

September 2018 · Dr. Stefan Heumann mit Nicolas Zahn

---

# Erfolgsmessung von KI-Strategien

Mit Indikatoren und Benchmarks  
die Umsetzung der Strategie  
erfolgreich steuern



Think Tank für die Gesellschaft im technologischen Wandel

## Executive Summary

Künstliche Intelligenz (KI) ist die Schlüsseltechnologie der Digitalisierung. Denn ohne automatisierte, intelligente Analyse- und Entscheidungssysteme kann die Fülle an im Zuge der digitalen Transformation von Wirtschaft und Gesellschaft entstehenden Daten nicht produktiv genutzt werden. Hinzu kommt das große Potenzial der Technologie für die Lösung wichtiger gesellschaftlicher Herausforderung von der Verbesserung der Medizin bis hin zur intelligenteren Steuerung von Verkehrs- oder Energiesystemen. Mittlerweile hat auch die Politik die Bedeutung von KI erkannt. Viele Länder haben nationale KI-Strategien veröffentlicht und fördern Entwicklung und Einsatz von KI-Technologien. Auch die Bundesregierung befindet sich mittlerweile in einem Strategieprozess. Die Ergebnisse sollen Anfang Dezember auf dem Digitalgipfel präsentiert werden. Während die im Juli veröffentlichten Eckpunkte der KI-Strategie erkennen lassen, wie ernst man das Thema mittlerweile im Kanzleramt und den Ministerien nimmt, wird ein wichtiger Punkt in der Debatte bisher zu wenig berücksichtigt. Die KI-Strategie folgt auf eine ganze Reihe ambitionierter digitalpolitischer Strategien, wie beispielsweise der Digitalen Agenda. Diese Strategien sind allgemein formuliert und enthalten nur vage Zielsetzungen. Das ist eine große Schwäche dieser bisherigen Strategien, denn ohne ausgewiesene Ziele konnte es nicht gelingen, aus den Strategien schnell konkrete Maßnahmen zu entwickeln und diese auch umzusetzen.

Wenn die Bemühungen der Bundesregierung diesmal nicht das gleiche Schicksal erleiden sollen, muss die KI-Strategie klare Zielsetzungen, Handlungsmaßnahmen und Erfolgskriterien enthalten. Dadurch würde zugleich ein weiteres Manko der KI-Strategie adressiert werden. Die Festlegung klarer Ziele, zum Beispiel in Bezug auf die Zahl der in Deutschland zukünftig auszubildenden KI-Expert:innen, setzt eine tiefgehende Auseinandersetzung mit dem KI-Begriff voraus. Es müsste festgelegt werden, welche Kompetenzen als KI-relevante Expertise gelten und welche nicht. Solche Benchmarks festzulegen, erfordert aber zunächst die tiefgehende Auseinandersetzung mit dem KI-Begriff. Eine solche Auseinandersetzung fehlt im politischen Diskurs und in den bereits veröffentlichten Eckpunkten der KI-Strategie der Bundesregierung.

Im nächsten Schritt ist bei der Entwicklung von konkreten Zielen die Unterscheidung zwischen Input- und Output-Indikatoren wichtig. Input-Indikatoren beinhalten alle Maßnahmen, die zur Zielerreichung eingesetzt werden. Output-Indikatoren ermöglichen die Überprüfung der Zielerreichung. Sie können dabei auf unterschiedliche Dimensionen der Strategie abzielen, wie zum Bei-



spiel die Steigerung der Zahl von KI-Forscher:innen oder die Erhöhung der Präsenz deutscher KI-Forschung auf den führenden internationalen Konferenzen.

Viele Staaten haben bereits KI-Strategien vorgelegt. Eine Analyse der Strategien Chinas, der Vereinigten Staaten von Amerika, Großbritanniens, Frankreichs, Finnlands und der europäischen Kommission zeigt, dass viele qualitative Input-Indikatoren enthalten sind. Die Maßnahmen sind allerdings in Regel recht abstrakt und vage gehalten und es fehlt generell an quantitativen Input-Indikatoren, wie zum Beispiel die Höhe der Haushaltsmittel, die zur Umsetzung der Maßnahmen veranschlagt werden. Noch weniger ist in den Strategien bezüglich Output-Indikatoren zu finden. Ausnahmen sind die Strategien Chinas und Großbritanniens. China macht klare Vorgaben, welchen Anteil KI an der zukünftigen Wertschöpfung in der Wirtschaft einnehmen soll. Großbritannien definiert konkrete Ziele bezüglich der Anzahl an auszubildenden KI-Wissenschaftler:innen und der Anzahl ausländischer KI-Expert:innen, die ins Land geholt werden sollen. Finnland erkennt die Bedeutung von Indikatoren zur Fortschrittsmessung zumindest an und will diese im Rahmen der Umsetzung der Strategie entwickeln. Trotz einiger guter Ansätze fällt das Ergebnis jedoch bescheiden aus. Keine der untersuchten Strategien setzt sich wirklich substantiell und systematisch mit der Ableitung von Input- und Output-Indikatoren aus den übergeordneten strategischen Zielen auseinander. Hier besteht die große Chance für Deutschland sich positiv abzusetzen und die Umsetzung der eigenen Strategie mit Hilfe von Input- und Output-Indikatoren voranzutreiben.

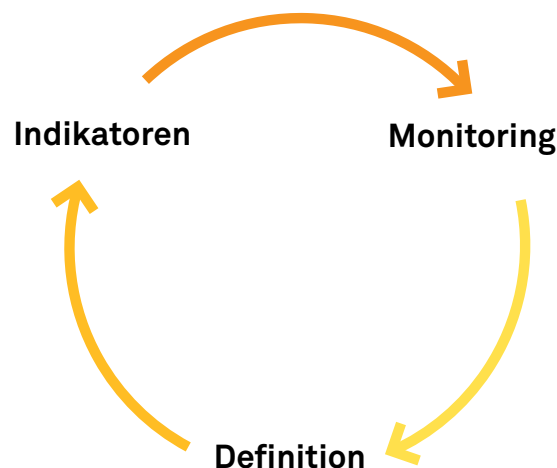
Die Entwicklung eines nationalen KI-Index mit zu messenden Output-Indikatoren müsste nicht bei Null beginnen. Die Analyse bestehender Indizes liefert wichtige Hinweise für die Entwicklung. Die hier untersuchten Indizes kommen vor allem von Unternehmensberatungen und lassen sich nach ihrem methodischen Ansatz in vier Kategorien mit jeweils eigenen Stärken und Schwächen einordnen. Die Umfrage eignet sich als Methode besonders gut, um neue Trends zu erfassen. Allerdings hängt die Aussagekraft von Umfragen stark mit der Auswahl der Befragten und der angewandten Befragungsmethodik zusammen. Eine zusammenfassende Auswertung bestehender Indizes ist zwar relativ leicht operativ umzusetzen; eigene und vor allem neue Fragestellungen lassen sich mit diesem Ansatz aber nur schwer bearbeiten. Bei Daten-getriebenen Ansätzen wird der Index an der Verfügbarkeit bestehender Datenquellen ausgerichtet. Je nachdem, wie aussagekräftig diese Datenquellen sind, steigt und fällt damit auch die Aussagekraft Daten-getriebener Indizes. Bei Indikatoren-getriebenen Ansätzen wird der Index aus einem theoretischen Verständnis des Untersuchungsgegenstands heraus entwickelt. Hier stellt sich aber häufig das Problem, dass es für die



entwickelten Indikatoren keine leicht erschließbaren Datenquellen gibt.

Grundsätzlich weisen die untersuchten KI-Indizes alle erhebliche methodische Schwächen auf. Das liegt zum einem an den genannten Nachteilen der jeweiligen Methodik. Es gibt nicht den einen perfekten Ansatz. Beim Umgang mit methodischen Problemen steht in der Regel nicht die beste Lösung, sondern Ressourceneffizienz im Mittelpunkt. Zum anderen gehen vor allem die Unternehmensberatungen nicht offen und ausführlich genug mit ihren methodischen Problemen um und bleiben somit hinter grundlegenden wissenschaftlichen Standards zurück. Die Politik kann daher nicht einfach bestehende Ansätze integrieren, sondern sollte die Beschäftigung mit den Stärken und Schwächen bestehender Indizes zur Grundlage der Entwicklung eines eigenen Index machen. Ein eigener Ansatz sollte dabei auch ein KI-Technologie-Monitoring integrieren, um den eigenen Index ständig in Bezug auf neueste Trends und Entwicklungen anpassen zu können.

Angesichts der Dynamik der KI-Entwicklung sollte der nationale KI-Index nicht statisch, sondern Prozess-orientiert angelegt werden. Dieser Prozess setzt sich aus drei Säulen zusammen, die sich ständig gegenseitig befruchten.



Grafik 4: Empirische Fundierung der KI-Strategie. Quelle: Stiftung Neue Verantwortung

Auf Basis einer tiefgehenden Auseinandersetzung mit dem KI-Begriff werden Output-Indikatoren abgeleitet, mit denen die zentralen Dimensionen des KI-Ökosystems erfasst werden. Neben Forschung, Bildung und Wirtschaft müssen auch Output-Indikatoren für die Bereiche Daten, Infrastruktur, Gesellschaft und Internationales definiert werden.



niert werden. Das KI-Trend-Monitoring bringt die Dynamik in den Prozess. Mit Hilfe des Monitorings können neue Trends identifiziert und bei der Weiterentwicklung der Indikatoren berücksichtigt werden.

Angesichts der zu erwartenden Höhe der öffentlichen Ausgaben für die Umsetzung der KI-Strategie sind die für den Aufbau eines Monitoring und Benchmarking benötigten Haushaltsmittel vergleichsweise gering. Es wäre aber sehr gut angelegtes Geld. Denn nur wenn die Entwicklung eines KI-Ökosystems in Deutschland mit Indikatoren erfasst wird, gibt es Anhaltspunkte zur Bewertung der Effektivität der im Rahmen der Strategie entwickelten Maßnahmen. Die Bundesregierung sollte für den Aufbau eines solchen Mess-Ansatzes ein unabhängiges Forschungsinstitut beauftragen. Auf Basis wissenschaftlicher Standards und hoher Methodenkompetenz würde das beauftragte Forschungsinstitut vor allem die Output-Indikatoren entwickeln und erfassen und den KI-Trend Report erstellen. Das Forschungsinstitut sollte dabei völlig frei und unabhängig von politischen Einflüssen agieren können. Und es sollte seine Arbeit regelmäßig veröffentlichen. Reports mit den Ergebnissen der Messungen der Indikatoren hinsichtlich der Entwicklung des deutschen KI-Ökosystems und zu globalen KI-Trends würden die gesellschaftspolitische Diskussion über KI in Deutschland auf eine empirisch unterlegte Grundlage stellen. Zusätzlich würden die Reports für die Bundesregierung eine wichtige Ressource für die erfolgreiche Umsetzung und kontinuierliche Weiterentwicklung der nationalen KI-Strategie darstellen.



## Inhalt

Einleitung	7
Erfolgsindikatoren in nationalen KI-Strategien	9
China	10
USA	13
Großbritannien	14
Frankreich	16
Finnland	17
Europäische Union (EU)	18
Konkrete Ziele und Indikatoren in den nationalen KI-Strategien	19
KI-Indizes	21
Aufwand vs. Aussagekraft	32
Empirische Fundierung der deutschen KI-Strategie	33
Input-Indikatoren	35
Output-Indikatoren	40
KI-Technologie- und Trend-Monitoring	42
Ausblick	43

## Einleitung

Im Juli hat die Bundesregierung ihre Eckpunkte für eine Strategie für Künstliche Intelligenz (KI) veröffentlicht.<sup>1</sup> In den Eckpunkten der Bundesregierung finden sich viele gute Ideen. Die Strategie setzt breit an und greift unter anderem auch unsere Forderung, die Maßnahmen der Bundesregierung auf den Aufbau und Entwicklung eines starken KI-Ökosystems auszurichten, auf.<sup>2</sup> Im Papier der Bundesregierung fehlt allerdings ein ganz zentraler Punkt: Es werden keine messbaren Ziele definiert. Der Mangel an konkreten Zielen und klaren Erfolgsindikatoren steht symptomatisch für viele Strategiepapiere und Ankündigungen in der deutschen Digitalpolitik. In der Regel fehlen eine Definition von klaren Zielen und eine Beschäftigung mit ihrer Überprüfung. So können Politiker:innen und Bürger:innen weder erkennen, was genau erreicht werden soll, noch wann es erreicht sein wird und ob wir Fortschritte machen. Die KI-Strategie der Bundesregierung bietet die Chance, es diesmal besser zu machen. Die Vorteile liegen auf der Hand. Ein Index macht klare, Ressort-übergreifende Ziele notwendig, ermöglicht eine agile Steuerung und schafft damit die Voraussetzung für eine schnelle Umsetzung der Strategie.

Eine intensive Auseinandersetzung mit KI ist angesichts der breiten und unbestimmten Nutzung des Begriffs in der öffentlichen Diskussion dringend erforderlich. Auch in der politischen Diskussion bleibt in der Regel unklar, was genau unter KI verstanden wird. Bei der Festlegung von bestimmten Zielen, zum Beispiel die Zahl der mit KI befassten Lehrstühle an deutschen Hochschulen um den Faktor  $x$  zu erhöhen, muss bedacht werden, wann ein Lehrstuhl als ein KI-Lehrstuhl gilt und wann nicht. Genauso verhält es sich mit der Forschungsförderung oder der Erhöhung der Anzahl der auf KI spezialisierten Startups. Und nicht nur die Beschäftigung mit Zielen führt zwangsläufig zu einer tieferen Auseinandersetzung mit KI. Auch die Frage nach der Messung von Fortschritten und Indikatoren zur Zielerreichung ist wichtig. Sie zwingt uns darüber nachzudenken, was ein starkes KI-Ökosystem überhaupt ist, wodurch es sich auszeichnet und wie wir seinen Zustand und seine Entwicklung empirisch erfassen könnten.

---

<sup>1</sup> Die Bundesregierung (2018). Eckpunkte der Bundesregierung für eine Strategie Künstliche Intelligenz,

[https://www.bmbf.de/files/180718%20Eckpunkte\\_KI-Strategie%20final%20Layout.pdf](https://www.bmbf.de/files/180718%20Eckpunkte_KI-Strategie%20final%20Layout.pdf)

<sup>2</sup> Harhoff, D. et al. (2018). Eckpunkte einer nationalen Strategie für Künstliche Intelligenz,

<https://www.stiftung-nv.de/de/publikation/eckpunkte-einer-nationalen-strategie-fuer-kuenstliche-intelligenz>



Eine solche Auseinandersetzung mit Zielen, Benchmarks und Indikatoren muss immer kritisch sein. Das Erreichen von Benchmarks und Indikatoren sollte niemals Selbstzweck werden. Stattdessen müssen wir ständig hinterfragen, ob die Indikatoren überhaupt das messen, was wir messen wollen und ob mögliche Fehler in unseren Indikatoren, den Datenquellen oder unserer Auswertung möglicherweise das Bild verzerren. Für einige Fragen mag es schwierig bis unmöglich sein, über einfach operationalisierbare Indikatoren Fortschritte und Zielerreichung zu überprüfen. Mängel in den Datenquellen und die Grenzen unserer Analysemethoden müssen verstanden und offen diskutiert werden – eine Diskussion, die wir selbst mit diesem Papier anstoßen wollen. Allerdings sind diese Limitierungen keine Rechtfertigung dafür, eine solche Auseinandersetzung grundsätzlich abzulehnen. Dafür wiegen die Vorteile doch zu schwer. Benchmarks und Indikatoren machen klare Zieldefinition notwendig und damit eine agile, auf Fortschritts- und Erfolgsmessung basierende politische Steuerung möglich.

Wir unterscheiden die Definition von politischen Maßnahmen und Erfolgskriterien als Input- und Output-Indikatoren. Bevor man diese entwickeln kann, müssen die übergeordneten Ziele definiert sein. Ein solches Ziel könnte zum Beispiel die Förderung eines KI-Ökosystems sein. Um daraus Input- und Output-Indikatoren abzuleiten, muss genauer definiert werden, was überhaupt ein KI-Ökosystem ist, was seine unterschiedlichen Dimensionen sind, und wie man diese fördern und ihre Entwicklung messen könnte.

	Input-Indikatoren	Output-Indikatoren
Quantitativ	Höhe der Fördergelder	Anzahl KI-Patente
Qualitativ	Agile Forschungsförderung	KI-Qualitätsstandards

Tabelle 1: Indikatormatrix mit Beispielen, Quelle: Stiftung Neue Verantwortung

Im Kontext einer KI-Strategie sind Input-Indikatoren demzufolge alle politischen Maßnahmen, die zur Stärkung des KI-Ökosystems unternommen werden. Dabei kann zwischen quantitativen und qualitativen Input-Indikatoren unterschieden werden. Quantitative und damit auch recht leicht operationalisierbare Input-Indikatoren sind zum Beispiel die Haushaltsmittel, die für die Forschungsförderung oder neue Investitionsfonds zur Verfügung gestellt werden sollen. Ein qualitativer Input-Indikator wäre eine politische Maßnahme, wie zum Beispiel neue Regelungen für die Vergabe von Forschungsmitteln, die Aufwand und Bürokratie reduzieren und zu mehr Wettbewerb führen sollen. Output-Indikatoren beziehen sich auf die Zielerreichung. Im Kern geht es darum, ob bei den Maßnahmen auch das herauskommt, was man sich zum Ziel gesetzt hat. Aber Output-Indikatoren müssen sich nicht zwangsläufig auf Input-Indikatoren beziehen. Man kann sie auch nutzen, um





unterschiedliche Dimensionen eines KI-Ökosystems zu erfassen – auch solche, die von der Politik nicht direkt adressiert werden, wie beispielsweise die Zahl von Startup-Gründungen. Auch hier kann man wieder zwischen einer quantitativen und einer qualitativen Dimension unterscheiden.

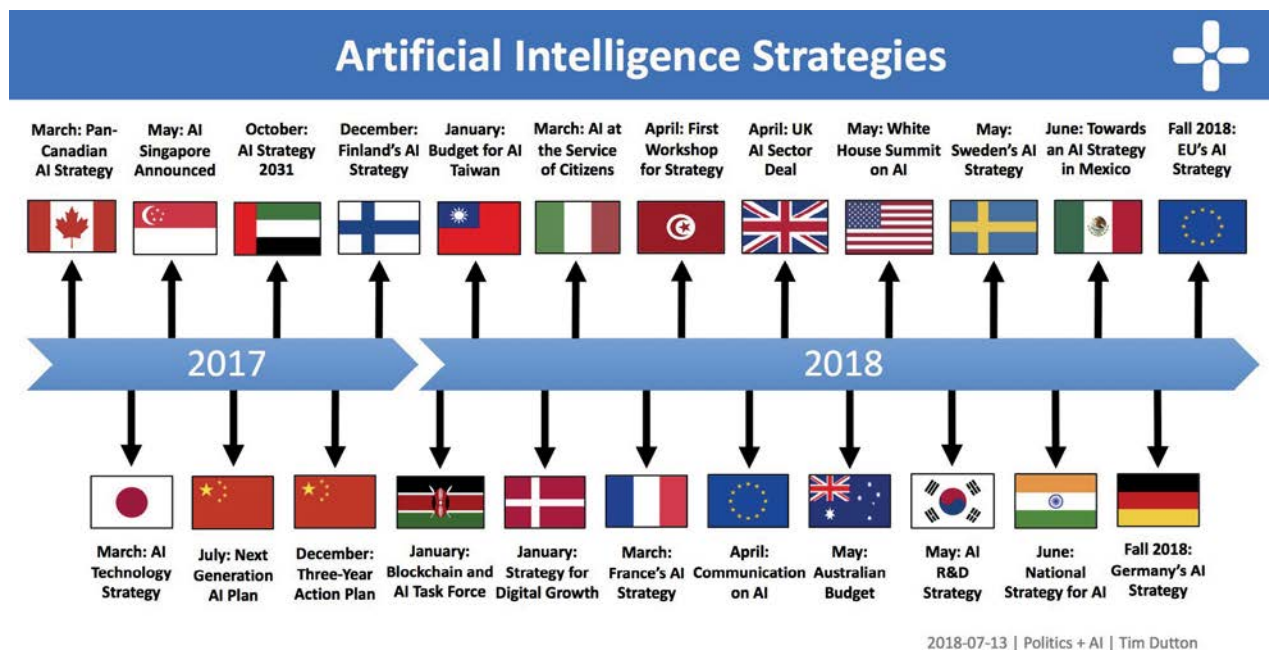
Mit dem hier vorgelegten Papier möchten wir eine Diskussion über In- und Output-Indikatoren nationaler KI-Strategien anstoßen. Um Ideen für die Entwicklung solcher Indikatoren zu generieren, haben wir uns angeschaut, ob und wie bereits veröffentlichte KI-Strategien Ziele definieren und Maßnahmen zur Zielüberprüfung enthalten. In einem weiteren Kapitel haben wir uns mit der Methodik bereits verfügbarer KI-Indizes befasst. In beiden Fällen ging es darum, in der Analyse wichtige Kernfragen herauszuarbeiten und einen ersten Überblick zu schaffen. Hierbei erheben wir keinen Anspruch auf Vollständigkeit. Die nationalen KI-Strategien liefern zwar vereinzelt gute Ansätze und Ideen, lassen aber eine tiefere und systematische Auseinandersetzung mit Indikatoren und Benchmarks vermissen. Um diese Lücke zu schließen, haben wir öffentlich zugänglich KI-Indizes und Benchmark-Reports zusammengetragen.

Der Großteil dieser Indizes und Reports weist erhebliche methodische Schwächen auf. In Medien und Öffentlichkeit werden diese Reports allerdings in der Regel völlig unkritisch rezipiert. Daher ist es uns auch ein Anliegen, eine kritische Diskussion über KI-Reports und Benchmarks anzustoßen. Allerdings wollen wir mit unserer Kritik den Sinn und Nutzen solcher Reports keineswegs in Frage stellen. Deshalb stellen wir in einem weiteren Kapitel auch einige eigene Ideen für die Entwicklung einer empirischen Fundierung der KI-Strategie vor. Unser eigener Ansatz setzt auf ein dynamisches Zusammenspiel zwischen einem KI-Trendmonitoring, das als Grundlage für die kontinuierliche Beschäftigung mit und Weiterentwicklung von Indikatoren dient. Wir hoffen, mit diesem Papier einen Beitrag für die Diskussion zu leisten, wie die Ziele der nationalen KI-Strategie definiert und für eine Überprüfung operationalisiert werden könnten.

## **Erfolgsindikatoren in nationalen KI-Strategien**

Rund um den Globus betrachten politische Entscheider:innen mittlerweile KI als eine zentrale Schlüsseltechnologie. Wie die Grafik weiter unten zeigt, wurden daher in den letzten Jahren in vielen Ländern Strategieprozesse angestoßen, um sich mit den Potenzialen und Risiken dieser Technologie auseinanderzusetzen. Die daraus hervorgegangenen KI-Strategiepapiere dienen als erste empirische Grundlage, um sich mit der Frage zu befassen, wie

Fortschritts- und Erfolgsindikatoren für eine nationale KI-Strategie entwickelt und definiert werden könnten. Für dieses kurze Impulspapier konnten wir keine systematische Auswertung aller bereits veröffentlichten Strategien leisten. Stattdessen haben wir uns einige prominente Beispiele etwas genauer hinsichtlich der Frage angeschaut, inwiefern die Strategien klare Ziele und Indikatoren in Bezug auf ihre Erfüllung definieren. Die in der Einleitung unterschiedenen Dimensionen zwischen Input und Output und quantitativ und qualitativ haben hierfür einen analytischen Rahmen gebildet.



Grafik 1: Quelle: <https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd>

## China - Planwirtschaft 4.0?

Im Juli 2017 hat die chinesische Regierung den “New Generation Artificial Intelligence Development Plan” präsentiert.<sup>3</sup> Aufgrund der dort formulierten globalen Ambitionen hat die chinesische KI-Strategie international viel Aufmerksamkeit erregt. Hinzu kommt, dass China in Bezug auf KI-Entwicklung strategisch bereits gut positioniert ist. Das liegt vor allem an der Größe des chinesischen Markts mit den weltweit meisten Internet-Nutzer:innen und den damit verbundenen riesigen Datenmengen für die KI-Entwicklung.

<sup>3</sup> Die Analyse stützt sich auf die Übersetzung des Originaldokuments wie sie hier verfügbar ist: New America (2017). A Next Generation Artificial Intelligence Development Plan, <https://www.newamerica.org/documents/1959/translation-fulltext-8.1.17.pdf>



Auch chinesische Internet- und Technologiefirmen haben eine rasante Entwicklung hinter sich. Nun überführt China seine Ambitionen auch in offizielle Regierungsziele. Bis 2030 möchte China zur global führenden KI-Nation werden. Die Strategie baut auf bestehende Initiativen zur Förderung von Internet-basierten Technologien und Smart Manufacturing<sup>4</sup> auf. Die chinesische Regierung setzt in ihrer KI-Strategie stark auf Ansätze des maschinellen Lernens, die große Datenmengen voraussetzen. Aktuelle Defizite in der Grundlagenforschung sollen durch Investitionen in die Forschungsförderung wettgemacht werden. Zusätzlich zur Stärkung von KI-Forschung und -Anwendung will man internationalen Einfluss auch über die Entwicklung und Durchsetzung global verwendeter KI-Standards<sup>5</sup> geltend machen.

2020	Technisches Know-how und Anwendung sind auf dem gleichen Niveau wie international führende Nationen KI Unternehmen sind ein wichtiger Wachstumstreiber
2025	China erzielt Durchbrüche bei der KI-Forschung und wird in bestimmten Technologien weltweit führend KI wird der Haupttreiber der wirtschaftlichen und gesellschaftlichen Entwicklung
2030	Chinesische KI-Forschung ist weltweit führend und China wird das weltweit beste KI-Innovationszentrum

Tabelle 2: Übersicht Chinesische Meilensteine der KI-Strategie, Quelle: Chinesische Strategie S. 5ff.

In der Strategie fehlt zwar eine Definition von KI, aber das Dokument zeugt von einer intensiven Auseinandersetzung mit Anwendungsfeldern und Forschungstrends. Neben datengetriebenen Ansätzen werden auch theoretische und methodische Ansätze erwähnt, die sich an der Gehirnforschung orientieren. In der Tradition großer industriepolitischer Maßnahmen folgt Chinas KI-Strategie einem klassischen Top-Down Ansatz. In Anlehnung

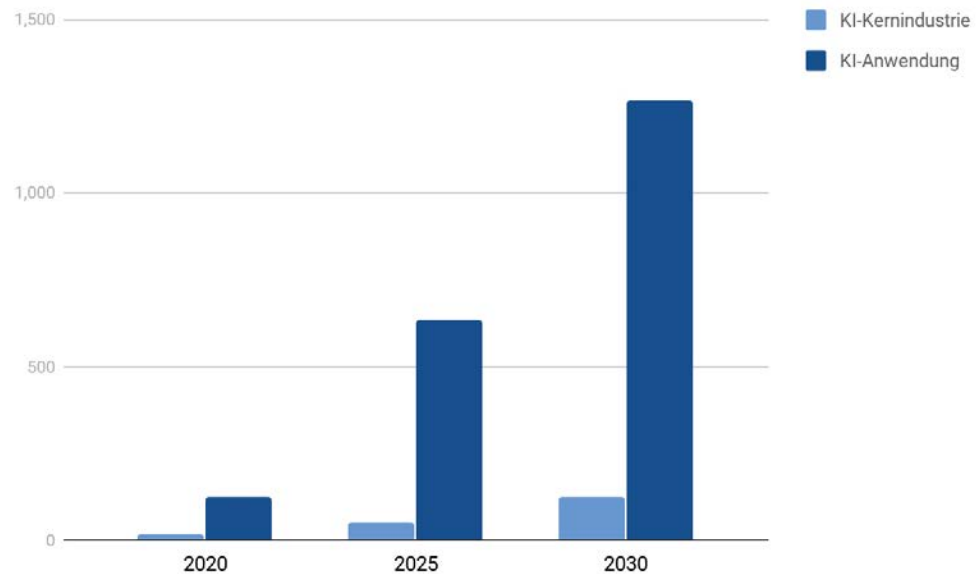
4 Siehe zum Beispiel die durch den Staatsrat der Volksrepublik China im Jahr 2015 veröffentlichte Strategie Made in China 2025 – 中国制造, <http://www.cittadellascienza.it/cina/wp-content/uploads/2017/02/loT-ONE-Made-in-China-2025.pdf>.

5 Vgl. <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/chinese-interests-take-big-seat-ai-governance-table/>



an vergangene Fünfjahrespläne wurden für die Jahre 2020, 2025 und 2030 wichtige Meilensteine definiert.

Bei den Zielvorgaben handelt es sich um Output-Indikatoren, die in der Grafik unten zusammengefasst sind. So soll die KI-Industrie ihren Beitrag zur Wirtschaft kontinuierlich auf über 1,3 Billionen EUR im Jahre 2030 steigern.



Grafik 2: Angestrebter Wert Chinesischer KI Unternehmen in Mrd. Euro. Quelle: Chinesische Strategie S. 5ff.

Zur Einordnung: Letztes Jahr betrug das BIP rund 11 Billionen EUR.<sup>6</sup> Auch wenn die chinesische Regierung bereits große finanzielle Mittel in die KI-Entwicklung steckt, finden sich interessanterweise in der KI-Strategie keine klassischen Input-Indikatoren, wie zum Beispiel Angaben über geplante staatliche Ausgaben für die Forschungsförderung.<sup>7</sup> Stattdessen listet die Strategie eine ganze Reihe von Maßnahmen auf, die in den nächsten Jahren umgesetzt werden sollen. So sollen zum Beispiel offene Plattformen die Entwicklung von KI-Anwendungen und die Kollaboration zwischen Forschung, Entwicklung und Industrie fördern. Die Strategie verweist auch auf bereits bestehende Pläne, jährlich tausende neue KI-Expert:innen auszu-

<sup>6</sup> National Bureau of Statistics of China (2018). Homepage, <http://data.stats.gov.cn/>

<sup>7</sup> New America (2018). Translation: Chinese government outlines AI ambitions through 2020, <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-government-outlines-ai-ambitions-through-2020/>

bilden.<sup>8</sup> Die Strategie wurde mittlerweile durch einen Dreijahresplan weiter konkretisiert, welcher beispielsweise die zu erreichende Genauigkeit (wie die Reduktion der Fehlerrate bei Diagnosen) für im Medizinbereich eingesetzte KI-Systeme vorschreibt.<sup>9</sup>

## Vereinigte Staaten von Amerika – zwischen Silicon Valley und Pentagon

In den Vereinigten Staaten wurde die KI-Debatte bereits unter Präsident Obama mit einigen international vielbeachteten Reports angestoßen. Neben einem nationalen Forschungs- und Entwicklungsplan zu Künstlicher Intelligenz im Oktober 2016<sup>10</sup> wurde nur zwei Monate später ein Bericht über die wirtschaftlichen Auswirkungen der Automatisierung veröffentlicht<sup>11</sup>. Unter Präsident Trump hat das Thema allerdings an Priorität verloren. Im Mai 2018 lud das Weiße Haus zwar führende Expert:innen zu einem Austausch, in dessen Nachgang auch ein kurzes Überblickspapier über bestehende Initiativen und Maßnahmen veröffentlicht wurde.<sup>12</sup> Dies kann allerdings nicht über den Eindruck hinwegtäuschen, dass die US-Regierung die Initiative weitgehend den großen Technologie-Firmen des Silicon Valley und in Bezug auf militärische Anwendungen dem Verteidigungsministerium überlässt.<sup>13</sup> Dazu gehört auch die Errichtung von Agenturen, die gezielt in Unternehmen investieren, die militärisch relevante (KI-)Technologie entwickeln. So verfügen das Pentagon mit der Defense Innovation Unit Experimental (DIUx)<sup>14</sup> und die CIA mit

---

8 Central Committee of the Communist Party of China (2018). The Thousand Talents Plan, <http://www.1000plan.org/en/plan.html>

9 Die chinesische KI-Strategie sticht auch durch ihre starke sicherheitspolitischen Dimension hervor. KI-Technologie, die im Privatsektor oder an staatlichen Universitäten und Forschungszentrum entwickelt wird, soll einen Übergang in das Militär erhalten. <https://www.cnas.org/publications/commentary/tsinghuas-approach-to-military-civil-fusion-in-artificial-intelligence>

10 Networking and Information Technology Research and Development (2016). The National Artificial Intelligence Research And Development Strategic Plan, [https://www.nitrd.gov/PUBS/national\\_ai\\_rd\\_strategic\\_plan.pdf](https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf)

11 Executive Office of the President (2016). Artificial Intelligence, Automation, and the Economy, <https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.PDF>

12 The White House (2018). Summary of the 2018 White House Summit on Artificial Intelligence for American Industry, <https://www.whitehouse.gov/wp-content/uploads/2018/05/Summary-Report-of-White-House-AI-Summit.pdf>

13 Krempl, S. (2018). Pentagon fordert KI-Strategie fürs Militär, <https://www.golem.de/news/usa-pentagon-fordert-ki-strategie-fuers-militaer-1808-136216.html>

14 Defense Innovation Unit (2018). Homepage, <https://www.diux.mil/>

In-Q-Tel<sup>15</sup> über finanziell potente Vehikel um dieses Ziel zu erreichen. Allerdings ist die Zusammenarbeit mit dem Militär in den Tech-Unternehmen sehr umstritten, wie die Kontroverse und der daraus resultierende Rückzug Googles aus der Zusammenarbeit mit dem Verteidigungsministerium im “Project Maven” zeigt.<sup>16</sup>

Angesichts der Vorstöße Chinas und vieler anderer Länder mehren sich auch in den Vereinigten Staaten die Rufe nach einer nationalen KI-Strategie.<sup>17</sup> Bisher gibt es nur einen nationalen KI-Forschungsplan.<sup>18</sup> Dieser ist allerdings breit konzipiert und adressiert auch übergreifende gesellschaftliche, wirtschaftliche und sicherheitspolitische Fragestellungen – ähnlich wie man es von einer nationalen KI-Strategie erwarten würde. Der Forschungsplan enthält ein eigenes Kapitel zur Bedeutung von Benchmarks und Standards, um die Entwicklung von KI strategisch zu steuern. Allerdings beziehen sich diese Benchmarks und Standards ausschließlich auf die Bewertung der Funktionalität und Performanz von KI-Anwendungen. Von Benchmarks in Bezug auf die erfolgreiche Umsetzung des Forschungsplans ist keine Rede.

### Großbritannien – Europas KI-Vorreiter

Die britische Regierung hat ihre KI-Ambitionen im sogenannten AI Sector Deal im April 2018 formuliert.<sup>19</sup> Im selben Zeitraum hat das britische Oberhaus auch an einem Bericht zu KI gearbeitet.<sup>20</sup> Auf letzteren ist die Regierung in einer eigenen Antwort eingegangen.<sup>21</sup> Die Fülle der Aktivitäten zeugt von der Bedeutung, die das Thema für die britische Regierung hat. Wir fokussieren hier auf den AI Sector Deal, da in diesem Papier auch einige Indikatoren und Benchmarks definiert worden sind. Im AI Sector Deal fehlt zwar eine

---

15 IQT (2018). Homepage, <https://www.iqt.org/>

16 Conger, K. und Cameron, D. (2018). Google Is Helping the Pentagon Build AI for Drones, <https://gizmodo.com/google-is-helping-the-pentagon-build-ai-for-drones-1823464533>

17 Delaney, J. (2018). France, China, and the EU All Have an AI Strategy. Shouldn't the US?, Wired, <https://www.wired.com/story/the-us-needs-an-ai-strategy/>

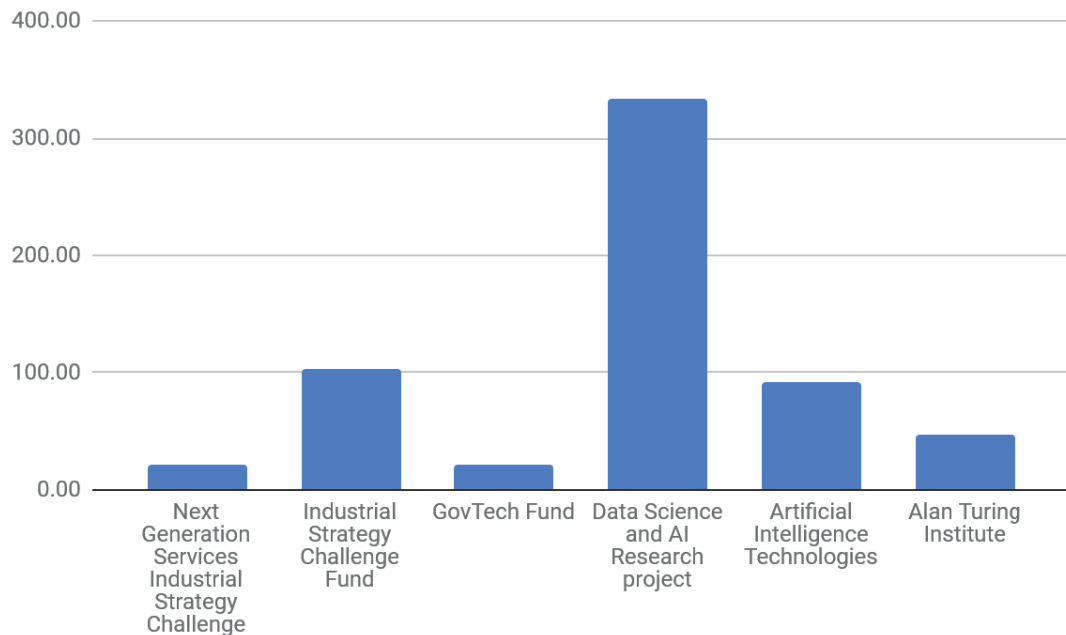
18 National Science and Technology Council (2016). The National Artificial Intelligence Research and Development Strategic Plan. [https://www.nitrd.gov/PUBS/national\\_ai\\_rd\\_strategic\\_plan.pdf](https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf)

19 Department for Business, Energy & Industrial Strategy and Department for Digital, Culture, Media & Sport (2018). Artificial Intelligence Sector Deal, <https://www.gov.uk/government/publications/artificial-intelligence-sector-deal>

20 House of Lords (2018). AI in the UK: ready, willing and able?, <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/10002.htm>

21 House of Lords (2018). Government responds to report by Lords Select Committee on Artificial Intelligence, <https://www.parliament.uk/business/committees/committees-a-z/lords-select/ai-committee/news-parliament-2017/government-response-to-report/>

Definition von KI, es wird allerdings hervorgehoben, dass KI eine Schlüsseltechnologie zur Datenverarbeitung ist. Auch der zurzeit besonders stark beachtete Machine Learning Ansatz des Deep Learning wird besonders betont. Der AI Sector Deal ist im Kern ein industriepolitisches Papier und enthält vor allem Input-Indikatoren, insbesondere bezüglich der Mittelvergabe für geplante Förderprogramme.



Grafik 3: Geplante Britische Investitionen in verschiedenen Bereichen in Mio Euro. Quelle: AI Sector Deal S. 13

Der britische Plan enthält zusätzlich noch weitere politische Maßnahmen in den fünf identifizierten Schwerpunktthemen der Strategie: Ideas, People, Infrastructures, Business Environment und Places. Hier finden sich vereinzelt auch Output-Indikatoren. So soll die Zahl der mit Regierungsmitteln geförderten Doktorand:innen in der KI-Forschung bis 2025 auf 2000 ansteigen. Mit einem speziellen Visa-Programm sollen noch einmal zusätzlich über 2000 KI-Expert:innen ins Land geholt werden. Um die für KI-Anwendungen notwendige Datenübermittlung zu erleichtern, sollen 95% Prozent aller Haushalte Zugang zu "superschnellen" Breitbandanschlüssen erhalten. Des weiteren wird das Ziel formuliert, in den nächsten 10 Jahren über 7,8 Milliarden Euro an zusätzlichen Wagniskapital für KI-Firmen zu mobilisieren. Zur Implementierung der Strategie wird die Regierung eigens ein neues KI-Büro (Office for Artificial Intelligence) ins Leben rufen. Dieses Büro soll innerhalb der einzelnen Maßnahmenfelder Erfolgskriterien entwickeln und überwachen.

## Frankreich – mit KI zu alter Stärke

Präsident Macron präsentierte die französische KI-Strategie auf der hier für eigens ausgerichteten AI for Humanity Konferenz im März 2018.<sup>22</sup> Die Strategie wurde unter Federführung des bekannten Mathematikers und Parlamentsabgeordneten Cédric Villani erarbeitet und erhielt vor allem auf Grund der prominenten politischen Unterstützung breite Aufmerksamkeit in Öffentlichkeit und Medien.<sup>23</sup> Der mehr als 140 Seiten umfassende Villani Report formuliert Frankreichs Vision für die Entwicklung und Nutzung von KI und benennt die wichtigsten gesellschaftlichen, wirtschaftlichen und politischen Herausforderungen für Frankreichs Weg zu einem der weltweit führenden KI-Standorte. Auch die französische Strategie sieht von einer Definition von KI ab und verweist stattdessen auf die historische vielschichtige Entwicklung von KI als Forschungsfeld mit Einflüssen aus Mathematik, Kognitionswissenschaften, Informatik und weiteren verbundenen Forschungsdisziplinen. Neben wissenschaftlichen Erfolgen – wie die jüngsten Entwicklungsfortschritte beim sogenannten Deep Learning – werden vor allem die Bemühungen in China und den USA und die massiven Investitionen der großen US-Technologiefirmen als Beleg für die strategische Bedeutung von KI angeführt.

Der Villani Report setzt sich ausführlich mit den unterschiedlichen Dimensionen von KI in Bezug auf Gesellschaft, Wirtschaft und Politik auseinander und benennt zahlreiche wichtige politische Handlungsfelder. Bis auf wenige Ausnahmen fehlt es allerdings an mit der Strategie verknüpften konkreten Input- und Output-Indikatoren. Der Report selbst enthält keine konkreten Investitionssummen. Bei der Vorstellung des Berichts nannte Präsident Macron die Summe von 1,5 Milliarden Euro, die die französische Regierung in den nächsten vier Jahren für die KI-Forschung bereitstellen will.<sup>24</sup> Die Input-Indikatoren der Strategie sind sehr vage gehaltene politische Maßnahmen wie zum Beispiel die Förderung von Daten-Pools, die Erhöhung der Sichtbarkeit französischer KI-Firmen oder die Nutzung öffentlicher Beschaffungsprogramme zur Förderung von KI-basierten Produkten. Andere Maßnahmen wie die Einrichtung eines Netzwerks von französischen KI-Forschungsinstituten

---

22 Gouvernement (2018). Artificial Intelligence - Making France a Leader, <https://www.gouvernement.fr/en/artificial-intelligence-making-france-a-leader>

23 Villani, C. (2018). For A Meaningful Artificial Intelligence - Towards A French and European Strategy, [https://www.aiforhumanity.fr/pdfs/MissionVillani\\_Report\\_ENG-VF.pdf](https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf)

24 Vinocur, N. (2018). Macron's €1.5 billion plan to drag France into the age of artificial intelligence, Politico, <https://www.politico.eu/article/macron-aims-to-drag-france-into-the-age-of-artificial-intelligence/>



werden ausführlicher diskutiert. Konkrete Output-Indikatoren finden sich in der Strategie nicht.

### **Finnland – intelligent Chancen identifizieren und besetzen**

Die finnische Regierung hat eine Arbeitsgruppe mit der Entwicklung einer KI-Strategie beauftragt. Dieser Prozess läuft noch. Im Dezember 2017 wurde ein erster Bericht der Arbeitsgruppe veröffentlicht.<sup>25</sup> In diesem Bericht setzen sich die Expert:innen – wenn auch nur kurz – mit dem Problem, KI zu definieren, auseinander und orientieren ihr eigenes Verständnis stark an der Übertragung menschlicher Intelligenz auf Maschinen. Die Strategie sticht im Vergleich mit anderen nationalen Strategien dadurch hervor, dass Entwicklung, Dynamik und Potenziale von KI klar analysiert werden. Weiterhin fällt auf, dass sie sich nicht nur mit Potenzial und bestehenden Stärken befasst, sondern auch klar Finnlands Schwächen im internationalen Wettbewerb benennt. Als relativ kleines Land sind die eigenen Ressourcen für KI-Forschung und der interne Markt für KI-Applikationen zwangsläufig begrenzt und die Autor:innen geben sich nicht der Illusion hin, weltweit einer der führenden KI-Standorte werden zu können. Vielmehr ist man bemüht, die Nischen zu identifizieren, in denen Finnland international eine wichtige Rolle spielen könnte. Auch Probleme bei der Kommerzialisierung von Forschung werden benannt.

Der Report definiert acht strategische Handlungsfelder. In jedem Handlungsfeld werden die wichtigsten Akteure aufgelistet und die Umsetzungsmöglichkeiten durch ein praktisches Beispiel illustriert. Damit hebt sich der Report auch von den in der Regel recht abstrakt gehaltenen Handlungsempfehlungen anderer Länder ab. Allerdings sind im Report kaum quantitative Input- und keine Output-Indikatoren enthalten, außer dem übergeordneten Ziel langfristig vier Prozent des BIP für Innovationsentwicklung auszugeben. Von der Aufstockung des Budgets zur Förderung von Innovationen um 100 Millionen Euro abgesehen finden sich im Bericht nur Handlungsempfehlungen als qualitative Input-Indikatoren.<sup>26</sup> Allerdings wird am Ende des Berichts auf die Notwendigkeit der Entwicklung von Indikatoren verwiesen, um Fortschritte bei der Umsetzung einzelner Handlungsmaßnahmen messen zu können. Auch mit dieser Erkenntnis sticht der Bericht unter den untersuchten nationalen KI-Strategien positiv hervor.

---

<sup>25</sup> Ministry of Economic Affairs and Employment (2017). Finland's Age of Artificial Intelligence, Publications of the Ministry of Economic Affairs and Employment 47/2017, ([https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap\\_47\\_2017\\_verkkojulkaisu.pdf?sequence=1&isAllowed=y](https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf?sequence=1&isAllowed=y))

<sup>26</sup> Ministry of Economic Affairs and Employment (2017) p. 53

## Europäische Union (EU) – zwischen Ethik und Industrie

Auch die EU-Kommission betrachtet KI als eine Schlüsseltechnologie, die besonderer Aufmerksamkeit bedarf. Angesichts des Wettbewerbs mit den Vereinigten Staaten und China liegt eine stärkere Bündelung der Anstrengungen auf EU-Ebene nahe. Um diesen Prozess anzustoßen, veröffentlichte die EU Kommission im April 2018 ein eigenes KI-Papier.<sup>27</sup> Die Kommission setzt darin auf eine stärkere Koordination der nationalen Initiativen der Mitgliedstaaten in einem gesamteuropäischen Rahmen und eine Einbettung der Anstrengungen in bereits laufende Projekte wie dem Digitalen Binnenmarkt. Auch die Kommission arbeitet mit einer sehr breiten Definition von KI, unter der die wichtigsten Ansätze in der KI-Entwicklung subsumiert werden können.

Das EU-Papier setzt auf zwei Schwerpunkte: Zu einem sollen Investitionen in KI-Technologien, sowohl in der Forschung wie auch in der Wirtschaft, stark ausgebaut werden, um die EU im globalen Wettbewerb vor allem mit China und den Vereinigten Staaten zu stärken. Hierbei wird besonders die Stärkung der Zusammenarbeit zwischen Forschung und Wirtschaft betont. Zum anderen erfahren auch die mit der Technologie verknüpften gesellschaftlichen und ethischen Fragestellungen eine große Aufmerksamkeit und sollen in den nächsten Jahren vertiefend bearbeitet werden. Das EU-Papier enthält ausschließlich Input-Indikatoren. Es handelt sich bei dem Papier allerdings nicht um eine umfassende KI-Strategie, sondern um einen Impuls, um die Debatte auf EU-Ebene voranzubringen und mit ersten konkreten Maßnahmen zu unterlegen. Die wichtigsten quantitativen Input-Indikatoren beziehen sich auf unterschiedliche Förderprogramme, mit denen die Kommission die KI-Forschung in der EU voranbringen möchte. Insgesamt sollen laut dem Papier die öffentlichen und privaten Investitionen in KI von insgesamt etwa vier bis fünf Milliarden Euro auf insgesamt 20 Milliarden Euro gesteigert werden. Um das Erreichen dieser und weiterer Zielsetzungen kontrollieren zu können, bedarf es allerdings entsprechender Indizes. Auch wenn die EU mit länderübergreifenden Benchmarks und Indizes viel Erfahrung hat, findet sich hierzu in dem EU Papier (noch) nichts Konkretes.<sup>28</sup>

---

<sup>27</sup> European Commission (2018). Communication Artificial Intelligence for Europe, <https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe>

<sup>28</sup> So zum Beispiel der Digital Economy and Society Index der EU European Commission (2018). The Digital Economy and Society Index, <https://ec.europa.eu/digital-single-market/en/desi>

## **Den nationalen KI-Strategien mangelt es an konkreten Zielen und Indikatoren**

In den untersuchten KI-Strategien finden sich viele Input-Indikatoren. Diese bestehen vor allem aus recht vage gehaltenen politischen Maßnahmen und Initiativen. Konkreter und damit besser messbar sind quantitative Input-Indikatoren. Bei diesen handelt es sich meistens um die Summe öffentlicher Gelder, die zur Umsetzung bestimmter Maßnahmen wie zum Beispiel den Ausbau der KI-Forschung ausgegeben werden sollen. Solche klar messbaren Input-Indikatoren findet man in den Strategien schon sehr viel seltener als die allgemeinen politischen Maßnahmen. Nur ganz wenige Strategien enthalten mit Output-Indikatoren konkrete Ziele (Outputs), die mit der KI-Strategie erreicht werden sollen.

Es gibt verschiedene Gründe, warum dies so ist. KI wird in den meisten Strategien, wenn überhaupt, dann nur sehr breit definiert. Diese Breite signalisiert Offenheit für die vielen unterschiedlichen Forschungs- und Entwicklungsansätze, die derzeit unter dem Schlagwort KI diskutiert werden. Die mangelnde klare Definition von KI macht es allerdings auch unmöglich, messbare Output-Indikatoren zu definieren. Denn wenn nicht klar ist, was genau unter den Begriff von KI fällt, kann auch nicht erhoben werden, wie viele KI-Forscher:innen oder KI-Startups es gibt beziehungsweise in Zukunft mehr geben soll.

Ein weiterer Grund für den Mangel an Output-Indikatoren ist, dass die nationalen KI-Strategien in erster Linie als Kommunikationsinstrumente zu verstehen sind. In gewisser Weise ist der Weg auch bereits das Ziel. Die Arbeit an einer nationalen KI-Strategie befördert eine strategische Auseinandersetzung mit dieser Schlüsseltechnologie innerhalb der Regierung. Hierbei stehen grundsätzliche Fragen im Mittelpunkt. Es geht um das Potenzial der neuen Technologien, um Chancen und Risiken für Wirtschaft und Gesellschaft und die Positionierung im internationalen Wettbewerb. Hiermit befassen sich nicht nur die zuständigen Ministerien. Der Strategieprozess wird in manchen Ländern auch dazu genutzt, breitere gesellschaftspolitische Diskussionen anzustoßen. Fragen nach Definitionen und Indikatoren wirken dabei eher kleinteilig und zu speziell.

Tabelle 3 gibt eine Übersicht über die in den nationalen KI-Strategien identifizierten quantitativen Input- und Output Indikatoren.<sup>29</sup> Sie zeigt, dass sich in den Strategien viel mehr Input- als Output-Indikatoren finden lassen. In der Regel sind die Input-Indikatoren nicht mit Ressourcen unterlegt. Es handelt sich vielmehr um vage formulierte, politische Maßnahmen. Zusätzlich fehlen einfach messbare Zielvorgaben, an denen sich Umsetzungsfortschritt und Erfolg der Strategie nachvollziehen lassen. Nur China und Großbritannien definieren quantitative Output-Ziele. Die finnische Strategie kündigt die Entwicklung von Indikatoren zumindest an.

Themenfeld	Input-Indikatoren	Output-Indikatoren
Forschung	<ul style="list-style-type: none"> <li>- Forschungsausgaben (CHN, EU, FIN, FRA, UK, USA)</li> <li>- Ankündigung neuer Forschungszentren (CHN, EU, FIN, FRA, UK, USA)</li> <li>- Generelle Attraktivität des Forschungsstandorts steigern (CHN, EU, FIN, FRA, USA)</li> </ul>	<ul style="list-style-type: none"> <li>- International sichtbare Beiträge zur KI- Grundlagenforschung (CHN)</li> <li>- Anzahl Doktorandenstellen (UK)</li> <li>- Anzahl Visas für Forscher (UK)</li> </ul>
Wirtschaft	<ul style="list-style-type: none"> <li>- Staatliche Förderung der KI-Wirtschaft (CHN, EU, FIN, FRA, UK, USA)</li> <li>- Staatliche KI-Beschaffung (CHN, FRA, UK, USA)</li> <li>- Staatliche Förderung der IT-Infrastruktur (CHN, EU, UK)</li> <li>- Abbau regulatorischer Hürden / Anreize für die KI-Nutzung durch Firmen (EU, FIN, USA)</li> <li>- Ausbau von Open Data Initiativen (EU, FIN, FRA)</li> </ul>	<ul style="list-style-type: none"> <li>- Produktivität und Wettbewerbsfähigkeit der Wirtschaft durch KI erhöhen (UK, FIN)</li> <li>- Benchmarks für KI-Wertschöpfung (CHN)</li> </ul>

<sup>29</sup> Die Auswertung für die Tabelle beruht auf einer Textanalyse der jeweiligen Länderstrategie. In der Regel werden Indikatoren in diesen Dokumenten nicht gesondert aufgeführt. Meistens finden sie sich stattdessen im Fließtext, was die klare Identifikation erschwert. Für die Nennung in der Tabelle genügte es zudem, wenn ein Land eine bestimmte Messzahl erwähnt, z.B. Forschungsausgaben, ohne einen konkreten Betrag zu nennen, z.B. X € bis 2030.

Themenfeld	Input-Indikatoren	Output-Indikatoren
Bildung / Fachkompe- tenzen	<ul style="list-style-type: none"> <li>- Investitionen in MINT in Schulen / Universitäten / Berufsausbildung / Weiterbildung (EU, FIN, FRA, UK, USA)</li> <li>- Einführung neuer Studiengänge mit KI-Fokus (EU, FIN, FRA, UK, USA)</li> </ul>	<ul style="list-style-type: none"> <li>- Führende internationale Rolle in "Ethischer" KI / Datennutzung (EU, FRA, UK)</li> <li>- Internationale technische KI-Standards definieren (CHN, USA)</li> <li>- Trendsetter für EU-Debatte sein (FIN)</li> </ul>

Tabelle 3: Quantitative Input- und Output-Indikatoren in zentralen Handlungsfeldern nationaler KI-Strategien, Quelle: Stiftung Neue Verantwortung

Wenn aus den Strategien mehr als Absichtserklärungen und Kommunikationsinstrumente werden soll, muss man sich konkreter mit den Zielen und den für die Zielerreichung einzusetzenden Ressourcen befassen. Dabei kommt man um eine Auseinandersetzung mit Definitionen und Indikatoren nicht herum. In einigen nationalen Strategien wird diese Herausforderung bereits berücksichtigt. Von den hier diskutierten Beispielen stechen vor allem China und Großbritannien hervor. Allerdings findet man auch in diesen beiden Strategien keine Ansätze für eine systematische Entwicklung von Input- und Output-Indikatoren.

## KI-Indizes

In den letzten Jahren sind eine Vielzahl von Berichten und Analysen über die Potenziale und Entwicklungsfortschritte in der KI veröffentlicht worden. Einige der Analysen beruhen auf empirischen Erhebungen mit dem Ziel, Stärken und Entwicklungen in der KI-Forschung, KI-Investitionen und KI-Anwendung systematisch zu erfassen. Wir betrachten hier eine Auswahl dieser KI-Indizes bezüglich ihrer Methodik. Dabei haben wir nur Indizes berücksichtigt, die laut eigener Darstellung auf KI fokussiert sind. Unser Ziel war es nicht, alle KI-Indizes zu erfassen und systematisch zu analysieren. Dies würde weit über die Zielsetzung dieses kurzen Papiers hinausgehen. Vielmehr ging es uns darum, aus der Auseinandersetzung mit der Methodik der untersuchten KI-Indizes erste Anregungen, Ideen und Herausforderungen für die Entwicklung eines Benchmarkings der deutschen KI-Strategie abzuleiten.

Wir haben eine große Bandbreite unterschiedlicher KI-Indizes identifiziert. Ansatz und Umfang der Indizes variieren sehr stark. Anhand der Methodik lassen sich aber vier unterschiedliche Ansätze unterscheiden.

Methodik	Vorteil	Nachteil
Umfrage: meist Expertenbefragung tiefergehende Einblicke, die öffentlich zugänglichen Daten in der Regel nicht ermöglichen	tiefergehende Einblicke, die öffentlich zugänglichen Daten in der Regel nicht ermöglichen	die Expertenauswahl und Antwortrate prägt die Ergebnisse vor (bias)
Komposition: Auswertung und Gewichtung mehrerer Indizes	relativ leicht umsetzbar, da keine neue Datenquellen erschlossen werden müssen	ungeeignet um neue Entwicklungen in den Blick zu nehmen, die Auswahl der genutzten Indizes prägt Ergebnisse vor (bias)
Daten-getrieben: Auswertung orientiert sich an verfügbaren Daten	macht in der Regel eine Auswertung mit statistischen Methoden (und damit den Ausschluss der Nachteile der Umfrage) möglich	Die Verfügbarkeit bestimmter Datenquellen prägt die Ergebnisse (bias)
Indikatoren-getrieben: Indikatoren werden eigens theoriegeleitet entwickelt	Auswertung beruht auf einem tieferen Verständnis des Untersuchungsgegenstandes, aus dem Indikatoren abgeleitet werden	oft mangelt es für die Umsetzung notwendigen Datenquellen oder ist eine saubere Umsetzung extrem aufwendig, da neue Datenquellen und -Analysen eigens entwickelt werden müssen

Tabelle 4: Übersicht über Index-Methodologien. Quelle: Stiftung Neue Verantwortung

Jede Methode hat ihre Stärken und Schwächen. Bei der Befragung von Expert:innen und Entscheider:innen können tiefergehende und komplexe Themen behandelt werden. Insbesondere Fragestellungen, für die es noch keine guten quantitativen Datenquellen beziehungsweise statistische Erhebungen gibt, lassen sich gut bearbeiten. Daher eignen sich Umfragen besonders gut, um neue Trends zu identifizieren und zu untersuchen. Allerdings hängt das Ergebnis der Befragung stark von der Auswahl der Expert:innengruppe, ihrer Antwortrate und -Qualität sowie der Auswahl und Ausrichtung der Fragen ab. Hinzu kommt, dass es sich in der Regel um nicht überprüfbare Auskünfte und subjektive Einschätzungen handelt. Ob ein Unternehmen zum Beispiel ausreichend in die KI-Entwicklung investiert, kann stark variieren, je nachdem, wen man fragt, wie gut die Person sich auskennt und wie sie die Situation einschätzt, beziehungsweise welche Informationen der Person zur Verfügung stehen und sie preisgeben möchte.

Komposition haben wir den Ansatz genannt, aus bestehenden Indizes einen neuen Index zu entwickeln. Dieser Ansatz ist relativ leicht umsetzbar, da keine neue Datenquellen erschlossen und keine eigenen Datenerhebungen durchgeführt werden müssen. Allerdings lassen sich so im Gegensatz zur Umfrage weder neue Trends erkennen, noch untersuchen. Die Qualität der Komposition ist zudem stark von der Qualität der genutzten Indizes und der Frage abhängig, ob und wie diese sinnvoll zu einem neuen Gesamtbild aggregiert werden können.

Daten-getriebene Ansätze richten ihre Indikatoren nach der Verfügbarkeit von Daten aus. Digitalisierung und neue Techniken der Datenerhebung und -Auswertung wie zum Beispiel das automatisierte Erfassen und Auswerten von Texten auf Webseiten (durch Einsatz sogenannter Webcrawler) machen neue Datenquellen für die Analyse zugänglich. Dabei neigt man aber stärker dazu, die Indikatoren nach verfügbaren Datenquellen auszurichten, statt nach der Frage, welche Indikatoren für die Messung bestimmter Fragestellungen am besten geeignet sind. Denn Datensätze, auch wenn sie leicht zugänglich und mit komplexen, statistischen Methoden ausgewertet werden, liefern keine guten Antworten, wenn sie die untersuchte Fragestellung nicht adäquat abbilden.

Anstatt die Entwicklung von Indikatoren von der Verfügbarkeit von Daten abhängig zu machen, stellen Indikatoren-getriebene Ansätze die Entwicklung der Indikatoren in den Mittelpunkt. Indikatoren-getriebene Ansätze leiten aus einem tieferen, theoretischen Verständnis des zu untersuchenden Phänomens die entsprechenden Indikatoren für den Index ab. Dies erscheint auf den ersten Blick als Königsweg der Indexentwicklung. Daher nützen die besten, aus einem tiefen theoretischen Verständnis des Untersuchungsgegenstands sauber abgeleiteten Indikatoren wenig, wenn man nicht über geeignete Messmethoden oder Zugang zu den hierfür benötigten Daten, Informationen und Auswertungsmöglichkeiten verfügt. Der perfekte Indikator, der sich nicht operationalisieren lässt, bringt uns für die Entwicklung von Benchmarks und Indizes auch nicht weiter.

Die Übersicht über die unterschiedlichen Ansätze zeigt, dass es nicht den einen perfekten Ansatz gibt. Die unterschiedlichen Ansätze haben ihre eigenen Stärken und Schwächen. Und oft wird nicht nur ein Ansatz gewählt, sondern eine Kombination der hier vorgestellten methodischen Ansätze. Elemente aller vier Ansätze lassen sich daher in unterschiedlichen Ausmaß über die von uns untersuchten Indizes verteilt finden. Allerdings lässt sich die Klassifizierung nicht immer ganz so sauber vornehmen, wie unten in der Tabelle dargestellt. Einige Indizes vereinigen mehrere Ansätze. So beruht



die Studie *AI: The New Frontier* des McKinsey Global Institutes sowohl auf einer Umfrage wie auch datengetriebener Analysen von Investitionen in KI. Und auch wenn Überlegungen zu einzelnen Indikatoren oft nicht transparent gemacht werden, so ist doch davon auszugehen, dass die Autor:innen aller Indizes über diese Frage bei der Entwicklung ihres Index oder Rankings nachgedacht haben.

Unsere Klassifikation beruht daher lediglich auf einer Einschätzung, welcher methodischer Ansatz im Index überwiegt. Keiner der Indizes kann als rein Indikatoren-getrieben gelten. Hierfür fehlt die entsprechende tiefergehende Auseinandersetzung mit dem Untersuchungsgegenstand (der Frage nach der Definition von KI) und darauf aufbauend die entsprechende Entwicklung von Indikatoren. Allerdings steckt in der Indikatorenauswahl immer auch eine theoretische Annahme über den Untersuchungsgegenstand und seiner Beziehung zum Indikator. Dieses Problem wird weiter unten noch etwas ausführlicher diskutiert.



Publikation / Plattform	Herausgeber	Erstmals publiziert	Aktuelle Publikation	Ziel	Abdeckung	Methodologie
AI Maturity	Infosys	2016	2016	Überblick der KI-Maturität verschiedener Industrien	7 Länder	Umfrage
AI: The Next Frontier	McKinsey Global Institute	2017	2017	Zeigt KI-Entwicklungen und mögliche wirtschaftliche Einsatzzwecke	10 Länder	Umfrage
AIQ	Accenture	2017	2017	Zeigt die Verbreitung/Bereitschaft von KI-Anwendungen in Privatunternehmen	10 Länder	Umfrage
Automation Readiness Index	Economist Intelligence Unit & ABB	2018	2018	Misst die Fähigkeit verschiedener Länder ihre Wirtschaft auf Automatisierung umzustellen	25 Länder	Komposition
Government AI Readiness Index	Oxford Insights	2017	2017	Zeigt wie eine Regierung aufgestellt ist um für ihre Dienstleistungen von KI zu profitieren	35 Länder (OECD)	Komposition
AI Index	AI 100	2017	2017	Überblick zu verschiedenen KI-Dimensionen (Forschung, Anwendung, Investitionen etc.)	Global	Datengetrieben
AI Progress Measurement	Electronic Frontier Foundation	2017	2017	Zusammenstellung diverser Probleme/Benchmarks um den Fortschritt der KI-Forschung zu messen	Global	Datengetrieben
The European Artificial Intelligence Landscape	Asgard & Roland Berger	2017	2017	Überblick KI-Startups	Europa	Datengetrieben
Global AI Talent Report	Jean-François Gagné	2018	2018	Überblick globaler KI-Talentpool	Global	Datengetrieben
AI Trends	CB Insights	2018	2018	Zeigt Trends für industrielle KI-Anwendungen	Global	Präsentation von Trends
State of AI	Nathan Benaich & Ian Hogarth	2018	2018	Überblick Entwicklungen in den Bereichen KI-Forschung, Talent, Industrie und Politik	Global	Präsentation

Tabelle 5: Übersicht führender Indizes zu Künstlicher Intelligenz. Quelle: Stiftung Neue Verantwortung

Beim Blick auf die identifizierten Indizes fällt auf, dass die meisten von Berater:innen oder Beratungsunternehmen erstellt wurden. Somit überrascht es wenig, dass die Analyse von Trends und Entwicklungen in der KI in fast allen Reports auch der Ableitung von strategischen Handlungsempfehlungen dient. Mit den Reports wollen sich die Beratungsunternehmen in neuen Themenfeldern positionieren. Wissenschaftlichen Standards halten die meisten Indizes nicht Stand und sie erheben auch nicht diesen Anspruch. Vor allem mangelt es an grundlegender Transparenz und kritischer Auseinandersetzung mit der Methodik. Anstatt die Grenzen der Aussagekraft ihrer Indikatoren offen zu diskutieren und die Ergebnisse des Index entsprechend kritisch einzuordnen, wird auf die Methodik nur spärlich eingegangen. Die Informationen, die sich zur Methodik finden lassen, sind oft in Fußnoten oder einem Appendix am Ende versteckt.

Grundlegendes Problem aller untersuchten Indizes ist, dass sie sich unzureichend mit ihrem Untersuchungsgegenstand auseinandersetzen. Die meisten beschränken sich auf eine sehr breite Definition von KI, die vor allem den Vergleich zu menschlicher Intelligenz in den Mittelpunkt stellt. So versteht *Accenture* unter KI Technologien, die “menschliche Fähigkeiten durch Wahrnehmung, Auffassungsgabe, Handeln und Lernen” erweitern.<sup>30</sup> *McKinsey* sieht in KI die “Fähigkeit von Maschinen menschenähnliche Intelligenz zu besitzen”.<sup>31</sup> Beide diskutieren allerdings nicht weiter, wie sie auf Basis dieses breiten KI-Verständnisses bestimmte Technologien der KI zuordnen oder nicht.

Eine solche Ableitung muss allerdings vorgenommen werden, wenn zum Beispiel KI-Investitionen erfasst werden. Hierfür muss klar abgegrenzt werden, welche Investition als KI-Investition gilt und welche nicht. Dies mag auch ein weiterer Grund für die Popularität der Umfrage-Methodik sein. Denn hier können Daten auch ohne klare Definitionen erhoben werden. Die Antworten, wie oder wie stark, sich laut einer Führungskraft ein Unternehmen im Bereich KI engagiert, hängt dann vom Verständnis von KI des:der Befragten ab. Wenn den Befragten hier keine klaren und für sie operationalisierbaren Definitionen und Eingrenzungen von KI vorgelegt werden, wirft dies allerdings

---

30 Accenture (2017). Boost your AIQ - Transforming into an AI Business, [https://www.accenture.com/t20170614T050454Z\\_\\_w\\_/us-en/\\_acnmedia/Accenture/next-gen-5/event-g20-yea-summit/pdfs/Accenture-Boost-Your-AIQ.pdf#zoom=50](https://www.accenture.com/t20170614T050454Z__w_/us-en/_acnmedia/Accenture/next-gen-5/event-g20-yea-summit/pdfs/Accenture-Boost-Your-AIQ.pdf#zoom=50), S. 5

31 McKinsey Global Institute (2017). Artificial Intelligence - The Next Frontier?, <https://www.mckinsey.com/~media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx>, S. 6

große Fragen bezüglich der Vergleichbarkeit der einzelnen Antworten auf.<sup>32</sup> Sie drücken dann nicht mehr als ein recht diffuses Stimmungsbild aus – können allerdings kaum als solide Grundlage genutzt werden, um Unternehmen in Bezug auf Aktivitäten zu klassifizieren und Rangordnungen zu erstellen.

Auch wenn *Accenture* in seinem Bericht die Frage, welche Technologien eigentlich unter KI fallen, nicht tiefergehend thematisiert, musste das Beratungsunternehmen sich mit dieser Frage beschäftigen. Denn neben Interviews wurde das Web auch nach relevanten Unternehmensaktivitäten in diesem Bereich durchsucht. Hierfür wurde eine Liste von KI-Technologien und -Anwendungen erstellt, die im Bericht auch am Ende dokumentiert ist.<sup>33</sup> McKinsey erklärt zur Operationalisierung seiner breiten KI-Definition, dass man sich auf fünf KI-Systeme fokussiert habe: Robotik und autonomes Fahren, Computer Vision, Sprachsysteme, virtuelle Agenten und Machine Learning. Warum genau diese Kategorien ausgewählt wurden und wie diese Kategorien zu verstehen sind, wird nicht weiter thematisiert. Ebenso wird nicht darauf hingewiesen, dass diese Technologien unterschiedliche Einsatzpotenziale für verschiedene Wirtschaftssektoren haben, die im Bericht verglichen werden, und dass sich damit wahrscheinlich auch ein Großteil der Varianz in Bezug auf den Grad des Einsatzes der Technologie in den jeweiligen Sektoren erklären ließe.<sup>34</sup>

Datengetriebene Ansätze haben gegenüber Umfragen den Vorteil, dass sie alle mit zwischenmenschlicher Kommunikation verbundene methodische Schwierigkeiten vermeiden können. Gerade quantitative Erhebungen und Auswertungen werden häufig mit Objektivität und hoher Aussagekraft assoziiert. Allerdings sind auch datengetriebene Ansätze nicht frei von Problemen und möglichen Verzerrungen. Ein klassisches Beispiel eines datengetriebenen Ansatzes ist die Erfassung und Auswertung der geographischen Verteilung von KI-Startups von *Asgard* und *Roland Berger*. Eine lange Excel-Tabelle mit KI-Startups von *Asgard* bildete die ursprüngliche Datengrundlage. Da es diese Datengrundlage bei *Asgard* schon gab, macht es für beide Unternehmen Sinn, den Fokus auf KI-Startups zu legen. Allerdings liegen diesem datengetriebenen Vorgehen starke Annahmen zugrunde, die nicht

---

<sup>32</sup> Ob und wenn ja in welcher Form dies im Rahmen der Befragung geschehen ist, lässt sich hier nicht nachvollziehen, da die Berichte ihre Umfragemethodik nur allgemein beschreiben aber keine Details offenlegen.

<sup>33</sup> Accenture (2017), S. 28

<sup>34</sup> Einen guten Überblick über unterschiedliche Einsatzpotenziale, bietet des Expertengutachten *Forschung und Innovation in seiner einführenden Übersicht über autonome Systeme*. S.65-81 [https://www.e-fi.de/fileadmin/Gutachten\\_2018/EFI\\_Gutachten\\_2018.pdf](https://www.e-fi.de/fileadmin/Gutachten_2018/EFI_Gutachten_2018.pdf)

weiter diskutiert werden. Denn der Vergleich der Stärken von KI-Ökosystemen auf Basis der Anzahl von Startups impliziert, dass KI-Startups für die Stärke und Potenzial der KI-Ökosysteme entscheidend sind. Universitäre Forschung und angewandte Forschung und Entwicklung in großen Unternehmen werden nicht erfasst – wahrscheinlich weil die zusätzliche Datenerfassung und -Auswertung zu aufwändig geworden wäre. Denn die Autor:innen der Analyse gestehen selbst ein, dass die Dynamik in der KI-Entwicklung nicht nur von Startups, sondern auch von den großen digitalen Plattformen und Forschungsinstituten kommt.

Positiv hervorzuheben ist, dass *Asgard* seinen Datensatz zur Verfügung stellt. Allerdings kann das nicht über methodische Probleme hinwegtäuschen. Es wird nicht weiter ausgeführt, wie die Daten erschlossen wurden. Neben Machine Learning, Deep Learning, Bilderkennung und Sprachverarbeitung wurden auch andere “Frontier AI technology” berücksichtigt, ohne zu spezifizieren, was darunter zu verstehen ist. Und wie wurde auf Systematik bei der Erschließung der Daten geachtet? Wurden bestimmte Quellen herangezogen oder beruht die Liste hauptsächlich auf Desktop-Recherchen? Wie wurde mit dem im Bericht angesprochenen Sprachbarrieren, zum Beispiel was die Erfassung chinesischer KI-Startups betrifft, umgegangen? Auch der Begriff Startup wird nicht weiter definiert. So sind im Datensatz KI-Startups mit hohen Marktvaluierungen und großen Entwickler:innenteams ebenso wie kleine “Garagen-Startups” vertreten. Beispielsweise findet sich die von Google für rund 580 Mio. Euro aufgekaufte<sup>35</sup> und bereits seit 2010 existierende Firma *DeepMind* mit über 50 Entwickler:innen ebenso in der Liste wie das Berliner Start-up *Bunch*, welches 2016 gegründet wurde, weniger als 10 Mitarbeitende hat und von Seed-Finanzierung lebt. Das wirft die Frage auf, wie aussagekräftig ein Vergleich basierend auf Anzahl der Startups ist, wenn sich zum Beispiel in einem Ökosystem nur wenige aber dafür bereits hoch bewertete und international aktive KI-Startups befinden, während ein anderes Ökosystem zwar viele, gleichzeitig aber nur schwach kapitalisierte und vornehmlich lokal tätige Startups aufweist. Ohne genauere Angaben zu Auswahlkriterien und Berücksichtigung weiterer Faktoren wie Mitarbeiter:innenzahl, Internationalisierungsgrad und Marktkapitalisierung besteht die Gefahr, dass Äpfel mit Birnen verglichen werden.

Der Global AI Talent Report ist auch datengetrieben<sup>36</sup>. Der Herausgeber Jean-François Gagné nutzt Daten des Karrierenetzerks LinkedIn für eine globale

---

<sup>35</sup> Economist (2016). What DeepMind brings to Alphabet, <https://www.economist.com/business/2016/12/15/what-deepmind-brings-to-alphabet>

<sup>36</sup> Gagne, J. (2018). Global AI Talent Pool Report, <http://www.jfgagne.ai/talent/>

Analyse der geographischen Verteilung von KI-Expert:innen. Zusätzlich werden auch die Programme führender KI-Konferenzen ausgewertet. Der Bericht geht mit den den Datenquellen zugrundeliegenden Einschränkungen bezüglich Aussagekraft und Interpretation der Daten sehr offen um. So beginnt die Analyse mit einer Auseinandersetzung mit der Frage, wie KI-Talent überhaupt zu definieren ist. Dabei setzt der Global AI Talent Report auf eine Mischung von relevanten Domänenwissen in KI-Subdisziplinen, Programmierkenntnissen und wissenschaftlicher Ausbildung (PhD). So versucht man mit dem Problem umzugehen, dass Nutzer:innen auf LinkedIn ihre Kompetenzen selbst angeben und ihre Fähigkeiten zu hoch oder falsch einschätzen könnten. Auch das Problem, das LinkedIn in unterschiedlichen Regionen der Welt unterschiedlich stark verbreitet ist, wird angesprochen. Auch mit diesen Einschränkungen lassen aus den Daten interessante Entwicklungen und Trends ableiten. Der Global AI Talent Report ist ein gutes Beispiel für das Potenzial, das in neuen Datenquellen, wie zum Beispiel Karrierenetzen, für die Indexentwicklung steckt.

Der von den Wissenschaftlern Yoav Shoham (Stanford University) und Eric Brynjolfsson (MIT) betreute AI Index enthält eine Vielzahl von Indikatoren. Die vielen Analysen im Index sind im Kern aber auch datengetrieben. Sie beruhen auf der Auswertung frei im Web zugänglicher Datenquellen. Die Analysen erstrecken sich von allgemeinen Entwicklungen in der KI bis hin zu spezielleren Analysen in Wissenschaft, Wirtschaft, Open Source Software, öffentlicher Wahrnehmung und Indikatoren technischer Performanz von KI-Anwendungen. Damit sticht der Bericht durch seine thematische Breite und Vielfalt unterschiedlicher Indikatoren und Benchmarks hervor. Die Autoren warnen allerdings eingangs direkt vor voreiligen Schlussfolgerungen, insbesondere bezüglich länderübergreifender Vergleiche und internationaler Trends, da die Datenquellen hauptsächlich aus den Vereinigten Staaten stammen und Informationen aus anderen Ländern oft unzureichend oder gar nicht erfasst werden.<sup>37</sup> Die Autoren verstehen ihren Bericht daher auch als langfristiges und kollaboratives Projekt. Zu Anregungen, Kritik und Hinweisen auf weitere Datenquellen wird explizit aufgefordert. Alle Datensätze werden auf der zum Bericht gehörenden Webseite veröffentlicht und jede Datenquelle wird im Appendix auch noch einmal in Bezug auf Definitionen und Aussagekraft diskutiert. Als einer der wenigen an einer Universität angesiedelten Projekte hebt sich der AI Index durch Transparenz und kritische Auseinandersetzung mit der Methodik hervor. Aufgrund der breit angelegten Ambitionen liefert dieser Bericht auch viele interessante Anhaltspunkte für die Entwicklung eines an einer nationalen KI-Strategie ausgerichteten

---

<sup>37</sup> AI Index (2017), 2017 Annual Report, <https://cdn.aiindex.org/2017-report.pdf>, S. 8



KI-Benchmarkings oder Index.

Das AI Progress Measurement der *Electronic Frontier Foundation* (EFF) vollzieht wichtige Fortschritte in der KI-Entwicklung nach.<sup>38</sup> Das Projekt der EFF dokumentiert Verbesserungen in der Performanz bei wichtigen KI-Problemen wie der Bild- oder Spracherkennung. Bei vielen dieser Probleme steht die Verringerung der Fehlerrate im Mittelpunkt, deren Reduktion zum Beispiel bei der Bilderkennung und der Spracherkennung über die letzten Jahre gezeigt wird. Aber das AI Progress Measurement erfasst auch Fortschritte von KI-Systemen bei der Bewältigung komplexer Strategiespiele.<sup>39</sup> Der Index ermöglicht Rückschlüsse auf den Entwicklungsfortschritt der KI in zentralen Problemfeldern. Er liefert damit einen wichtigen Ansatz, um Entwicklungen und Fortschritt im KI-Forschungsfeld zu erfassen. Weitere Projekte und Ansätze, die sich mit technischen Benchmarks für KI-Soft- und Hardware befassen, sind in Tabelle 6 aufgeführt.

Kaggle Wettbewerbe	<a href="https://www.kaggle.com/competitions">https://www.kaggle.com/competitions</a>
Stanford DAWNbench	<a href="https://dawn.cs.stanford.edu/benchmark/">https://dawn.cs.stanford.edu/benchmark/</a>
TensorFlow Benchmarks	<a href="https://www.tensorflow.org/performance/benchmarks">https://www.tensorflow.org/performance/benchmarks</a>
MLPerf Benchmark	<a href="https://mlperf.org/">https://mlperf.org/</a>
ImageNet Large Scale Visual Recognition Challenge	<a href="https://www.image-net.org/challenges/LSVRC">https://www.image-net.org/challenges/LSVRC</a>
Diskussion Benchmarks	<a href="https://siliconangle.com/2018/05/07/challenge-finding-reliable-ai-performance-benchmarks/">https://siliconangle.com/2018/05/07/challenge-finding-reliable-ai-performance-benchmarks/</a>

Tabelle 6: Technische Benchmarks und Wettbewerbe. Quelle: Stiftung Neue Verantwortung

In unserer Liste finden sich auch zwei Indizes, die für ihre Analyse auf bereits bestehenden Indizes zurückgreifen. Diesen Index-Typ nennen wir, wie bereits erwähnt, Komposition. Der Vorteil dieses Ansatzes ist, dass er sich recht leicht umsetzen lässt, da keine neuen Datenquellen eigens erschlossen werden müssen. Der Nachteil ist, dass Schwächen und mögliche Fehler der genutzten Indizes reproduziert werden, und sich keine wirklich neuen Fragestellungen in den Blick nehmen lassen. Der Government AI Readiness Index der Beratungsfirma *Oxford Insights* beruht auf einem Score, der die Regierungen aller OECD Länder bezüglich der Voraussetzungen bewertet, vom Einsatz von KI-Systemen in der eigenen Regierungsarbeit und im öf-

<sup>38</sup> Electronic Frontier Foundation (2018). AI Progress Measurement, <https://www.eff.org/de/ai/metrics>

<sup>39</sup> Open AI Five ist eines der prominenten Beispiele, in der Fortschritt in der KI-Entwicklung über Erfolge bei Strategiespielen dokumentiert und gemessen werden. <https://blog.openai.com/openai-five/>

fentlichen Sektor profitieren zu können.<sup>40</sup> Der Index benötigt keine tiefere Auseinandersetzung mit der Definition von KI, weil es im Kern gar nicht um KI geht. Stattdessen beruht der Index auf nicht weiter ausgeführten, aber sicherlich auch nicht unplausiblen Annahmen, dass Länder mit mehr digitalen Kompetenzen, besserer digitaler Infrastruktur und einem innovativen öffentlichen Sektor eher in der Lage sind, KI-Anwendungen in die eigene Arbeit zu integrieren als Länder, in denen diese Voraussetzungen fehlen. Zwar werden die zur finalen Score herangezogenen Indizes genannt. Wie diese Indizes aber genau kombiniert und gewichtet worden sind, wird nicht erklärt. Allerdings wären weitere Erklärungen angebracht und wünschenswert. Denn die Ergebnisse sind überraschend: Beispielsweise wird das im Bereich e-Government weltweit führende Estland erst auf Platz 23 hinter Mexiko und vor Polen geführt.

Der Automation Readiness Index der Intelligence Unit des *Economist* (EUI) rankt Länder bezüglich ihrer Fähigkeit und Bereitschaft, mit der von KI erwarteten Transformation von Wirtschaft und Gesellschaft politisch umzugehen. Dabei fokussiert die Analyse auf drei Bereiche: Innovationskraft, Bildungspolitik und Arbeitsmarktpolitik. Der Automation Readiness Index erstellt jeweils eine 25 Länder umfassende Rangliste für jeden dieser drei Teilbereiche und errechnet daraus auch einen Gesamtscore. Im Gegensatz zum Government AI Readiness Index werden nicht nur die genutzten Indikatoren, sondern auch die Gewichtung angegeben. Die Gesamtscore setzt sich aus 52 Indikatoren zusammen, von denen EUI 45 selbst entwickelt hat. Zusätzlich werden sieben weitere Indikatoren anderer Organisationen genutzt. Allerdings enthält der Bericht nur wenig Auskunft über die 45 selbst entwickelten Indikatoren. Die meisten Indikatoren beziehen sich nur darauf, ob es bestimmte Regierungsprogramme, zum Beispiel zur Forschungsförderung oder im Bildungsbereich überhaupt gibt. Wie viele Ressourcen für diese Programme jeweils aufgewendet werden und ob sie wirksam sind, wird nicht weiter berücksichtigt. Der Index vermittelt damit einen groben Eindruck, ob Länder sich mit Fragen von KI und Automatisierung in unterschiedlichen Politikfeldern befassen. Aus den erhobenen Daten lässt sich allerdings, anders als suggeriert, nicht wirklich ableiten, wie gut diese Länder vorbereitet sind.

Ein weiteres, häufig anzutreffendes Format ist der KI Trend-Report. Diese Berichte lassen sich zwar nicht klar unserer Index-Methodik zuordnen, sie bieten aber wichtige Einblicke, insbesondere bezüglich neuer Entwicklungen. Daher sollten sie hier nicht unerwähnt bleiben. Der von *CBInsights* her-

---

<sup>40</sup> Oxford Insights (2017). Government AI Readiness Index, <https://www.oxfordinsights.com/government-ai-readiness-index/>

ausgegebene “Top AI Trends to Watch in 2018”-Bericht identifiziert 13 Trends im Bereich der KI.<sup>41</sup> Diese Trends könnten Vorboten größerer Entwicklungen sein. Ein Trend handelt zum Beispiel von der Frage, wie durch Automatisierung gesenkte Lohnkosten zu einer Rückführung industrieller Produktion in die USA führen könnten. Andere Trends erfassen neue Anwendungsbeispiele für den Einsatz von KI oder Vergleiche zwischen den USA und China bei der Zahl von Patentanmeldungen. Auch neue Forschungstrends, zum Beispiel bezüglich der Architekturen neuronaler Netzwerke oder Edge-Computing, werden vorgestellt. Die Präsentation “State of AI” von Nathan Benaich und Ian Hogarth gibt einen noch umfassenderen Überblick über neueste KI-Trends. Sie ist in mehrere Kapitel unterteilt: signifikante Forschungsfortschritte in Hard- und Software, Humankapital (AI Talent), Wirtschaft und Anwendungen, Gesellschaft und Politik und Thesen zur weiteren Entwicklung (Prognosen). Die gesamte Präsentation umfasst 156 einzelne Slides und schließt mit persönlichen Prognosen der beiden Autoren. Die Trend-Reports fallen hier ein wenig aus dem Rahmen, da sie keine systematische Methodik verwenden, sondern stark von Expertise und Einschätzungen ihrer Autoren geprägt sind. Allerdings sind diese Reports für alle, die ein Technologie-Monitoring betreiben wollen, eine wertvolle Quelle.

### **Aufwand vs. Aussagekraft: das Dilemma von KI-Indizes**

Eine kurze Analyse einiger der bekanntesten KI-Indizes liefert viele interessante Ansatzpunkte, die bei der Entwicklung eines an einer nationalen KI-Strategie angelehnten Index weiterhelfen können. Vor allem wird klar, dass bei der Entwicklung eines solchen Index komplexe methodische Probleme adressiert werden müssen. Diese Auseinandersetzung muss mit der Frage, was alles unter dem Begriff KI verstanden und subsumiert wird, beginnen. Darauf aufbauend muss geklärt werden, was wir im Bereich der KI eigentlich messen wollen und welche Indikatoren sich dafür am besten eignen. Und schließlich müssen wir uns mit den Stärken und Schwächen der Indikatoren offen und kritisch befassen. Gemäß unseres methodischen Klassifikationsschemas würde ein solches Vorgehen einem Indikator-getriebenen Ansatz entsprechen.

Es gibt gute Gründe, warum sich in unserer Übersicht kein Indikator-getriebener Ansatz findet. Ein solches Vorgehen wäre extrem aufwendig. Gleichzeitig besteht die Gefahr, dass aus diesem Vorgehen heraus zwar sehr gute

---

<sup>41</sup> Fußnote mit Link zur Report - Weitere Hinweise über wichtige Trends liefert der CBInsights Report über 100 KI Startups. “AI 100: The Artificial Intelligence Startups redefining Industries” <https://www.cbinsights.com/research/artificial-intelligence-top-startups/>



Indikatoren entwickelt werden, diese sich aber in der Praxis, wenn es um Datenerfassung und -Analyse geht, nicht umsetzen lassen. Die hier diskutierten Ansätze weisen zwar alle methodische Schwächen auf, sie zeigen aber auch, wie man mit begrenztem Ressourceneinsatz und unter Nutzung bestehender, und in vielen Fällen öffentlich zugänglicher Daten, wertvolle Erkenntnisse über unterschiedliche Dimensionen im Bereich der KI gewinnen kann. An einem pragmatischen Ansatz für die Indikatoren-Entwicklung führt kaum ein Weg vorbei. Sich von bestehenden Datenquellen leiten zu lassen oder bei bestimmten Fragestellungen auf Umfragen zu setzen, ist vollkommen legitim, solange man sich kritisch mit der Aussagekraft der Datenquellen auseinandersetzt und die Ergebnisse entsprechend einordnet. Indikatoren sind wichtige Hilfsmittel zur Beurteilung von Entwicklungstrends. Sie müssen allerdings auch ständig kritisch hinterfragt werden.

## Ideen für eine empirische Fundierung der deutschen KI-Strategie

Keine der von uns untersuchten nationalen KI-Strategien beschäftigt sich systematisch mit der Definition klarer Ziele und der dafür benötigten Ressourcen und Maßnahmen (Input-Indikatoren). Auch die Frage, wie die Effektivität dieser Maßnahmen (Output-Indikatoren) erfasst und überprüft werden könnte, bleibt unbeantwortet. Hier kann Deutschland sich positiv abheben, wenn die Bundesregierung in ihre KI-Strategie die Entwicklung eines KI-Indexes integriert. Aufbauend auf unserer Analyse in den beiden vorherigen Kapiteln skizzieren wir hier einige eigene Ideen für die Entwicklung eines KI-Indexes. Als Bezugspunkt dienen die von der Stiftung Neue Verantwortung veröffentlichten Eckpunkte einer nationalen Strategie für Künstliche Intelligenz.<sup>42</sup> In diesem Papier schlagen wir vor, dass die KI-Strategie der Bundesregierung die Förderung eines starken KI-Ökosystems in den Mittelpunkt stellen sollte. Ein solches Ökosystem zeichnet sich durch starke Netzwerke und enge Kooperation zwischen Wissenschaft, Wirtschaft, Gesellschaft und Politik aus.<sup>43</sup>

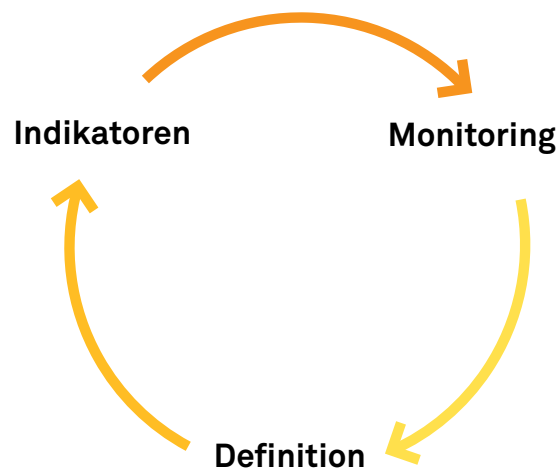
Die Entwicklung von Indikatoren zur Erfassung der unterschiedlichen Dimensionen des KI-Ökosystems stellt allerdings aus unserer Sicht nur eine von drei Aufgabenbereichen dar. Sie würde aber den Kern einer empirischen

---

<sup>42</sup> Harhoff, D. et al. (2018). Eckpunkte einer nationalen Strategie für Künstliche Intelligenz, <https://www.stiftung-nv.de/de/publikation/eckpunkte-einer-nationalen-strategie-fuer-kuenstliche-intelligenz>

<sup>43</sup> Beispiel für den Versuch das kanadische KI-Ökosystem zu erfassen. <http://www.jfgagne.ai/canadian-ai-ecosystem-2018-en>

Fundierung der KI-Strategie ausmachen. Diese empirische Fundierung sollte, wie in Grafik 4 dargestellt, einen fortlaufenden Prozess bilden. Am Anfang müsste eine Auseinandersetzung mit der Definition von KI stehen und insbesondere der Frage, welche Forschungs- und Anwendungsfelder unter diese Definition fallen. Darauf aufbauend können die Ziele der KI-Strategie und die Maßnahmen, die zu ihrer Erreichung ergriffen werden müssen, formuliert werden. Dabei ist, wie bereits weiter oben schon ausführlicher diskutiert, zwischen Input- und Output-Indikatoren zu unterscheiden. Die Input-Indikatoren gehören entweder in die KI-Strategie oder sollten zumindest eng mit ihr verknüpft werden. Sie sollen möglichst konkret die aus der Strategie abgeleiteten Maßnahmen der Bundesregierung benennen sowie die Ressourcen, die für ihre Umsetzung zur Verfügung gestellt werden. Die Output-Indikatoren wiederum beschreiben konkrete Ziele der Strategie. Damit stellen sie die Grundlage für die Definition von Benchmarks dar, deren Erreichung mit Hilfe von Indikatoren überprüfbar werden. Neben der Beschäftigung mit Input- und Output-Indikatoren sollte ein KI-Technologie-Monitoring in den Prozess integriert werden. Dieses Monitoring soll dazu dienen, einen Überblick über wichtige, globale Trends der KI zu behalten und frühzeitig neue, für die KI-Strategie relevante Entwicklungen zu erkennen.



Grafik 4: Empirische Fundierung der KI-Strategie. Quelle: Stiftung Neue Verantwortung

Die drei Bereiche einer empirischen Fundierung einer KI-Strategie greifen ineinander und befruchten sich gegenseitig. Input- und Output-Indikatoren leiten sich aus einer Analyse, was ein starkes KI-Ökosystem auszeichnet, ab. Unter KI werden nicht nur eine Vielzahl an Forschungsansätzen subsumiert. Das Forschungsfeld der KI entwickelt sich auch sehr dynamisch. Das Trend- und Technologie-Monitoring unterstützt die Regierung dabei, neue

Trends frühzeitig zu erkennen. Diese Erkenntnisse sollten in die Weiterentwicklung der KI-Strategie und der mit ihr verbundenen In- und Output-Indikatoren fließen. Mit diesem umfassenden Ansatz kann verhindert werden, dass die KI-Strategie zu einer reinen Absichtserklärung verkommt. Denn die empirische Fundierung würde aus der KI-Strategie einen KI-Strategie-Prozess machen, bei dem nicht nur klare Handlungsmaßnahmen definiert und umgesetzt werden müssen, sondern auch auf neue Entwicklungen reagiert werden kann und anhand von konkret messbaren Ergebnissen die Ergebnisse des Strategieprozesses bewertet und Maßnahmen gegebenenfalls angepasst werden können.

### **Input-Indikatoren**

Input- und Output-Indikatoren lassen sich konzeptionell zwar klar voneinander unterscheiden. In der praktischen Entwicklung bedingen sich beide jedoch gegenseitig und leiten sich aus dem übergeordneten Verständnis des KI-Ökosystems ab. Um die wissenschaftlichen, gesellschaftlichen, wirtschaftlichen und politischen Dimensionen dieses Ökosystems klar zu beschreiben, ist die Auseinandersetzung mit einer Definition von KI und der Operationalisierung des Begriffs in konkreten Forschungs- und Anwendungsbereichen unerlässlich. Aus guten Gründen arbeiten alle KI-Strategien mit einem sehr breit gefassten Verständnis von KI. Schließlich handelt es sich um ein Forschungsfeld, das viele unterschiedliche Disziplinen (Informatik, Kognitionswissenschaften, Statistik, etc.) miteinander vereint und sich sehr dynamisch entwickelt. Allerdings können weder klare Ziele definiert, noch ihre Erreichung überprüft werden, wenn nicht geklärt wird, welche Forschungsansätze und Anwendungen unter KI fallen.

KI-Anwendungen	Beispielanwendungen
<b>Predictive Analytics</b>	<ul style="list-style-type: none"> <li>Überwachung und Wartung der Produktionsanlagen, um auf der Basis von Sensordaten auf kritische Zustände wie Überhitzung einer Produktionsanlage Rückschlüsse ziehen und proaktiv auf mögliche Ausfälle reagieren zu können</li> <li>Beschaffungsplanung unter der Berücksichtigung der Absatzschwankungen</li> </ul>
<b>Optimiertes Ressourcenmanagement</b>	<ul style="list-style-type: none"> <li>Optimierung von Produktions- und Fertigungsplänen</li> <li>Personalplanung</li> <li>Optimierung von Prozessen in der Ein- und Ausgangslogistik</li> </ul>
<b>Qualitätskontrolle</b>	<ul style="list-style-type: none"> <li>Prüfung der Beschaffenheit von Bauteilen oder sonstigen Produktionsstoffen</li> <li>Überprüfung von Korrektheit der Montageprozesse anhand von Video-, Bild- oder auch Sensordaten</li> </ul>
<b>Intelligente Assistenzsysteme</b>	<ul style="list-style-type: none"> <li>Einarbeitung in Verwaltungsprozesse</li> <li>Montageeinleitungen</li> <li>Unterstützung bei Fertigungsprozessen</li> <li>Unterstützung in der Weiterbildung</li> </ul>
<b>Wissensmanagement</b>	<ul style="list-style-type: none"> <li>Management von unternehmensinternen Informationen und Prozessen</li> <li>Datenmodelle für komplexe Engineering-Prozesse</li> <li>Produktkonfigurationen und Beschreibung von Schnittstellen zwischen verschiedenen Bauteilen und Produkten</li> </ul>
<b>Robotik</b>	<ul style="list-style-type: none"> <li>Adaptive, lernende industrielle Robotersysteme in der Produktion und Fertigung</li> <li>Adaptive Service-Roboter</li> <li>Lernende, selbstregulierende Greifsysteme und Montageroboter</li> </ul>
<b>Autonomes Fahren und Fliegen</b>	<ul style="list-style-type: none"> <li>Fahrerlose Transportsysteme wie Reinigungsroboter oder autonom fliegende Drohnen für Bestückung von Regalen in Lagerhallen</li> </ul>
<b>Intelligente Automatisierung</b>	<ul style="list-style-type: none"> <li>Automatisierung von Routineprozessen in Fertigung und Montage durch selbstregulierende Anpassung der Steuerungsparameter</li> <li>Automatisierung von Arbeitsschritten in IT-gestützten Unternehmensprozessen (Robotic Process Automation) inklusive Entscheidungen, die bisher nur von Menschen getroffen wurden, wie E-Mail-Antworten auf Kundenanfragen</li> </ul>
<b>Intelligente Sensorik</b>	<ul style="list-style-type: none"> <li>Umgebungswahrnehmung (Bild, Laserscan) und Vorverarbeitung der Daten für Kollisionsvermeidung der fahrerlosen Transportsysteme</li> <li>Vorverarbeitung der Daten beim Monitoring von Produktionsanlagen</li> </ul>

Tabelle 2: KI-Anwendungen und Beispiele (Quelle: itt)

Beispielhaft sei hier die vom Institut für Innovation und Technik (itt) entwickelte Übersicht über KI-Anwendungen angeführt.<sup>44</sup> In einer Studie für das Bundeswirtschaftsministerium hat sich das itt mit den Potenzialen von KI für das produzierende Gewerbe befasst. Hierzu musste bestimmt werden, welche KI-Technologien Gegenstand der Studie sind. In der Tabelle geben die Autoren der Studie einen Überblick und führen die wichtigsten KI-Technologien auf. Bei jeder Technologien werden zugehörigen Verfahren und Methodik kurz beschrieben. Eine solche Operationalisierung des KI-Begriffs ist auch für die Entwicklung der Input- und Output-Indikatoren notwendig. Ansonsten bleibt unklar, auf was sie sich konkret beziehen.

Die Tabelle unten führt erste Ideen für Input- und Output-Indikatoren für die einzelnen Dimensionen des KI-Ökosystems auf. In Bezug auf die Forschung sollte die KI-Strategie Fördersummen benennen und beschreiben, wie För-

<sup>44</sup> Potenziale der Künstlichen Intelligenz im produzierenden Gewerbe in Deutschland, Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie (BMWi), [https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/potenziale-kuenstlichen-intelligenz-im-produzierenden-gewerbe-in-deutschland.pdf?\\_\\_blob=publicationFile&v=16](https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/potenziale-kuenstlichen-intelligenz-im-produzierenden-gewerbe-in-deutschland.pdf?__blob=publicationFile&v=16) S. 14.



derprogramme agiler, weniger bürokratisch und wettbewerbsorientierter aufgesetzt werden könnten. Es sollten möglichst konkrete Maßnahmen zur Vermittlung von KI-Kompetenzen über die gesamte Bildungskette hinweg aufgeführt werden. Input-Indikatoren bezüglich der Verbesserung der Bereitstellung von Daten betreffen sowohl den weiteren Ausbau des Open-Data-Ansatzes wie auch Maßnahmen, die das Teilen von Daten über unterschiedliche Unternehmen und Sektoren hinweg ermöglichen könnten. Bei der KI-Infrastruktur geht es vor allem um die Einführung eines Monitoring, um frühzeitig mögliche Probleme beim Zugang zu wichtiger KI-Hardware erkennen und adressieren zu können. Input-Indikatoren zur Stärkung der Wirtschaft im KI-Ökosystem könnten zum Beispiel spezielle Förderprogramme für KMU oder eine Strategie für den Einkauf von KI-Systemen durch die öffentliche Hand sein. Input-Indikatoren zur gesellschaftlichen Dimension könnten sich auf die Stärkung des Dialogs mit der Zivilgesellschaft und die Förderung der Forschung zu ethischen Fragen und den Auswirkungen von KI auf Arbeit und Beschäftigung beziehen. Auf der internationalen Ebene stellen der Aufbau länderübergreifender Forschungskooperation und außenpolitische Initiativen zur Entwicklung internationaler Normen wichtige Input-Indikatoren dar.

Themenfeld	Input-Indikatoren	Output-Indikatoren
Forschung	<p>Aufgeschlüsselt nach Anwendungs- &amp; Spitzenforschung</p> <ul style="list-style-type: none"> <li>- Finanzielle Ressourcen</li> <li>- Agilere Förderprogramme</li> </ul>	<p>Aufgeschlüsselt nach Anwendungs- &amp; Spitzenforschung</p> <ul style="list-style-type: none"> <li>- Anzahl der Wissenschaftler in der KI-Forschung</li> <li>- Anzahl der Doktoranden bzw. Promotionen</li> <li>- Anzahl an Konferenzteilnahmen, Publikationen und Zitations</li> <li>- Attraktivität des Forschungsstandortes</li> <li>- Agilität der Forschungspolitik</li> </ul>
Kompetenzen	<ul style="list-style-type: none"> <li>- Vermittlung von Basiswissen in Schulen</li> <li>- KI-relevante Anteile in Studienfächern</li> <li>- KI-relevante Anteile in Weiterbildungen</li> </ul>	<ul style="list-style-type: none"> <li>- Verbreitung von KI-Kompetenzen ausserhalb der Spitzenforschung</li> <li>- z.B. an Fachhochschulen und in Unternehmen</li> <li>- KI-Kompetenzen von Absolventen von MINT Studienfächern</li> </ul>
Daten	<ul style="list-style-type: none"> <li>- Zugang zu Daten (Staat: Open Data / Wirtschaft: Daten Pools) in maschinenlesbaren Formaten</li> <li>- Verbesserung von Datenqualität und Depersonalisierungsverfahren</li> <li>- Standardisierung von Datenformaten</li> </ul>	<ul style="list-style-type: none"> <li>- Fortschritte Deutschlands in internationalen Open Data Indizes</li> <li>- Verbreitung und Nutzung von Datenpools</li> </ul>
Infrastruktur	<ul style="list-style-type: none"> <li>- KI-Hardware Technologie- Monitoring</li> <li>- Förderung strategisch wichtiger KI-Hardware Entwicklung</li> <li>- Zugänglichkeit von mit KI verknüpften Cloud-Infrastrukturen</li> </ul>	<ul style="list-style-type: none"> <li>- Erkenntnisse aus dem Hardware Technologie-Monitoring</li> <li>- Erfassung der wichtigsten Hardware-Anbieter und ihrer Marktanteile</li> <li>- Marktanteile in Deutschland entwickelter KI-Hardware</li> </ul>
Wirtschaft	<ul style="list-style-type: none"> <li>- Förderung von Programmen speziell für KMU</li> <li>- Staatliche auf KI-fokussierte Wagniskapitalfonds bzw. Maßnahmen zur Förderung privater KI-Investitionen</li> <li>- Staatliches Beschaffungswesen</li> </ul>	<ul style="list-style-type: none"> <li>- Höhe der KI-Investitionen in der Wirtschaft (sowohl Investitionen in eigene Projekte / Kompetenzen wie auch der Aufkauf von KI-Startups)</li> <li>- Grad des Einsatzes von KI in unterschiedlichen Branchen nach Unternehmensgröße</li> <li>- Zahl der angemeldeten KI-Patente</li> <li>- Zahl der geförderten Experimentierräume und Labs und der mit den Maßnahmen erreichten Unternehmen</li> <li>- Zahl und Performanz von KI-Startups (Erfassung von Gründungen innerhalb und außerhalb von Universitäten)</li> <li>- Wagniskapital-Investitionen in KI</li> <li>- Zahl von auf KI fokussierten Acceleratorenprogramme und Innovation-Labs</li> </ul>

Themenfeld	Input-Indikatoren	Output-Indikatoren
Gesellschaft	<ul style="list-style-type: none"> <li>- Einbindung von Zivilgesellschaft in die politische Debatte</li> <li>- Forschung zu ethischen Prinzipien und Regulierung fördern und ggf. gesetzgeberisch umsetzen</li> <li>- Auswirkungen von KI auf Beschäftigung untersuchen</li> </ul>	<ul style="list-style-type: none"> <li>- Austausch mit Zivilgesellschaft z.B. über Einladungen zu Konferenzen, Workshops und Expertengremien</li> <li>- Förderung von Forschung an ethischen und sozialen Fragen rund um KI</li> <li>- Öffentlichkeitsarbeit rund um KI bzw. Umfragen zum Kenntnisstand in der Bevölkerung</li> <li>- Untersuchungen zu Auswirkungen auf Beschäftigung im KI-Bereich</li> </ul>
International	<ul style="list-style-type: none"> <li>- Kooperationen mit anderen Ländern und EU vorantreiben (z.B. über gemeinsame Forschungsprojekte)</li> <li>- Mitgestaltung internationaler Normen zu KI</li> <li>- KI auf die außenpolitische Agenda setzen</li> </ul>	<ul style="list-style-type: none"> <li>- Anzahl internationaler Forschungsk Kooperationen</li> <li>- Strategische KI-Investitionen deutscher Unternehmen im Ausland</li> <li>- Aktivitäten des Auswärtigen Amts zu zentralen internationalen Debatten rund um KI</li> <li>- Fortschritte bei der Entwicklung und Umsetzung KI-Strategie der EU</li> </ul>

Tabelle 7: Vorschläge zur Operationalisierung einzelner Indikatoren. Quelle: Stiftung Neue Verantwortung

## Output-Indikatoren

Die Output-Indikatoren liefern Hinweise, wie es um die einzelnen Dimensionen des Ökosystems bestellt ist. Werden die Output-Indikatoren regelmäßig über einen längeren Zeitraum hinweg erhoben, sind auch Aussagen über die Entwicklung des KI-Ökosystems möglich. Während die Input-Indikatoren von der Regierung relativ klar definiert werden können und erstmal keine eigene Datenerhebung notwendig machen, spielen bei der Entwicklung der Output-Indikatoren Fragen nach der Methodik, den Datenquellen und der Aussagekraft eine wichtige Rolle. Hier müssen vor allem die im vorherigen Kapitel zu KI-Indizes herausgearbeiteten Herausforderungen und Probleme berücksichtigt werden, um brauchbare Output-Indikatoren zu entwickeln und richtig zu interpretieren. Ein kritischer Umgang mit der Methodik ist notwendig, um die Ergebnisse ständig neu bewerten und hinterfragen zu können. Indikatoren geben uns Hinweise auf ein Lagebild und wichtige Entwicklungen. Aufgrund der bereits angesprochenen Probleme und Beschränkungen darf ihnen allerdings auf keinen Fall blind vertraut und gefolgt werden.

Die Messung von Forschungs-Output ist in der Wissenschaft umstritten. Dennoch gibt es wichtige Kennziffern, um Umfang und Qualität von Forschungsaktivitäten statistisch zu erheben und zu bewerten. Einen ersten Anhaltspunkt bietet die Anzahl der in der KI-Forschung tätigen Personen. Den Output der Wissenschaftler:innen könnte man auf mehreren Ebenen bewerten: Anzahl der Doktorand:innen beziehungsweise erfolgreich betreuter Promotionen, Anzahl der Präsentationen auf führenden KI-Konferenzen und Anzahl und Sichtbarkeit (zum Beispiel über Zahl und Qualität der Zitationen) der Forschungspublikationen. Die Attraktivität des Forschungsstandortes könnte man über die Anzahl der von Top-Instituten und Universitäten aus dem Ausland angeworbenen Wissenschaftler:innen oder Umfragen unter KI-Expert:innen erfassen. Ebenso könnte man über Umfragen, Ausrichtung und Durchführung der Antragsverfahren von Forschungsförderungsprogrammen untersuchen und bewerten.

Eine wichtige in unserem Strategiepapier beschriebene Herausforderung ist es, KI-Kompetenzen nicht nur in der Spitzenforschung zu fördern, sondern auch in die Breite zu tragen. Hier könnte man zum Beispiel erfassen, welche Fachhochschulen für KI relevante Kompetenzen vermitteln und wie viele Studierende damit erreicht werden. Auch die Vermittlung von KI relevanten Kompetenzen in MINT-Studienfächern sollte erfasst werden. Um Fortschritte bei Open Data zu messen, müssten nicht unbedingt neue Daten erhoben werden. Stattdessen könnte man Deutschlands Performanz aus



bestehenden Indizes wie zum Beispiel dem Global Open Data Index der Open Knowledge Foundation ableiten.<sup>45</sup> In Bezug auf die Verfügbarkeit von Daten im Privatsektor wäre auch eine Erfassung und Analyse von von unterschiedlichen Unternehmen gemeinsam initiierten und genutzten Daten-Pools hilfreich. Im Bereich Hardware sollte bei Output-Indikatoren der Aufbau eines KI-Hardware-Technologie-Monitoring bewertet werden. Eine erfolgreiche Umsetzung würde Daten über die wichtigsten Hardware-Anbieter und ihrer Marktanteile liefern. Hieraus könnten Schlüsse gezogen werden, wie groß Abhängigkeiten und Risiken bezüglich Zugang zu kritischen Hardware Komponenten sind.

Die Entwicklung und Einsatz von KI in der deutschen Wirtschaft könnte über eine ganze Reihe von Indikatoren hinweg erfasst und nachvollzogen werden. Begrenzungen ergeben sich hier vor allem durch den Aufwand der Datenerhebung und Auswertung. So sind nicht nur die Höhe staatlicher Ausgaben zur Förderung von KI relevant, sondern auch die von deutschen Unternehmen aufgebrachten Investitionen. Der Grad der Entwicklung und des Einsatzes von KI in Unternehmen könnte wahrscheinlich am leichtesten über Umfragen erfasst werden. Hier wäre auch die Erhebung wichtig, wie viele Personen in einem Unternehmen mit KI-relevanten Tätigkeiten befasst sind. Umfrageergebnisse zum Umfang und Akzeptanz möglicher staatlicher Förderprogramme, wie zum Beispiel den vorgeschlagenen Labs und Experimentierräumen für KMUs, wären ebenfalls wichtig. Auch die Aktivitäten im Startup-Bereich müssten systematisch erfasst werden. Neben der Zahl von Neugründungen, Exits und der Größe der Investments wären auch die Größe (Mitarbeiter:innenzahl) und gegebenenfalls internationale Verbindungen (zum Beispiel über Investoren oder Marktaktivitäten) interessant.

Die gesellschaftliche und die internationale Dimension sollten auch mit Output-Indikatoren in den Blick genommen werden. Beide Dimensionen geraten bei der Diskussion um nationale KI-Strategien bei der üblichen Fokussierung auf Forschungsförderung und Industriepolitik schnell ins Hintertreffen. Für den Erfolg der KI-Strategie sind sie aber ebenfalls von großer Bedeutung. So sollten von Indikatoren die Einbindung von Zivilgesellschaft in unterschiedliche Diskussionsformate der Bundesregierung zu KI systematisch erfasst werden (zum Beispiel über statistische Erfassung der Einbindung zivilgesellschaftlicher Organisationen in Dialog- und Austauschformate). Darüber hinaus wären konkrete Benchmarks bei der Förderung von auf die ethischen und sozialen Aspekte von KI fokussierten Forschungsprojekte

---

<sup>45</sup> The Global Open Data Index (2017). Tracking the state of open government data, <https://index.okfn.org/>

wichtig. Über große repräsentative Umfragen könnte Wissen, Hoffnungen und Sorgen der Bürger:innen rund um diese neue Technologie erfasst werden. Auf der internationalen Ebene sollte es vor allem um die Erfassung der grenzüberschreitende Vernetzung von deutscher Forschung und Wirtschaft gehen. Außerdem könnten die Aktivitäten des Auswärtigen Amts bei der Erarbeitung internationaler Normen erfasst werden. Und schließlich wird der Erfolg der deutschen KI-Strategie davon abhängen, wie gut eine Integration der nationalen Ansätze in eine EU-weite Strategie gelingt. Hier ließen sich Fortschritte über den Stand der Planung, Finanzierung und Umsetzung konkreter EU-weiter Maßnahmen und Förderprogramme erfassen.

Die Entwicklung von Output-Indikatoren sollte mit einer grundsätzlichen Auseinandersetzung mit der Datengrundlage verknüpft werden. Denn der deutsche Staat erhebt bereits viele relevante Daten und Statistiken.<sup>46</sup> Anstatt völlig neue Verfahren zur Datenerhebung aufzusetzen, könnten in vielen Fällen sicherlich durch eine Anpassung bestehender statistischer Erhebungen und die Integration von für die KI-Strategie relevante Kategorien schnell große Fortschritte erzielt werden.

### **KI-Technologie- und Trend-Monitoring**

Die Input- und Output-Indikatoren sollten zusätzlich von einem KI-Technologie und Trend-Monitoring flankiert werden. Die Einführung eines Technologie-Monitoring ist auch im Eckpunkte-Papier der Bundesregierung erwähnt, ohne, dass die Idee genauer beschrieben wird. Ein solches Monitoring ist sehr sinnvoll. Denn sowohl in der Forschung als auch in der Anwendung ist KI nicht nur ein sehr breites Feld sondern auch ein sehr dynamisches. Neue Forschungsansätze haben sich in den letzten Jahren sehr schnell etabliert und auch bei den Anwendungsfeldern sieht man viel Bewegung. Daher macht eine möglichst offene und breite Definition von KI Sinn. Allerdings muss bei der Operationalisierung genauer spezifiziert werden, welche Forschungsansätze und Anwendungsfelder unter KI fallen und daher erfasst werden sollen.

Ein kontinuierliches Trend-Monitoring dient daher nicht nur dazu, neue Entwicklung frühzeitig zu erkennen, sondern kann auch genutzt werden, um die Operationalisierung der Indikatoren, zum Beispiel durch Hinzufügen neuer Forschungsfelder, aktuell und relevant zu halten. Als Grundlage könnte eine

---

<sup>46</sup> Die Bundesregierung (2018). Förderung von Forschung und Entwicklung im Bereich der Informations- und Kommunikationstechnologie, <https://dipbt.bundestag.de/dip21/btd/19/038/1903831.pdf>

Auswertung bereits bestehender Trend-Reports, wie dem oben vorgestellten *CBInsights* Trend Report oder der State of AI Präsentation dienen. Falls es die Ressourcen zulassen, könnte eine Auswertung von bereits bestehenden Reports durch eigene Datenerhebungen und Analysen erweitert werden. So könnte eine eigene Datenerhebung und Auswertung der wichtigsten Performanz-Wettbewerbe und Challenges in Bereich der KI wichtige Aufschlüsse über die Forschungsentwicklung liefern. Auch die Implementierung von KI-Applikationen in unterschiedlichen Branchen und die damit verbundenen Einflüsse auf Wertschöpfungsketten sind wichtige Themenfelder für das KI-Trendmonitoring. Zusätzlich bietet eine eigene Umfrage mit national und international führenden KI-Expert:innen die Möglichkeit, neue Erkenntnisse zu den neuesten Trends und Entwicklungen liefern.

## Ausblick

Mit der Veröffentlichung ihrer Eckpunkte für eine nationale KI-Strategie folgt die Bundesregierung dem Beispiel zahlreicher Länder, die bereits nationale KI-Strategien veröffentlicht haben. Unsere Analyse zeigt, dass die Strategien zwar ambitioniert, aber vor allem auch sehr unkonkret sind. Es fehlen klar definierte Ziele und daraus abgeleitet In- und Output-Indikatoren, welche die zu ergreifenden Maßnahmen und die hierfür zur Verfügung gestellten Ressourcen sowie konkrete Benchmarks, die es zu erreichen gilt, festlegen. Hier bietet sich die Chance für Deutschland, es besser zu machen. Die Fokussierung der Diskussion auf die Umsetzung klar definierter Ziele und der dazu notwendigen Erfolgskriterien kann dem bisher sehr breiten und unkonkret geführten gesellschaftlichen Diskurs nur guttun. Die Analyse bestehender KI-Indizes zeigt die Hürden und Herausforderungen auf, die es dabei zu meistern gilt.

Angesichts der zu erwartenden Höhe der öffentlichen Ausgaben für die Umsetzung der KI-Strategie, sind die für den Aufbau eines Monitoring und Benchmarking benötigten Haushaltsmittel vergleichsweise gering. Es wäre aber sehr gut angelegtes Geld. Denn nur wenn die Entwicklung eines KI-Ökosystems in Deutschland mit Indikatoren erfasst wird, gibt es Anhaltspunkte zur Bewertung der Effektivität der im Rahmen der Strategie entwickelten Maßnahmen. Letztendlich ist eine agile und zielorientierte Umsetzung der KI-Strategie ohne eine empirische Fundierung nicht möglich. Die Bundesregierung sollte für den Aufbau eines solchen Mess-Ansatzes ein unabhängiges Forschungsinstitut beauftragen. Auf Basis wissenschaftlicher Standards und hoher Methodenkompetenz würde das beauftragte Forschungsinstitut vor allem die Output-Indikatoren entwickeln

und erfassen und den KI-Trend Report erstellen. Das Forschungsinstitut sollte dabei völlig frei und unabhängig von politischen Einflüssen agieren können. Die Input-Indikatoren zu definieren ist hingegen eine politische Aufgabe. Schließlich können nur Regierung und Parlament über die Vergabe von Haushaltsmitteln und die Entwicklung und Umsetzung politischer Maßnahmen entscheiden.

Die von einem beauftragten Forschungsinstitut regelmäßig herauszugebenden Reports mit Angaben zu den eigens festgelegten Indikatoren bezüglich der Entwicklung des deutschen KI-Ökosystems und zu globalen KI-Trends würde die gesellschaftspolitische Diskussion über KI in Deutschland auf eine empirisch unterlegte Grundlage stellen. Zusätzlich würden die Reports für die Bundesregierung eine wichtige Ressource für die erfolgreiche Umsetzung und kontinuierliche Weiterentwicklung der nationalen KI-Strategie darstellen. Natürlich müsste sich die Bundesregierung auch an den Reports messen lassen. Das ist nicht ganz ohne politisches Risiko. Die Bereitschaft, dieses Risiko einzugehen würde allerdings ein sehr starkes Signal senden: nämlich, dass es die Bundesregierung mit der Umsetzung der KI-Strategie wirklich ernst meint.



**Dr. Stefan Heumann mit Nicolas Zahn**

**September 2018**

**Erfolgsmessung von KI-Strategien**

## **Über die Stiftung Neue Verantwortung**

Think Tank für die Gesellschaft im technologischen Wandel

Neue Technologien verändern Gesellschaft. Dafür brauchen wir rechtzeitig politische Antworten. Die Stiftung Neue Verantwortung ist eine unabhängige Denkfabrik, in der konkrete Ideen für die aktuellen Herausforderungen des technologischen Wandels entstehen. Um Politik mit Vorschlägen zu unterstützen, führen unsere Expertinnen und Experten Wissen aus Wirtschaft, Wissenschaft, Verwaltung und Zivilgesellschaft zusammen und prüfen Ideen radikal.

## **Über die Autoren**

[Dr. Stefan Heumann](#) ist Politikwissenschaftler und Mitglied des Vorstands der Stiftung Neue Verantwortung. Er ist Sachverständiger der KI-Enquete des Deutschen Bundestages.

Nicolas Zahn ist Politikwissenschaftler und war im Juli 2018 als Fellow des Mercator Kolleg für internationale Aufgaben bei der SNV.

## **So erreichen Sie die Autoren**

Dr. Stefan Heumann

Mitglied des Vorstands der Stiftung Neue Verantwortung

[sheumann@stiftung-nv.de](mailto:sheumann@stiftung-nv.de)

+49 (0)30 81 45 03 78 80



## Impressum

Stiftung Neue Verantwortung e. V.

Beisheim Center  
Berliner Freiheit 2  
10785 Berlin

T: +49 (0) 30 81 45 03 78 80

F: +49 (0) 30 81 45 03 78 97

[www.stiftung-nv.de](http://www.stiftung-nv.de)

[info@stiftung-nv.de](mailto:info@stiftung-nv.de)

Design:

Make Studio

[www.make-studio.net](http://www.make-studio.net)

Layout:

Johanna Famulok

Free Download:

[www.stiftung-nv.de](http://www.stiftung-nv.de)



Dieser Beitrag unterliegt einer CreativeCommons-Lizenz (CC BY-SA). Die Vervielfältigung, Verbreitung und Veröffentlichung, Veränderung oder Übersetzung von Inhalten der Stiftung Neue Verantwortung, die mit der Lizenz „CC BY-SA“ gekennzeichnet sind, sowie die Erstellung daraus abgeleiteter Produkte sind unter den Bedingungen „Namensnennung“ und „Weiterverwendung unter gleicher Lizenz“ gestattet. Ausführliche Informationen zu den Lizenzbedingungen finden Sie hier:

<http://creativecommons.org/licenses/by-sa/4.0/>