



AI-JAHRESRÜCKBLICK 2020

Werner Bogula, Dezember 2020

AI-Jahresrückblick 2020

Das Jahr 2020 wurde überschattet von der Corona-Krise. Hinter den unser Leben umkrepelnden Maßnahmen der Virenbekämpfung liefen Forschung und Entwicklung in der AI weiter auf Hochtouren. In diesem kurzen Essay soll ein Überblick über einige der wichtigen und prägenden AI-Entwicklungen des Jahres 2020 gegeben werden. Auch wenn einige Einzelprojekte herausgestellt werden, soll es dabei um die großen aktuelle Entwicklungsstränge gehen.

1 Forschung und Entwicklung

Neue Transformer-Modelle des sprachlichen Verstehens

Im Bereich der Sprachverarbeitung gab es im Jahr 2020 einige Furore machende Entwicklungen. An erster Stelle dürfte dabei das aufsehenerregende Sprachmodell GPT-3¹ stehen. GPT-3 ist eine umfangreichere Version des GPT-2-Modells von Open AI und war mit 175 Milliarden trainierbaren Parametern das bei seinem Erscheinen umfangreichste Sprachmodell.

GPT-3 ist ein sogenanntes Transformer-Sprachmodell, das Texte generieren kann. Auf der Basis von gewaltigen Textmengen sind in diesem Modell statistische Zusammenhänge über das gemeinsame Auftreten von Wörtern errechnet worden. Die Leitidee dabei ist, dass sich allein durch die Menge der verarbeiteten Daten und eine spezielle Aufmerksamkeitsarchitektur² in dem Modell textliche aber auch logische Zusammenhänge abbilden lassen, die für die Textgenerierung genutzt werden können. Die zunächst ziemlich simpel klingende Aufgabe von GPT-3, nämlich bei vorgegebenen Text das jeweils nächste Wort vorherzusagen, führt aufgrund der immensen Anzahl an gespeicherten Zusammenhängen in GPT-3 dazu, dass Texte erzeugt werden können, die z.T. von menschlichen sprachlichen Leistungen nicht mehr unterschieden werden können.

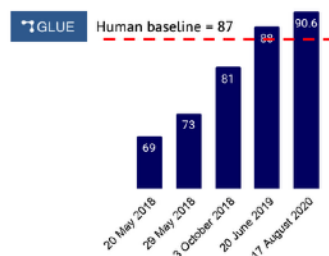


Abb. 1 - Sprachmodelle übersteigen die Fähigkeiten von Menschen bei synthetischen Sprachperformance-Tests um einige Punkte. (Quelle: stateof.ai, 27)

Die Kosten für die Erstellung des Modells werden auf mehrere Millionen US-Dollar geschätzt, und zur Monetarisierung bietet die Entwicklungsfirma Open AI, den Access zu diesem Modell entgegen der ursprünglichen Idee einer Open-Source-AI als kommerzielles Abonnement an. Microsoft hat sich die Exklusivrechte auf den Sourcecode dieses Modells gesichert³. Man kann sich gut vorstellen, dass eine das Textverständnis einbeziehende Rechtschreibprüfung oder eine in Word eingebaute Sprach-KI als ein ertragreiches potenzielles Abonnement für die Abermillionen Word-User angesehen wird.

¹ <https://openai.com/blog/openai-api/>

² <https://towardsdatascience.com/attention-is-all-you-need-discovering-the-transformer-paper-73e5ff5e0634>

³ <https://www.technologyreview.com/2020/09/23/1008729/openai-is-giving-microsoft-exclusive-access-to-its-gpt-3-language-model/>

Insgesamt setzt sich der Trend zu großen Sprachmodellen fort. Gerade im Bereich des Sprachverstehens sind große Schätze zu heben, da der überwiegende Teil der menschlichen Kommunikation durch Sprache erfolgt. So laufen bei Google in 2020 so ziemlich alle Suchanfragen über das Sprachmodell BERT⁴, das mit einer ähnlichen Architektur wie GPT-3 bereits 2018 vorgestellt wurde. Die Hoffnung, vom Natural Language Processing (NLP) zu einem Natural Language Understanding (NLU) zu gelangen, beflügelt neue Anwendungsideen wie: bessere Recherche in Texten, automatisches Routing von Emails und Bearbeitung von Kundenanfragen. Die Modelle BART, mT5⁵ und GPT ermöglichen NLU-Anwendungsfälle wie verbesserte Text-Zusammenfassungen, Übersetzungen und Textgenerierung. Die Idee des Transfer-Learnings trägt dazu bei, dass die zugrundeliegenden Transformer-Architekturen über die Sprachverarbeitung hinaus in der Bilderkennung ein neues Paradigma etablieren⁶. Allerdings hat mittlerweile bei den AI-Giganten ein Wetttrüsten eingesetzt, mit immer größeren und teureren Sprachmodellen die Vorherrschaft beim NLU zu erlangen, und damit Standards zu schaffen, die weniger kapitalkräftige Wettbewerber nicht mehr einholen können. Aus der AI-Ethik gibt es Einwände, die auf die Transparenz, den Bias und die Klimateffizienz der Riesenmodelle abzielen. (s.u.)

Als Gegenbewegung zu dieser Entwicklung schälen sich drei neue Entwicklungen heraus:

- a) die Entwicklung kleinerer, spezialisierter Modelle, die in bestimmten Bereichen mit den Riesenmodellen bei den Kosten und der Performanz konkurrieren können⁷
- b) der Einbezug von außersprachlichem Wissen⁸ und kausalem Denken, basierend auf der Beobachtung, dass die reine Sprachstatistik der Riesenmodelle inhaltliche Zusammenhänge in Texte nicht immer logisch und richtig abbilden kann.⁹
- c) Comeback der symbolischen AI und Hybridformen aus symbolischen und neuronalen Architekturen. Seit der von Gary Marcus aufgeworfenen Kontroverse¹⁰ um die Bedeutung der symbolischen Verarbeitung (dazu auch Kahneman¹¹) erweitert sich die Forschungsperspektive über mathematische Modelle hinaus auf den Einbezug von Erkenntnissen aus Psychologie und Kognitionswissenschaften und führt zu neuen hybriden Architekturen¹². Insgesamt ist Sprache derzeit im Fokus der Forschung¹³, da mit dem Erschließen weiterer kognitiver menschlicher Leistungen die Tür zu weiteren Anwendungsfällen aufgestoßen wird, in denen das Verstehen von Informationen eine Rolle spielt.

⁴"Google: BERT now used on almost every English query". Search Engine Land. 2020-10-15. Retrieved 2020-11-24.

⁵ <https://arxiv.org/pdf/2010.11934.pdf>

⁶ <https://openreview.net/pdf?id=YicbFdNTTy>

⁷ <https://www.polyai.com/polyais-convert-model-outperforms-bert-and-gpt-based-models/>

⁸ <http://www.ifis.cs.tu-bs.de/node/3246>

⁹ <https://digitalminds2016.wordpress.com/category/turing-test/>

¹⁰ <https://montrealartificialintelligence.com/aidebate/>

¹¹ https://en.wikipedia.org/wiki/Thinking,_Fast_and_Slow

¹² <https://thenextweb.com/podium/2020/01/15/what-are-neural-symbolic-ai-methods-and-why-will-they-dominate-2020/>

¹³ <https://hbr.org/2020/09/the-next-big-breakthrough-in-ai-will-be-around-language>

2 AI in Biologie und Medizin

In den Life Sciences werden seit Jahren große Hoffnungen auf die Massenverarbeitung von medizinischen Daten aber auch auf das AI-gestützte Research gesetzt. Laut stateof.ai 2020 erfuhr die Biologie in 2020 ihren „AI moment“¹⁴, belegt durch über 21.000 Papers mit AI Fokus in der National Library of Medicine¹⁵.

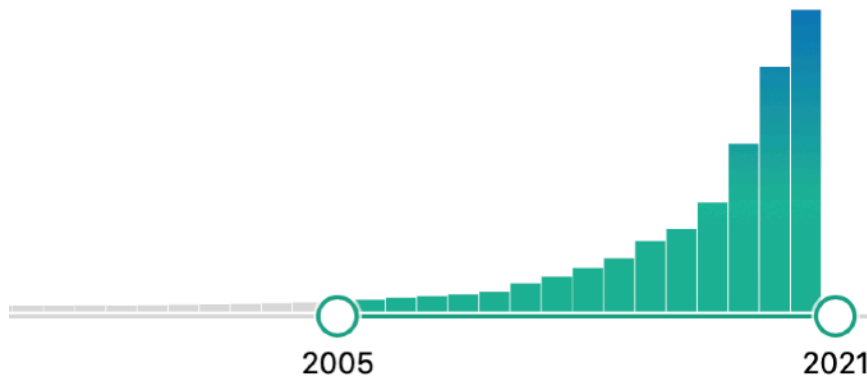


Abb. 2 - Zuwachs an Papers in der *Library of Medicine* mit dem Schlagwort „Machine Learning“ (Quelle: pubmed.gov)

Google Alpha Fold 2

Als größter Durchbruch in den Life Sciences wurde im Dezember 2020 Google Alpha Fold 2 gefeiert. GAF-2 löst eins der schwierigsten Probleme in der Life Science Forschung: das Proteine Folding. Hierbei geht es um die Transformation der Abfolge von Protein-Bausteinen in eine dreidimensionale Faltungsstruktur. Diese Anordnung im Raum determiniert die Funktion der Proteine in ihrer biochemischen Umgebung und konnte bisher nur für einen kleinen Ausschnitt der möglichen Proteine und ihrer Molekülstruktur experimentell erforscht werden. Der riesige Ergebnisraum von geschätzten 100 hoch 300 möglichen Lösungen macht eine brute force Bestimmung der Faltungsstrukturen unmöglich. Das federführende Unternehmen Deepmind und ihr charismatischer Leader Demis Hassabis¹⁶ machen in der Tradition von Alpha Go (dem Programm, das eine Winning-Strategie für den Jungle der 100 hoch 360 möglichen Go-Züge umsetzte) das Versprechen wahr, AI für wirklich harte wissenschaftliche Probleme der Menschheit anzuwenden.

Die herausragende Entwicklung von GAF2 lenkt den Blick auf die Life Science Industries als neues Hot Bed für AI. Gerade in dieser von extrem teurerer Forschung und langwierigen und kostspieligen Experimenten geprägten Industrie ermöglicht AI völlig neue Perspektiven, in denen gewaltige Effizienzsteigerungen in eine neue Qualität der Forschung umschlagen.

„Combining ML with carefully designed experiments has enabled LabGenius to increase the number of potential drug candidates by up to 100,000 fold.“ (stateof.ai 2020, p87)

¹⁴ „**Biology is experiencing its “AI moment”**: From medical imaging, genetics, proteomics, chemistry to drug discovery.“ stateof.ai 2020, 30

¹⁵ <https://pubmed.ncbi.nlm.nih.gov/?term=machine+learning&filter=years.2005-2021&timeline=expanded&sort=date>

¹⁶ <https://www.ft.com/content/048f418c-2487-11e7-a34a-538b4cb30025>

Keine Wissenschaft dürfte aufgrund der Komplexität der untersuchten Systeme, also lebender Organismen, so sehr vom Erkennen von Korrelationen der Ausprägungen bestimmter Features und dahinter liegender multikausaler Zusammenhänge profitieren wie die Life Sciences. Das beginnt bei der Erforschung individueller Reaktionen des Metabolismus auf unterschiedliche Umwelteinflüsse z.B. durch das Start-up Zoe¹⁷ und endet in der Entwicklung personalisierter Wirkstoffe, die für genau einen Patienten entwickelt werden¹⁸.



Abb. 3 - Die Modellierung multifaktorieller Einflüsse auf den Organismus sind perfekt für ML Modelle, (Quelle joinzoe.com)

Ein Hinweis, zu welchen Leistungen ein relativ simples ML-Modell in der Lage ist, gibt die extrem hohe Erkennungsrate eines Modells¹⁹ zur Prognose einer vorliegenden Covid-19 Infektion aus der Tonaufnahme von Hustengeräuschen oder auch die Covid19-Erkennung anhand von Sprachdaten²⁰.

Die hohe Verantwortung, die mit solchen Verfahren für die Gesundheit von Menschen entsteht, stellt auch neue Anforderungen an die Sicherheit und Erklärbarkeit der zugrundeliegenden AI-Modelle. Die Themen *Explainability* und *Causal Reasoning* geraten im medizinischen Umwelt in den Fokus. Denn erst eine klare Unterscheidung der reinen Korrelation von Features von der kausalen Bedingtheit kann erklärende Wirkung²¹ entfalten.

¹⁷ <https://joinzoe.com>

¹⁸ <https://www.nytimes.com/2019/10/09/health/mila-makovec-drug.html>

¹⁹ <https://news.mit.edu/2020/covid-19-cough-cellphone-detection-1029>

²⁰ <https://www.uni-augsburg.de/de/campusleben/neuigkeiten/2020/11/09/3204/>

²¹ <https://faculty.ai/blog/neurips-building-more-explainable-ai-through-asymmetric-shapley-values/>

3 AI Hardware

Die immer ausgefeilteren neuen AI-Modelle (speziell die oben erwähnten Transformer) benötigen immer leistungsfähigere Hardware für das Training und die Anwendung (Inferenz); sowohl beim möglichst nah an die Prozessorhardware angebondenen Speicher als auch bei Anzahl und Leistung der parallel arbeitenden GPUs. Der führenden Hersteller NVIDIA hat 2020 mit seiner Ampere-Reihe²² neue Leistungsrekorde aufgestellt und hat sich vom Anbieter für Spiele-Grafikkarten zum weltweit tätigen Ausrüster von Rechenzentren weiterentwickelt.

Auch beim Edge-Computing bietet NVIDIA mit den Jetson Nanos erschwingliche Hardware im Raspberry Pi Segment zwischen 60 und 100 Dollar an und liefert damit vom Rechenzentrum, über den PC bis zum IoT-Bereich GPU-gestützte Hardware. Im Highend konnte 2020 einzig Graphcore mit der M2000 eine noch leistungsfähigere Architektur vorlegen.²³

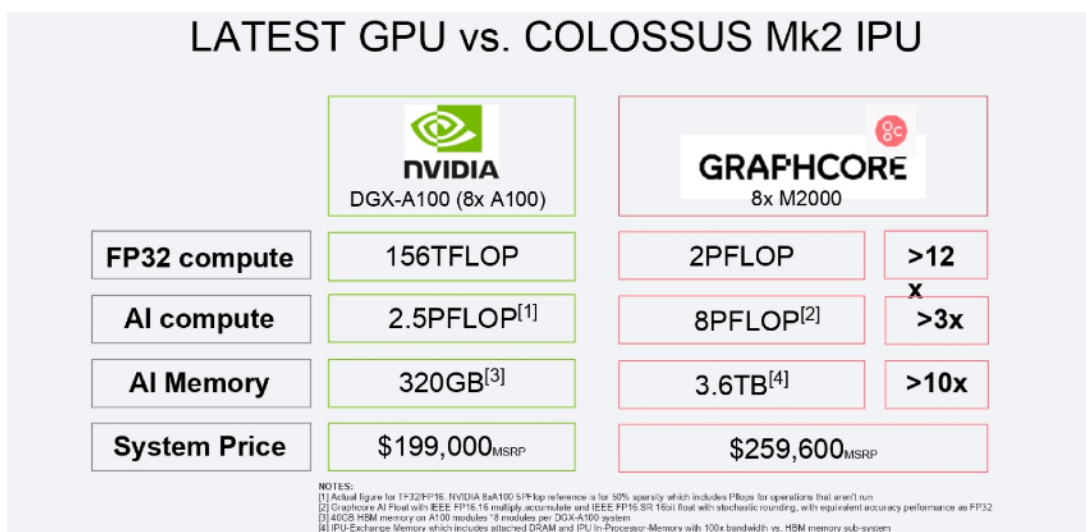


Abb. 4 - Die leistungsfähigsten AI - Prozessoren²⁴

In den zurückliegenden Jahren haben immer mehr Smartphone-Hersteller bestimmte Hardwareeigenschaften mit dem Label AI aufwerten wollen, so wimmelte es schon im letzten Jahr besonders bei chinesischen Telefonen nur so von ‚AI-Kameras‘. Tatsächlich geht der Einbau dedizierter ML-fähiger Architektur in den Smartphones aber zurück, da die für die Inferenz nötige Leistung bei Bild- und Spracherkennung durch Ressourcen schonende vorberechnete Modelle mit den herkömmlichen Chips erreicht wird. Google z.B. verzichtet 2020 auf den Einbau sogenannter *Neural Cores* in seine Pixel Phones. Einzig Apple arbeitet sowohl bei den iPhones als auch bei seinen neuen Computer-Chips an neuer und leistungsfähigerer hochintegrierter Parallelhardware (*Neural Engine* im M1 und A13 Chip).

Interessant ist, dass für all die genannte Hardware Versionen von TensorFlow²⁵ verfügbar sind, so dass für die Entwickler eine einheitliche Schnittstelle zu einer Fülle von ML-Power verfügbar ist. Gerade diese Fülle von Hardware, lokal und cloudbasiert, hat den Bedarf von Performance-Standards und -Vergleichen wie z.B. *MLPerf.com* nötig gemacht. Insgesamt brachte 2020 also eine

²² <https://www.nvidia.com/de-de/data-center/nvidia-ampere-gpu-architecture/>

²³ Weitere Entwickler solcher Hardware sind Google und Huawei

²⁴ <https://wccfttech.com/graphcores-colossus-mk2-gc200-7nm-ai-chip-rivals-nvidia-a100-gpu/>

²⁵ Allerdings hat aufgrund der übersichtlichen Struktur das Framework PyTorch den Standard TensorFlow in 2020 überholt.

Professionalisierung ,zunehmende Abstrahierung und Vergleichbarkeit bei der Anwendung von ML-Hardware. Der Ruf nach Vergleichbarkeit von ML-Leistung und der erzielten Ergebnisse deutet auf eine weitere Professionalisierung im Übergang von der experimentellen zur industriellen AI an. Vorherrschend bleiben im Cloudbereich bleiben große Anbieter wie Google Cloud Services, Amazon AWS, Microsoft Azure und IBM Watson.

Quantum Computing und Neuromorphic Computing

Der Bereich des Quantum Computing gilt als der ambitionierteste und komplexeste Bereich der AI Forschung. Nach dem im letzten Jahr noch anhaltenden Rennen um die sogenannte *Quantum Supremacy* gibt es in 2020 sowohl bei Amazon AWS als auch bei Microsoft Azure Quantum mittlerweile Quantum-Rechenleistung und Quanten-Stacks zu mieten. Projekte dazu befinden sich 2020 noch im Experimentalstadium. Bis zur Erlangung des Heiligen Grals, dem Brechen aller herkömmlichen Verschlüsselungen, wird es also noch eine Jahre dauern. Gerade wegen dieser strategischen Verheißung bzw. Bedrohung bleibt Quantum Computing allerdings eine wichtige Komponente eines internationalen Rüstungswettlauf zwischen den USA und China.

Auch beim Neuromorphic-Computing gab es graduelle Weiterentwicklung und immerhin schon einige praktische Anwendungen. Diese Chips sind am sogenanntem Edge das Gegenteil der riesigen AI-Rechenzentrum und versprechen mit geringem Stromverbrauch in externen autonomen Systemen AI-Leistungen z.B. zur Navigation oder schnellen Reaktion vor Ort bereitzustellen²⁶. Hier versucht u.a. Intel²⁷ eine Nische zu besetzen.

4 Problematische Entwicklungen

Die Durchbrüche und Stärken bei immer weiter ausufernden Ressourcen z.B. durch riesige Transformer-Modelle führten 2020 auch zu Kritik. Zum einen beschleunigt die Fokussierung auf immer leistungshungrigere ML-Modelle eine weitere Konzentration der AI-Entwicklung bei großen Anbietern wie Google, Facebook, Open AI, Amazon und Microsoft. Zu der ohnehin schon magnetischen Anziehungskraft auf das knappe weltweite AI-Talent zementieren diese Anbieter ihre Macht durch die Zentralisierung von Ressourcen wie Kapital und Rechenleistung.

Auch die internationale Zusammenarbeit in einer offenen Forschungsgemeinschaft leidet unter der extremen Kommerzialisierung. Laut stateof.ai 2020 werden nur noch 15% aller Forschungspaper mit Code veröffentlicht. Lieber hält man diese unter Verschluss für eine potenzielle Exklusivnutzung oder Patentierung. Pluralität und freier Austausch von Forschung und Entwicklung sind die Verlierer dieser Entwicklung. Typisch für beide Entwicklungen ist die Exklusiv-Lizenzierung des Open AI Modells GPT-3 an Microsoft.

Zwei weitere Kritikpunkte an der Big AI kulminierten zum Ende des Jahres im exemplarischen Fall vom Timnit Gebru. Die Ethical-AI-Forscherin hatte bei Google darauf hingewiesen, dass zentralisierte Transformer-Modelle wie BERT Biases and Stereotype verstärken und zudem durch die immense Hardwareleistung umweltschädlich sein können²⁸. Das einmalige Training eines Transformer-Modells wie BERT soll etwa so viel CO² verbrauchen wie ein Hin- und Rückflug SF-NYC.

²⁶ <https://www.forbes.com/sites/patrickmoorhead/2020/12/10/intel-labs-moving-mountains-with-neuromorphic-computing-and-photonics-technologies/?sh=508cdfab182f>

²⁷ <https://www.intel.de/content/www/de/de/research/neuromorphic-computing.html>

²⁸ <https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/>

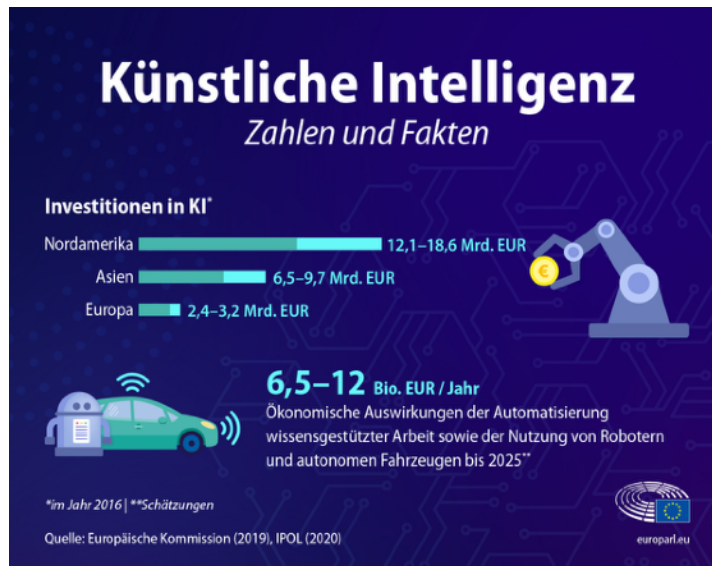


Abb. 5 - Investment in AI. (Quelle: Europäische Kommission)

Zu diesen Kritikpunkten kommen aus europäischer Perspektive erhöhte Anforderungen an Security und Privacy für die in der AI erforderlichen Daten. Die EU versucht sich in einem Spannungsfeld zwischen dem totalen Datenautoritarismus Chinas und dem Surveillance Capitalism der USA zu verorten und nach eigenem Selbstverständnis eine werteorientierte am Menschen orientierte KI zu etablieren²⁹. Allerdings befürchten Vertreter der Industrie und Forschung, dass eine Überregulierung bei gleichzeitiger Unterfinanzierung den Wettbewerbsnachteil der EU gegen über den beiden AI Superpowers USA und China weiter zementieren wird. Angesichts dieser Entwicklungen und einer gefährlichen Zögerlichkeit sprach AI-Topforscher Hans Uszkoreit bei der Rise of AI Conference 2020 von einer „severe situation“.



Abb. 6 - Hans Uszkoreit: „The Global AI Landscape“ at Rise of AI 2020³⁰

Zwei Entwicklungen, die sich aus den oben erwähnten Kritikpunkten ergeben und in den nächsten Jahren an Bedeutung gewinnen werden, sind *Explainable AI* und *Federated Learning*. Der Verdacht oder die Bedrohung maschineller Unvoreingenommenheit durch unverstandene Biases ruft ebenso

²⁹ <https://www.europarl.europa.eu/news/de/headlines/society/20201015STO89417/ki-regeln-wofur-das-europaische-parlament-eintritt>

³⁰ <https://youtu.be/Ba0XgeQz2ol>

wie die erhöhten gesetzlichen Transparenzanforderungen die sogenannte Explainable AI (XAI³¹) auf den Plan. Diese soll anstelle des Black-Box-Modells unverstandener Gewichtungen in den tiefen Schichten eines Neuronalen Netzes für den Menschen nachvollziehbare Faktoren bzw. Gründe, die zu einer bestimmten Entscheidung führen, verdeutlichen. Das Federated Learning hingegen versucht den Ansatz einer zentralen Datenspeicherung für das ML und die damit verbundenen Privacy Probleme durch verteiltes Lernen aufzulösen. Technisch verbleiben die Daten in den verteilten Speichern (auch Mobile oder Edge Devices), das Modell wird lokal mit diesen Daten trainiert, und erst am Ende des Lernvorgangs wieder zum zentralen Server geschickt. So können die Daten an der Peripherie verbleiben und dennoch umfangreiche Modelle errechnet werden. Ein Framework wie z.B. Flower³² ermöglicht die Verteilung bestehender ML-Verfahren unter Nutzung bestehender Modelle in PyTorch oder Keras.

5 Enterprise und Start-ups in AI

Angesichts der im letzten Abschnitt skizzierten Hegemonie der beiden AI Supermächten³³ USA und China in den Bereichen Entwicklung, Consumer-AI und AI mit Massendaten sieht die deutsche Industrie ihre Stärken in der industriellen AI. Mit SAP hat Deutschland zwar ein führendes Digital-Enterprise-Unternehmen von internationalem Rang. Doch bei einer Bewertung, die etwa den US-Konkurrenten bei industrieller Software: Oracle und Salesforce, entspricht spielt SAP nicht in der Liga der großen AI-Firmen mit. Über alle Nutzungsformen: Consumer, Research und Enterprise sind die US-amerikanischen und zunehmend die chinesischen AI-Leader allen europäischen Verfolgern enteilt.

Ein beträchtlicher Teil der deutschen Industriefirmen hängt an der Automobilbranche³⁴. Der technologische Großangriff durch die gleichzeitige Elektrifizierung und KI-Ausstattung von Autos, symbolisiert im Erfolg von Tesla, führt zwangsläufig zu einem Umdenken bei Herstellern wie VW, Daimler und BMW. Die Aufholjagd beim intelligenten Fahren (mit dem Fernziel des (teil-)autonomen Fahrens hat begonnen, und die Software-Schwächen der deutschen Industrie werden offenbar³⁵. Im Maschinenbau wird weiterhin versucht unter dem Label *Industrie 4.0*, sich einen Bereich der (fertigungs-)industrienahen Digitalisierung und damit auch der AI-Anwendung zu sichern. Ankündigungen wie die von Bosch, man wolle bis 2025 100% der Produkte mit AI herstellen oder aufwerten, passen in dieses Paradigma³⁶. Doch erfährt der Begriff Industrie 4.0 auch vielfältige Kritik, insbesondere wegen der Denkweise, dass primär graduelle Verbesserung bei Effizienz und Kostensenkung akzentuiert werden, statt dass völlig neue Dienstleistungen und Businessmodelle hervorgebracht werden. Im Gegensatz dazu gewinnen amerikanische digitale Plattformen gerade auch in der Corona Krise noch einmal einen riesigen Push. Ein Versuch, dieser Vorherrschaft etwas entgegen zu setzen, ist das Europäische Projekt Gaia-X³⁷, das

³¹ <https://chetanambi.medium.com/5-explainable-ai-xai-frameworks-92359b661e33>

³² <https://flower.dev>

³³ https://en.wikipedia.org/wiki/AI_Superpowers

³⁴ <https://www.listenchampion.de/produkt/top-1-000-industrieunternehmen-deutschland-liste-der-groessten-industriefirmen/>

³⁵ Mangels Angebots hat VW begonnen, selber IT-Kräfte heranzubilden. <https://www.webundmobile.de/beyond-dev/webinar/vw-stellt-2020-verstaerkt-software-experten-2449958.html>

³⁶ <https://www.handelsblatt.com/technik/forschung-innovation/kuenstliche-intelligenz-bosch-kaempft-fuer-einen-ki-kodex/26676670.html?ticket=ST-15797993-mFUfTXm4ofea9tluqxMZ-ap2>

³⁷ <https://www.bmwi.de/Redaktion/DE/Dossier/gaia-x.html>

Datensouveränität mit einer federated Cloud-Struktur zu verbinden sucht³⁸ und damit auch AI-Applikationen in einem rechtssicheren Umfeld ermöglichen soll. Gerade das Datensourcing, selbst für essentielle Forschung, wird zu einem zunehmenden Problem. So wurden für eine an der Uni Augsburg entwickelte Covid-Erkennungsapp, die initialen Sprach-Trainingsdaten aus China bereitgestellt³⁹. Mittlerweile werden auch deutsche Daten verwendet, aber das Beispiel zeigt, wie China zunehmend nicht nur auf der Hard- und Softwareseite, sondern auch bei der Datenbeschaffung uneinholbar an die Spitze eilt und ein breites und rechtssicheres Datenhandling in Europa für Forschung und Entwicklung zur Überlebensfrage im Sektor AI werden kann.

Company	Country	Market	Funding (\$)
1. SenseTime	China	Computer Vision, Facial Recog.	3.3 B
2. Megvii	China	Internet of Things	1.4 B
3. Nuro	USA	Autonomous Vehicles, Robotics	1 B
4. OpenAI	USA	Natural Language Proc., Research	1 B
5. UiPath	USA	Robotic Process Automation	1 B
6. Zoox	USA	Autonomous Vehicles	955 M
7. Samsara	USA	Internet of Things	930 M
8. Automation Anywhere	USA	Robotic Process Automation	840 M
9. MiningLamp	China	Data Analytics	786.6 M
10. ThoughtSpot	USA	Data Analytics	743.7 M

Source: Crunchbase Jun 2020
DISFOLD

Abb.7 - Top 10 AI Start-ups 2020 Funding (Quelle Disfold.com)

Ein weiterer Erfolgsfaktor ist die Investmentlandschaft⁴⁰. Der von Disfold zusammengestellte Überblick über die größten AI-Start-ups 2020 zeigt, dass die Finanzierung eines einzelnen Unternehmens in den USA oder China die Höhe des gesamten mehrjährigen Europäischen AI-Fördertopfes erreichen kann. (Allerdings kann man einige der oben aufgelisteten Unternehmen mit einer Historie von z.T. mehr als 10 Jahren kaum mehr als Start-ups verstehen.)

Ansonsten gab es in 2020 in Deutschland weitere Anstrengungen in den Bereichen Process Automation, Business Intelligence und AI Education⁴¹. Der Anteil selbstentwickelter AI liegt in der deutschen Industrie laut Deloitte⁴² in 2019/20 bei 8%, der Großteil der AI-Leistungen wird zugekauft und über externe Dienste (Cloud Services) abgewickelt. Die deutsche Industrie sieht AI primär als Business-Technik für die weitere Optimierung in Produktion und Produktentwicklung an.

„Deutsche Unternehmen befinden sich beim Einsatz von Künstlicher Intelligenz weitgehend auf Augenhöhe mit der internationalen Konkurrenz. Besonders erfreulich ist die vergleichsweise kurze Amortisationszeit von KI-Projekten hierzulande, die häufig auch auf den Einsatz von externen KI-Experten zurückzuführen ist, was im Regelfall die Implementierung beschleunigt und die Kosten senkt.“ (Milan Sallaba - Partner und Leiter Technology Sector, Deloitte Deutschland)

³⁸ <https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html>

³⁹ <https://www.uni-augsburg.de/de/campusleben/neuigkeiten/2020/11/09/3204/>

⁴⁰ <https://www.cbinsights.com/research/artificial-intelligence-top-startups/>

⁴¹ <https://kipodcast.de>

⁴² <https://www2.deloitte.com/de/de/pages/technology-media-and-telecommunications/articles/ki-studie-2020.html>

Das Zitat unterstreicht, dass die etablierte deutsche Industrie wie schon bei der Diskussion um Industry 4.0 angesprochen AI als Tool, nicht aber als Vision ansieht, um völlig neue Produkte und Businessmodelle zu entwickeln. Letzteres ist auch in Europa die Domäne von AI-Start-ups. Hier gibt es auch in 2020 in Deutschland und Europa viele Newcomer in einem riesigen Markt. Applied AI⁴³ kartiert für 2020 in Deutschland über 200 AI-Start-ups, für Europa weit über 500. Netzwerke, Enabler und Aggregatoren wie AppliedAI und ARIC Hamburg werden damit zum Dreh- und Angelpunkt für Start-ups bei der Vernetzung und als Schnittstelle zwischen Start-ups und etablierter Industrie. Einen weiteren ambitionierten Versuch Forschung, Start-ups und Investment in einem starken AI-Ökosystem zusammenzubringen, findet man in Berlin mit dem Merantix-AI-Venture-Studio⁴⁴. Hier sollen forschungsorientierte Deep-Tech-Produkte zu erfolgreichen Start-ups skaliert werden. Ein vielversprechender Ansatz in der nach wie vor extrem forschungstarken deutschen und europäischen AI-Landschaft.

Fazit 2020 - Ausblick 2021

Als Fazit für 2020 lässt sich sagen, dass die AI-Entwicklung trotz der Corona-Sondersituation weiter mit Warp Speed vorangeschritten ist. Die AI-Modelle wurden noch größer und leistungsfähiger und haben gezeigt, dass quantitatives Upscaling in qualitative Sprünge umschlagen kann. Die Übertragung der Transformermodelle aus der NLP in die visuelle AI als Alternative zu Convolutional Neural Networks zeigt, dass der Transfer von Modellen neue noch leistungsfähigere Paradigmen hervorbringen kann. Ebenso zeigt AI in 2020, dass ganze Industrien wie Pharma und Life Science ihre Entwicklungspipeline auf völlig neue Verfahren umstellen können und damit Produkte und Dienstleistungen möglich sind, die mit herkömmlicher Labortechnik kaum erreichbar wären.

Aber auch die fachlich fundierte Kritik an den hoch aggregierten Riesen-AI-Modellen in den Rechenzentren von Google, Open AI, Microsoft, Facebook und Amazon führt zu neuer Kreativität bei hybriden Modellen aus Deep Learning und symbolischer AI, zu neuen verteilten Ansätzen wie beim Federated Learning und der langfristigen Weiterentwicklung der klassischen GPU-Architektur im Neuromorphic- und Quanten-Computing.

Die Situation Deutschlands und Europas im Spannungsfeld der AI Supermächte USA und China ist prekär und braucht eine neue klare AI-Perspektive. Es gilt eine toxische Mischung aus technophober Überregulierung bei gleichzeitiger Unterfinanzierung zu vermeiden. Wird es in Europa unmöglich, Trainingsdaten zu bekommen, werden diese halt in China beschafft, oder die Start-ups wandern gleich mit ihrem ganzen Know-How in weniger regulierte Regionen ab.

Die auf den ersten Blick zersplitterte AI-Landschaft in Europa wird durch Netzwerke und das Bemühen um gemeinsame Standards zusammengehalten. Die EU hat das ganze Jahr um neue allgemeine Regeln gerungen und will im Q1/2021 einen Entwurf für verbindliche KI-Richtlinien vorlegen. Die Vernetzung einer freien und offenen Forschung, die rechtssichere Verfügbarmachung von Daten und die nachhaltige Finanzierung zur Skalierung erfolgreicher Start-up-Ideen weisen den Weg in eine europäische AI der Zukunft. Der durch hohe Bildung und persönliche Freiheiten extrem vielfältige und kreative Forschungsstandort Europa braucht Mittler und Netzwerke wie ARIC-Hamburg, Applied AI oder Merantix, damit aus Ideen Produkte und Unternehmen werden.

⁴³ <https://www.ai-startups.de>

⁴⁴ <https://www.merantix.com>