternlanguage/>

<sup>31</sup> <https://modelzoo.co/>

<sup>32</sup> <https://voice.mozilla.org/>

<sup>33</sup> <https://kiprotect.com/>

<sup>34</sup> <https://aircloak.com>

## 3. Vorschlag des Themenschwerpunkts

### 3.1 Beschreibung

Aus der Analyse geht hervor: es gibt noch einiges zu entwickeln und zu erforschen, bis die Gesellschaft und Wirtschaft das volle Potenzial von intelligenten Systemen ausschöpfen kann. Für den Themenschwerpunkt der 5. Runde des Prototype Fund schlagen wir daher vor, die gesellschaftlichen Auswirkungen und Chancen von neuen Technologien zu erforschen und die Systeme, im Zeichen der Wissenschaftskommunikation, besser zu erklären und Außenstehenden zugänglich zu machen. Wir schlagen vor, monetäre und wissenschaftliche Ressourcen darauf zu verwenden, Künstliche Intelligenz und neue Technologien hinsichtlich ihrer gesellschaftlichen Potenziale zu untersuchen und die Entwicklung von Anwendungen für das Gemeinwohl zu fördern. Ziel ist auch die Relevanz des Themas „intelligente Systeme“ hervorzuheben und mehr Menschen dazu zu motivieren, die Potenziale zu erforschen.

Der Schwerpunkt baut auch gut auf vorangegangene Themenschwerpunkte auf, Diversity, als auch Privatsphäre, spielen beide eine wichtige Rolle in der Entwicklung von KI-Systemen.

### 3.2 Themenvorschlag

Maschinen lernen lassen: Technologien für die Zukunft– Making machine learning, AI and emerging technologies that work for us.

**Building the tech we want:** Wie können neue Technologien uns dabei helfen, existierende Ungerechtigkeit zu adressieren (und zu reduzieren), statt diese zu verstärken?  
Welche Rolle kann die Entwicklung von Open-Source-Infrastruktur spielen?

**Maschinelles Lernen für das Gemeinwohl:** Welche gesellschaftlichen Themen können mit Hilfe neuen Technologien besser erschlossen und bearbeitet werden, und wie?

**Die schwarze Box ausleuchten:** Neue Technologien erklären und verstehen: Wie funktioniert Maschinelles Lernen oder Künstliche Intelligenz? Was sind Gefahren, Mythen und Chancen?

Bestehende Beispiele:

- R2D3 „Visual Intro to Machine Learning“<sup>35</sup>,
- Distill.pub, eine Zeitschrift, die sich dem Erklärungen des Maschinellen Lernens widmet<sup>36</sup>
- Explorable Explanations<sup>37</sup>,

---

<sup>35</sup> <http://www.r2d3.us/visual-intro-to-machine-learning-part-1/>

<sup>36</sup> <https://distill.pub/about/>

<sup>37</sup> <https://explorabl.es/>

- Deep Dream Generator, visual content creation with AI enabled tools<sup>38</sup>.

Anwendungsbeispiele:

- Face Privacy via Pixelation<sup>39</sup>,
- Algorithm Observatory<sup>40</sup>,
- DoNotPay, a artificial-intelligence lawyer chatbot<sup>41</sup>,
- Clear My Record<sup>42</sup>,
- Open Schufa<sup>43</sup> von Algorithm Watch,
- Cognimates<sup>44</sup>,
- Machine learning is being used to uncover the mass graves of Mexico's missing<sup>45</sup>, Human Rights Data Analysis Group,
- Generative models for medical imaging<sup>46</sup>,
- Feminist Dataset<sup>47</sup>,
- KI Protect<sup>48</sup>,
- Uwazi<sup>49</sup>,
- VFrame<sup>50</sup>.

### 3.3 Chancen des Schwerpunktes

Trotz der oben genannten Herausforderungen bietet das Thema KI innerhalb des Wirkungsbezugs des Prototype Funds eine Reihe von Möglichkeiten:

- EntwicklerInnen und ForscherInnen in Deutschland wird die Möglichkeit geboten, ihre Fähigkeiten im Bereich des Maschinellen Lernens und der Künstlichen Intelligenz zu verbessern – derzeit besteht eine reale Kapazitätslücke, wobei Länder wie beispielsweise China und Kanada stark in die Entwicklung in diesem Bereich investieren.

---

<sup>38</sup> <https://deepdreamgenerator.com/>

<sup>39</sup> <https://marketplace.acumos.org/#/marketSolutions?solutionId=b0da3778-fc9f-4453-8bc4-78a09ae43d4c>

<sup>40</sup> <https://algoritmi.pybossa.com/>

<sup>41</sup> <https://www.theguardian.com/technology/2016/jun/28/chatbot-ai-lawyer-donotpay-parking-tickets-london-new-york>

<sup>42</sup> <https://www.clearmyrecord.org/>

<sup>43</sup> <https://algorithmwatch.org/en/openschufa-shedding-light-on-germanys-opaque-credit-scoring/>

<sup>44</sup> <https://www.media.mit.edu/projects/cognimates/overview/>

<sup>45</sup> <https://qz.com/958375/machine-learning-is-being-used-to-uncover-the-mass-graves-of-mexicos-missing/>

<sup>46</sup> [https://arxiv.org/pdf/1806.05034.pdf@DeepMind\\_health](https://arxiv.org/pdf/1806.05034.pdf@DeepMind_health)

<sup>47</sup> <http://soho20gallery.com/caroline-sinders-feminist-data-set/>

<sup>48</sup> <https://kiprotect.com/>

<sup>49</sup> <https://prototypefund.de/project/uwazim/>

<sup>50</sup> <https://prototypefund.de/project/vframe/>

- Schaffung von Kooperationsmöglichkeiten zwischen denjenigen mit technischen Fähigkeiten und jenen mit kontextuell oder sozial einordnender Kompetenz, um eine angemessene und verantwortungsvolle Nutzung von Künstlicher Intelligenz und Maschinellem Lernen zu entwickeln, die deren Nutzung als soziotechnisches System und nicht nur als technisches System widerspiegelt.
- Wiederverwendung von Open-Source-Modellen im Dienste des öffentlichen Interesses und des Gemeinwohls – die Community rund um das Maschinelle Lernen hat sich relativ klar dazu geäußert, dass sie das Maschinelle Lernen offen halten will. Tausende ForscherInnen haben beispielsweise eine neue Zeitschrift zum Maschinellen Lernen boykottiert, weil diese nur als kostenpflichtiges Abonnement geplant war.

### 3.4 Trends

Welche anderen Förderprogramme fokussieren sich auf die Unterstützung von Open-Source-Entwicklungen im Bereich Künstliche Intelligenz und Machine Learning?

- Die Mozilla Foundation hat Anfang Juni einen Fund zur Erforschung von Künstlicher Intelligenz gestartet.<sup>51</sup>
- Die Linux Foundation startet ein Projekt für Open-Source-KI.<sup>52</sup>
- ACUMOS AI Challenge.<sup>53</sup>
- ThoughtWorks Residency zum Thema „Implikationen von künstlicher Intelligenz“.<sup>54</sup>

## 4. Empfehlungen für die Umsetzung

Angesichts der Komplexität dieses Themas geben wir einige Empfehlungen, wie es erklärt werden könnte:

**Co-Kreation:** Wer soll an der Gestaltung der Zukunft, beteiligt sein? Vielfältige Erfahrungen und Fachkenntnisse sind entscheidend, um sicherzustellen, dass neue Technologien den sozialen Kontext der Menschen widerspiegeln. Zum Beispiel die Paarungen: Data Scientist + Community OrganizerIn; Software-EntwicklerIn + AktivistIn. Wir empfehlen, diese Arten von Zusammenarbeit in der Ausschreibung explizit anzugeben und darauf hinzuweisen, dass transdisziplinäre Partnerschaften zu besseren Tools und Produkten führen können.

---

<sup>51</sup> <https://blog.mozilla.org/blog/2018/06/04/mozilla-announces-225000-for-art-and-advocacy-exploring-artificial-intelligence/>

<sup>52</sup> <https://www.golem.de/news/deep-learning-linux-foundation-startet-projekt-fuer-open-source-ki-1803-133549.html>

<sup>53</sup> <https://www.linuxfoundation.org/blog/compete-in-the-acumos-ai-challenge-for-a-chance-to-win-50000/>

<sup>54</sup> <http://www.newmediacaucus.org/thoughtworks-arts-residency-open-call-the-implications-of-ai/>

Künstliche Intelligenz und Maschinelles Lernen (und viele andere neue Technologien) können auf viele verschiedene Probleme angewendet werden – das bedeutet jedoch nicht, dass es immer notwendig ist sie einzusetzen, oder dass nicht eine einfachere Lösung gefunden werden könnte. Es erscheint von Dringlichkeit, schon über die Ausschreibung klarzustellen, dass es in dieser Runde nicht nur um den Versuch geht, eine emergente Technologie einzusetzen, sondern auch um das langfristige öffentliche Interesse und Teilhabe.

Wie oben beschrieben ist ein Mangel an Verständnis dafür, was diese Technologien bieten können und was nicht, ein wachsendes Problem. Mit dem Fokus auf Prototypen und Visualisierungen, die sich auf die Verbesserung des Verständnisses des Betrachters oder Anwenders konzentrieren, könnte diesem Problem begegnet werden.

Obwohl der skizzierte Schwerpunkt einen Fokus auf KI hat, schlagen wir vor, den Themenschwerpunkt breiter zu fassen, um eine Vielzahl von kreativen Anwendungen zuzulassen. Die Beschränkung auf KI birgt die Gefahr, dass Projekte sich bei der Lösung von Problemen auf diese Technologie versteifen, anstatt die am besten geeignete Technologie einzusetzen.

Wir empfehlen, den ExpertInnenpool der Jury um ExpertInnen aus den Bereichen Künstliche Intelligenz und Maschinelles Lernen zu erweitern.

## 5. Ziele

### 5.1 Interne Ziele

- Gezielt neue Entwicklungen fördern und Beziehungen zu neuen Akteuren aufbauen,
- Trends und Themen im Bereich Open-Source-Entwicklung verstehen und durchdringen,
- Netzwerke zu anderen Förderprogrammen aufbauen und weitere Stakeholder dazu anregen, in den Bereich Open Source und Civic Tech zu investieren.

### 5.2 Externe Ziele

- Ein Thema ausrufen, das zu vielfältigen Bewerbungen anregt und potenzielle BewerberInnen aktiviert;
- Die EntwicklerInnen in Deutschland sollen durch die Themenwahl dazu ermutigt werden, kritisch über Themen nachzudenken, die für die Technik/Gesellschaft in ihrem Land wichtig sind;
- Einen Beitrag zur Vermittlung komplexer technischer Themen leisten;
- Entwicklung von Begrifflichkeiten und einer klaren Sprache fördern, um technische Zusammenhänge zu erklären;
- Die Erforschung des gesellschaftlichen Nutzens und Mehrwerts von neuen Technologien; die Erforschung von Chancen und Risiken neuer Technologien.