



## Kurzbriefing

# ChatGPT und andere Generative KI: Einordnung, Herausforderungen, Handlungsräume

Juni 2023

Der VDE hat schon ab 2019 auf ihre Bedeutung aufmerksam gemacht und erschließt seitdem vorausschauend Handlungsräume für Politik und Wirtschaft: **Generative Künstliche Intelligenz ist 2023 ganz oben auf die Agenda geklettert.**

In diesem Briefing geht es um die **Fähigkeiten und Unfähigkeiten dieser Technologie, um ihre Chancen und Risiken sowie um Möglichkeiten, die politische, gesellschaftliche und wirtschaftliche Disruption aktiv zu gestalten.**

### (1) Kontext: Worüber reden wir?

Unter Generativer KI versteht man ein ganzes Bündel von Werkzeugen und Methoden, um automatisiert Texte, Bilder und Videos für unterschiedlichste Anwendungen und in beliebiger Menge zu erzeugen. Dazu gehören beispielsweise **Deepfakes** (Video/Audio, seit 2018), **GPT 2+3** (Texte, seit 2019/20), Github Copilot (Programmcode, seit 2021), **Stable Diffusion, MidJourney** und **DALL-E** (Bilder aus Textprompts, seit 2022), **ChatGPT** und **Bard** (Texte/Dialoge, seit Ende 2022) und viele weitere.

Generative KI bringt Umwälzungen unter anderem für den Arbeitsmarkt, für den Bildungssektor, für das Rechtswesen, für die innere und äußere Sicherheit, für Medien, für Handelsplattformen und nicht zuletzt für den gesellschaftlichen Diskurs und Zusammenhalt.

Ein konstruktiver Umgang wird durch den Hype erschwert, der seit Ende 2022 insbesondere rund um ChatGPT zu Konstatierungen ist und zu Übertreibungen und Fehleinschätzungen führt.

Die folgenden Abschnitte umreißen Antworten auf einige grundsätzliche Fragen:

- Welche Auswirkungen sind zu erwarten?
- Welche konstruktiven Potenziale lassen sich nutzen?
- Welche offensichtlichen und versteckten Gefahren birgt Generative KI?
- Wie kann die Politik reagieren und den Umgang mit der sich rasend schnell entwickelnden Technologie aktiv gestalten?

# 5 Tage

bis 1 Million  
Nutzerinnen und Nutzer  
erreicht wurden

Chat GPT

150 Tage

Spotify

300 Tage

Facebook

730 Tage

Twitter

## Verbreitung in rasendem Tempo

ChatGPT verbreitet sich in ungekannter Geschwindigkeit – auch ein Grund für die teils unsachliche Debatte.

## (2) Eine rasante Entwicklung

Die Bereitstellung von ChatGPT Ende 2022 wird vielfach als „iPhone-Moment“ der KI-Technologie bezeichnet: Innerhalb von nur fünf Tagen erreichte die Anwendung eine Million Menschen – ein nie dagewesenes Tempo der Verbreitung.

Das wirtschaftliche Potenzial gilt als erheblich. Microsoft will Medienberichten zufolge den ChatGPT-Hersteller OpenAI mit bis zu 10 Milliarden US-Dollar ausstatten. SAP will einen dreistelligen Millionenbetrag in Generative KI investieren, und das McKinsey Global Institute (MGI) sieht den jährlichen Produktivitätszuwachs durch die Technologie bei 2,4 bis 4,1 Billionen Euro.

Spiegelbildlich dazu sieht sich Google in seinem Kerngeschäft durch den Erfolg von ChatGPT massiv gefährdet und rief Anfang des Jahres einen Alarm „Code Red“ aus.

Aus europäischer Perspektive stellen sich hier wie auch bei KI insgesamt Fragen der **digitalen Souveränität**: Werden in Europa die notwendigen private und staatlichen Investitionen mobilisiert, um eigene Recheninfrastrukturen und Kompetenzen aufzubauen und Köpfe zu halten? In welche Abhängigkeiten geraten europäische Unternehmen und Institutionen durch aktuelle und mögliche künftige Geschäftsmodelle der großen Anbieter Generativer KI sowie durch möglicherweise vertrauliche Daten europäischer kommerzieller und privater Nutzer, die als Prompts abfließen? Welche Rolle können deutsche und europäische Firmen wie Aleph Alpha, HuggingFace oder nyonic AI, Initiativen wie OpenGPT-X und LEAM sowie Inkubatoren wie der KI Park spielen?

Diese Fragen sind dringlich, und der VDE steht zur Unterstützung und zum neutralen Sparring dabei gerne zur Verfügung.

## (3) ChatGPT & Co.: Fähigkeiten und Unfähigkeiten

ChatGPT ist prominentestes Beispiel eines sogenannten **Sprachmodells** (englisch „**Large Language Model** (LLM)“), das in Sekundenbruchteilen beispielsweise den zweiten Weltkrieg in Versform zusammenfassen, brauchbare Website- und Marketingtexte formulieren, positive oder negative Produktbewertungen schreiben, aktuelle Nachrichten kommentieren oder auch Neuntklässlern Schulvorträge über Napoleon Bonaparte bereitstellen kann.

Zugrunde liegen künstliche neuronale Netzwerke, sogenannte Transformermodelle, die hocheffizient sprachliche Muster und Stile nachahmen. ChatGPT wurde mit praktisch allen im Internet verfügbaren Texten trainiert. Ein Fokus liegt auf der Fähigkeit, die generierten Texte aufgrund von Nutzerreaktionen anzupassen und weiterzuspinnen, so dass letztlich ein Dialog entstehen kann.

Die nachstehende Tabelle gibt einen Überblick über die tatsächlichen Fähigkeiten als auch die Unfähigkeiten von ChatGPT und ähnlichen Sprachmodellen:

ChatGPT und ähnliche Sprachmodelle können ...	ChatGPT und ähnliche Sprachmodelle können <b>nicht</b> ...
<ul style="list-style-type: none"> <li>■ sprachlich korrekte, inhaltlich plausible und oft nützliche Inhalte in vorgegebener Länge als Antwort auf umgangssprachliche Eingaben („Prompts“) generieren.</li> <li>■ mit fast beliebigen Themen umgehen.</li> <li>■ große Textmengen inhaltlich zusammenfassen.</li> <li>■ aus Stichpunkten längere Texte erstellen, auch in mehreren Varianten</li> <li>■ die produzierten Texte stilistisch in jeder beliebigen Form variieren.</li> <li>■ Hypothesen mit oft erstaunlich hoher Trefferquote aufstellen.</li> </ul>	<ul style="list-style-type: none"> <li>■ Texte inhaltlich in irgendeiner Weise verstehen.</li> <li>■ logisch und stringent denken.</li> <li>■ ein Bewusstsein entwickeln oder selbständig Ziele verfolgen.</li> <li>■ die inhaltliche Korrektheit und Qualität von Texten garantieren, insbesondere nicht bei Fachthemen.</li> <li>■ Output produzieren, der sich ohne fachliche Prüfung und Überarbeitung durch kompetente Personen nutzen lässt, insbesondere nicht in kritischen Anwendungen.</li> <li>■ verlässliche Quellen für Aussagen angeben und würdigen.</li> </ul>

### (4) Arbeitsmarktpolitische Implikationen

Wie jede disruptive Technologie wird Generative KI die Rolle menschlicher Arbeitskräfte in vielen Branchen stark verändern und teilweise ersetzen. So werden beispielsweise Softwarehäuser, Kundendienst-Call-Center, Versicherungen, Personalvermittler, Übersetzungsbüros, Medienhäuser, Handbuchautoren oder auch Werbeagenturen unter Druck geraten. IBM, British Telecom und Axel-Springer haben bereits Stellenstreichungen im Zusammenhang mit der verstärkten Nutzung von ChatGPT ins Spiel gebracht.

Gleichwohl: Gerade in Deutschland hat die technologiegetriebene Automatisierung vor allem in der

Industrie immer wieder die Wettbewerbsfähigkeit erhöht und dadurch unter dem Strich zu mehr statt weniger Arbeitsplätzen geführt. Dies ist auch für Generative KI zu erwarten, nicht zuletzt auch durch neue Synergien zwischen Menschen und Maschinen: So können qualitativ hochwertige Fachtexte durch Fachleute und KI-Systeme gemeinsam schneller formuliert werden. Im Pharmabereich kann Generative KI Kandidaten für wirksame Wirkstoffe zielgenau identifizieren, die dann von Menschen wie bisher geprüft werden. In der Softwareentwicklung kann Generative KI im Tandem mit menschlichen Programmierern die Produktivität steigern.

### (5) Bildungspolitische Implikationen

Sprachverarbeitende KI-Modelle verändern die Art und Weise, wie gelernt und gelehrt wird. Schülerinnen und Schüler nutzen ChatGPT, um Aufsätze schreiben zu lassen, Referate zu skizzieren oder Texte zusammenzufassen. Studierende lassen sich bei Hausarbeiten umfassend unterstützen. Die Qualität der Arbeiten speist sich damit weniger aus erlerntem Wissen als vielmehr aus der Fähigkeit, dem Sprachmodell die richtigen Aufgaben zu stellen und dessen Antworten nachzujustieren. Zugleich können KI-generierte Texte mit technischen Hilfsmitteln nicht zuverlässig als solche erkannt werden. Auch Plagiatsoftware fällt aus, da es sich gar nicht um Plagiate im engeren Sinne handelt.

Durch die Verfügbarkeit von ChatGPT und ähnlichen Werkzeugen werden Prüfungsformate, die sich auf das eigenständige Verfassen von Texten stützen, grundsätzlich in Frage gestellt.

Generative KI kann auch in der Lehre unterstützen, beispielsweise bei der dialogischen Vermittlung von Lerninhalten durch Avatare auf individuell zugeschnittenem Niveau. Aufgabe von Schulen und Hochschulen wirdes sein, diese Formen der Zusammenarbeit zu trainieren. Gleichzeitig wird zu klären sein, inwieweit bestehende Curricula anzupassen und welche Fähigkeiten weniger stark zu vermitteln sind.

### (6) Sicherheits- und gesellschaftspolitische Implikationen

Bei allem vorsichtigen Optimismus birgt Generative KI einige ernstzunehmende Gefahren, die bis hin zur **Gefährdung unseres demokratischen Gemeinwesens reichen und dringende politische Aufmerksamkeit erfordern**.

Mithilfe Generativer KI können nämlich auch **gefälschte und irreführende Inhalte automatisiert und zielgruppenspezifisch** erstellt werden. Durch die Kombination von Video, Audio, Bildern, Texten und Dialogen können außerdem **nichtexistente „Personen“ in praktisch beliebiger Zahl vorgetäuscht** werden. Internetforen und Social-Media-Kanäle lassen sich in völlig neuem Ausmaß und neuer Qualität mit solchen Bots fluten. Gezielte Einflussnahmen wie seit längerem die Aktivitäten der russischen „Internet Research Agency“ (IRA) oder die Aktionen von Cambridge Analytica in jüngerer Zeit werden beliebig skalierbar, einfacher durchzuführen und schwieriger abzuwehren, insbesondere wenn sie mit „Microtargeting“ kombiniert sind.

Generative KI dient in den Händen böswilliger politisch oder kommerziell motivierter Akteure als Brandbeschleuniger für Desinformationen und Fakes im digitalen Raum und bedroht das demokratische Gemeinwesen. Selbst abseits des gezielten Missbrauchs stellt sich die Frage, welche Folge das

sprunghafte Wachstum der Menge verfügbarer Inhalte durch Generative KI haben wird, während gleichzeitig die menschliche Aufnahmekapazität gleichbleibt.

Zu den politischen Gefahren kommen außerdem **Bedrohungen für das Funktionieren unserer Wirtschaft**: Ein immer leichteres Fälschen von überzeugend klingenden Bewertungen von Online-Aktivitäten (z. B. für Produkte, Dienstleistungen, Inhalte, Dienstleistungsanbieter, Nutzer usw.) untergräbt sowohl den traditionellen Handel auf großen digitalen Plattformen als auch die Sharing Economy und damit ein entscheidendes Element einer nachhaltigeren Wirtschaft und Gesellschaft.

**KI-generierte Inhalte lassen sich nicht zuverlässig von solchen unterscheiden, die Menschen erstellt haben**. Traditionelle Ansätze wie Faktenchecker oder Plagiatsoftware sind weitgehend machtlos. Kennzeichnungspflichten und Wasserzeichen sind kaum durchsetzbar und zum Teil technisch gar nicht möglich, insbesondere bei kürzeren Texten und bei Mischformen, in denen KI-Output noch durch Menschen modifiziert wurde. Auch die Hoffnung auf etwaige Erkennungswerkzeuge ist verfehlt. Insoweit solche Erkennungswerkzeuge überhaupt machbar sind, könnte die KI kurzfristig gezielt weitertrainiert werden, um sie zu überlisten.

### (7) Lösungsweg des VDE: Infrastrukturen für Vertrauen im digitalen Raum

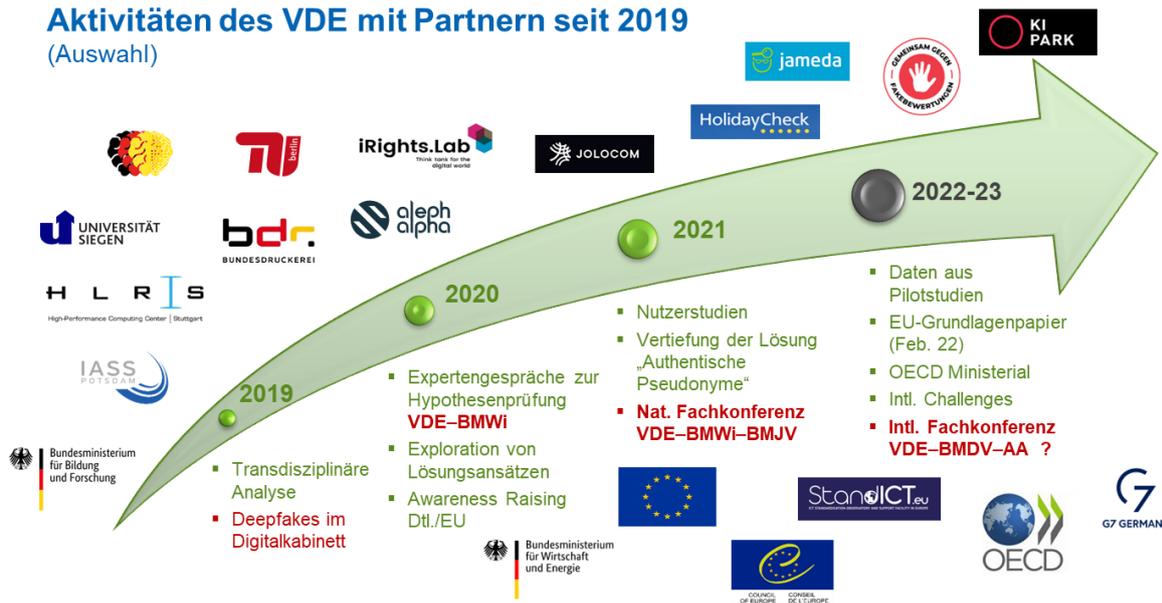
Um die negativen Auswirkungen von Generativer KI, insbesondere die gesellschafts- und sicherheitspolitischen Implikationen, in den Griff zu bekommen, hat der VDE bereits seit 2019 Lösungskonzepte entwickelt und diese in Forschungs- und Pilotprojekten, in mehreren Fachgesprächen und Fachkonferenzen gemeinsam mit Bundesministerien sowie auf EU- und OECD-Ebene validiert.

Im Mittelpunkt steht die Frage, wie es in Zeiten Generativer KI, in der Inhalte und Personen nach Belieben gefälscht werden können, möglich bleibt, Vertrauen im digitalen Raum zu schaffen. Denn ohne Vertrauen funktionieren weder der politische Diskurs, noch der freie Markt, noch unsere Medienlandschaft. Im Zentrum stehen dabei Mechanismen und Infrastrukturen, mit denen sich die Vertrauenswürdigkeit und Urteilskraft von Quellen von Informationen einschätzen lässt. Dazu

gehört auch die Möglichkeit, Informationsquellen eine Identität zuschreiben zu können, ohne dass damit ein Verlust von Privatsphäre verbunden sein darf. Technische Lösungen existieren bereits, die nun jedoch auf in der Breite umzusetzen und dabei präzise auf Nutzerbedürfnisse und -verhalten zuzuschneiden sind.

In diesem Sinne baut der VDE gemeinsam mit Partnern ein Trusted Information & Identity Lab (TIIL) auf – ein deutsches und europäisches Zentrum für Forschung, Entwicklung und praktische Pilotierung von Lösungen rund um Vertrauen im digitalen Raum. Zu den Lösungskonzepten gehören unter anderem sogenannte „authentische Pseudonyme“ als Privatsphäre schützende, kontext-singuläre Identitäten. Übergreifendes Ziel ist es, trotz aller Fakes ein vertrauensvolles, konstruktives, sicheres und faires Zusammenleben im digitalen Raum zu ermöglichen.

## Aktivitäten des VDE mit Partnern seit 2019 (Auswahl)



Zuvor hat der VDE bereits gemeinsam mit dem Land Hessen den „AI Quality & Testing Hub“ gegründet. Seine Aufgabe ist es, die Qualität von KI-Systemen messbar und prüfbar zu machen und dazu Kompetenzen, Werkzeuge und Netzwerke aufzubauen. Ihm kommt im Zeitalter von ChatGPT und Generativer KI die wichtige Rolle zu, die bisher nicht verlässliche Qualität solcher Systeme so weit in den Griff zu bekommen, dass sich die vielfältigen in diesem Papier skizzierten positiven Anwendungen tatsächlich sicher umsetzen lassen.

Der VDE freut sich auf den weiteren Austausch mit Stakeholdern aus der Politik auf Landes-, Bundes- und europäischer Ebene, um Generative KI aktiv, vorausschauend und zukunftssicher zu gestalten. Wie schon beim Thema KI-Ethik steht der VDE auch bei Generativer KI für das Ziel, Brücken zwischen Innovation und Verantwortung sowie zwischen grundlegenden Prinzipien und deren praktischen Umsetzung zu schlagen.

## Über den VDE

Der VDE, eine der größten Technologie-Organisationen Europas, steht seit mehr als 130 Jahren für Innovation und technologischen Fortschritt. Als einzige Organisation weltweit vereint der VDE dabei Wissenschaft, Standardisierung, Prüfung, Zertifizierung und Anwendungsberatung unter einem Dach. Das VDE Zeichen gilt seit mehr als 100 Jahren als Synonym für höchste Sicherheitsstandards und Verbraucherschutz.

Wir setzen uns ein für die Forschungs- und Nachwuchsförderung und für das lebenslange Lernen mit Weiterbildungsangeboten „on the job“. Im VDE Netzwerk engagieren sich über 2.000 Mitarbeiter\*innen an über 60 Standorten weltweit, mehr als 100.000 ehrenamtliche Expert\*innen und rund 1.500 Unternehmen gestalten im Netzwerk VDE eine lebenswerte Zukunft: vernetzt, digital, elektrisch. Wir gestalten die e-diale Zukunft.

Sitz des VDE (VDE Verband der Elektrotechnik Elektronik Informationstechnik e.V.) ist Frankfurt am Main. Mehr Informationen unter [www.vde.com](http://www.vde.com)

### Ansprechpartner zum Thema Generative KI:

Dr. Sebastian Hallensleben  
Leiter Digitalisierung & Künstliche Intelligenz

Tel. +49 69 6308 305  
[sebastian.hallensleben@vde.com](mailto:sebastian.hallensleben@vde.com)

VDE Verband der  
Elektrotechnik Elektronik Informationstechnik e.V.