Deutscher Bundestag
Ausschuss für Digitales

Ausschussdrucksache
20(23)147

17.05.2023

Bundesamt für Sicherheit in der Informationstechnik

Deutschland Digital•Sicher•BSI•

Antworten des Bundesamtes für Sicherheit in der Informationstechnik (BSI)

Fragenkatalog zur Öffentlichen Anhörung des Ausschusses für Digitales am 24.05.2023 „Generative Künstliche Intelligenz“

1) **Die Regulierung generativer KI ist derzeit Gegenstand der Verhandlungen um den europäischen AI Act (AIA). Wie kann Ihrer Einschätzung nach generative KI wirksam im AIA einbezogen und reguliert werden und wie beurteilen Sie vorgeschlagene Differenzierungen innerhalb generativer KI zwischen „general purpose AI“ und „foundation models“?**

Antwort BSI:

Die Regulierung der generativen KI soll aus BSI-Sicht risikobasiert und anwendungsbezogen erfolgen. Ein solcher Ansatz im AI Act wird seitens BSI begrüßt. Je nach Anwendung unterliegt ein KI-System bzw. eine anwendungsbezogene Erweiterung eines ggf. nicht-anwendungsspezifischen KI-Systems unterschiedlichen Anforderungen, die nachweisbar erfüllt werden müssen.

Zur Erstellung dieser Nachweise fehlen derzeit noch geeignete Prüfkriterien, Prüfmethoden und Prüfwerkzeuge. Zusammen mit Partnern aus Forschung und Wirtschaft arbeitet das BSI aktiv an der Definition solcher Prüfgrundlagen. Diese sollen in einem weiteren Schritt in die entsprechenden internationalen Normen und Standards eingebracht werden, auf die u. a. der aktuelle Entwurf des AI Acts verweist. Aus BSI-Sicht ist eine Kombination aus horizontalen Sicherheitsstandards, die für alle KI-Anwendungen auf dem EU-Markt gelten sollen, und branchen- und anwendungsspezifischen vertikalen Standards der geeignete Weg, um ein angemessenes Sicherheitsniveau zu erreichen.

2) **Generative KI bietet zahlreiche Anwendungsmöglichkeiten in den unterschiedlichsten Berufsständen und kann für Entlastungen am Arbeitsmarkt sorgen. Wie schätzen Sie die Potenziale und Risiken generativer KI für die Arbeitswelt ein und wo sehen Sie Regelungsbedarf?**

Antwort BSI:

Generative KI-Modelle bieten ein großes Potential, die Digitalisierung in Deutschland und Europa voranzutreiben, Arbeitsabläufe zu beschleunigen und die Qualität von Arbeitsergebnissen zu verbessern. Auch im Bereich der IT-Sicherheit gibt es viele Anwendungsmöglichkeiten, wie z. B. bei der Berichtserstellung zu Sicherheitsvorfällen oder der Analyse und Verbesserung von Programmcode.

Gleichzeitig entstehen durch die neuen Technologien aber auch neue Risiken und Missbrauchsmöglichkeiten. Mögliche Missbrauchsszenarien sind zum Beispiel:

- Generierung von Texten für Angriffe aus dem Bereich Social Engineering
- Generierung und Ausführung von Malware
- Generierung von Falschmeldungen (z. B. Desinformationen, Propaganda oder Hassnachrichten)

Darüber hinaus ergeben sich nach aktuellem Stand der Technik mindestens folgende Risiken für Anwendungen, die generative KI verwenden:

- Fehlende Faktizität („Halluzinieren“) und Reproduzierbarkeit der Ausgaben
- Fehlende Sicherheitseigenschaften von generiertem Programmcode
- Zum Teil fehlende Aktualität der Trainingsdaten
- Fehlerhafte Reaktion der Modelle auf sehr spezifische Eingaben
- Zum Teil fehlende Vertraulichkeit der eingegebenen Daten (z. B. werden diese ggf. für das weitere Training der Modelle verwendet)
- Abhängigkeit vom Hersteller/ Betreiber des Modells
- Anfälligkeit für KI-spezifische Angriffe, z. B.
    - "versteckte" Eingaben mit manipulativer Absicht (Adversarial Attacks, Indirect Prompt Injections),
    - die Beeinflussung des Modells durch Modifikation von Trainings- oder Feedbackdaten (Poisoning Attacks)
    - oder die Extraktion von vertraulichen Informationen aus dem Modell (Privacy Attacks).

Wenn Sprachmodelle in kritischen Bereichen eingesetzt werden, können die genannten Risiken signifikante Auswirkungen haben und sollten entsprechend im Rahmen einer systematischen und anwendungsfallabhängigen Risikoanalyse berücksichtigt werden.

Ein ausführliche Darstellung der oben genannten Inhalte ist in der BSI-Publikation „Große KI-Sprachmodelle: Chancen und Risiken für Industrie und Behörden“[1] nachzulesen.

3) Inwieweit können sich Anwendungen aus staatlichen oder wirtschaftlichen Systemen, die nicht immer demokratische und freiheitliche Werte teilen, auf die europäische Gesellschaft auswirken und wie sollten die EU und Deutschland damit umgehen?

Antwort BSI:

Es sollte sichergestellt werden, dass KI-Systeme, die in Deutschland oder der EU eingesetzt werden, unseren Gesetzen und Werten folgen. Es sind mehrere Maßnahmen vorstellbar, wie man solchen gesellschaftlichen Herausforderungen begegnen kann:

[1] https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/Grosse_KI_Sprachmodelle.pdf

- Aktive Forschungs- und Wirtschaftsförderung: Die aktive Förderung der Forschung an KI-Aspekten, welche die Einhaltung von europäischen Werten unterstützen (wie z. B. Transparenz, Nicht-Diskriminierung etc.) kann die globale Entwicklung von zukünftigen KI-Systemen in unserem Sinne beeinflussen. Der Auf- und Ausbau europäischer Unternehmen, deren KI-Systeme von Grund auf die europäischen Werte und Vielfalt abbilden, sollte unterstützt werden. Dieses kann Wettbewerber aus dem Nicht-EU-Ausland dazu veranlassen, solche Aspekte ebenfalls in ihren Systemen zu berücksichtigen.
- Sensibilisierung der Bürger: siehe Antwort auf Frage 6
- Europäische Normen und Standards: Durch die Verankerung europäischer Werte in nationalen und internationalen Normen und Standards und durch die Zulassung nur solcher KI-Systeme auf den EU-Markt, bei denen die Konformität zu diesen Normen und Standards wirksam nachgewiesen ist, können potenzielle negative Auswirkungen auf die europäische Gesellschaft minimiert werden.

KI-Technologien entwickeln sich derzeit rasant. Um diese Entwicklung effektiv beeinflussen zu können, müssen entsprechende Maßnahmen beschleunigt und weiter ausgebaut werden.

4) **Bisher gibt es einige Überlegungen und Projekte von Wasserzeichen bis Tools, die KI-generierte Texte markieren bzw. erkennen sollen – beides wird angesichts mangelnder Beständigkeit oder Treffsicherheit kritisch kommentiert. Wie könnte eine sichere und wirksame Kenntlichmachung von Inhalten, die durch generative KI entstanden sind, konkret aussehen? Und welche flankierenden Informationen könnten Nutzer:innen zwecks Aufklärung bereitgestellt werden?**

Antwort BSI:

Derzeit existieren keine vollständig zuverlässigen und wirksamen Technologien, um KI-generierte Texte zu erkennen oder zu markieren.

Folgende grundsätzlichen Möglichkeiten zur Detektion werden derzeit diskutiert:

- Werkzeuge, welche die statistischen Eigenschaft der Texte oder die Parameter eines KI-Modells verwenden, um einen Indikator zu berechnen, der anzeigt, ob ein Text maschinengeneriert ist
- Wasserzeichen, welche von den Herstellern der Modelle implementiert werden und spezifische Muster in den Ausgaben erzeugen, anhand derer die Ausgabe später identifiziert werden kann

Die Detektionsleistung der vorhandenen Werkzeuge, insbesondere wenn diese keinen direkten Zugriff auf die Parameter des Modells haben, ist oft begrenzt. Einschränkungen bestehen insbesondere bei kurzen Texten und Texten, die nicht auf Englisch verfasst sind. Grundsätzlich können die generierten Texte, ggf. auch automatisiert, nachbearbeitet werden, sodass statistische Auffälligkeiten und Wasserzeichen unkenntlich gemacht werden. Die Ergebnisse der genannten Ansätze können daher in der Regel nur einen Hinweis geben und stellen keine an sich belastbare Aussage dar.

Dennoch ist es wünschenswert, dass in diesem Bereich weiter geforscht wird und Hersteller sowie Betreiber von KI-Modellen mögliche Ansätze, wie z. B. die Verwendung von Wasserzeichen, umsetzen, um den Aufwand für einen Missbrauch der Modelle zu erhöhen. Für Modelle einer gewissen Größe und Bedeutung für Wirtschaft und Gesellschaft, ist es zudem wünschenswert, dass ausgewählte Stellen, z. B. staatliche Stellen oder Forschende, auf eine vertrauliche Art und Weise Zugriff auf die Parameter eines solchen Modells erhalten, um weitere Detektionsmechanismen zu erforschen und zu entwickeln.

Abgesehen von den Missbrauchsmöglichkeiten von KI-generierten Texten sollte betont werden, dass es umgekehrt aber wirksame kryptografische Methoden gibt, um menschengeschriebene Texte kenntlich zu machen, wie dies z. B. auch bei einer E-Mail Signierung verwendet wird.

Nutzerinnen und Nutzer sollten dafür sensibilisiert werden, dass es zukünftig umso wichtiger wird Medieninhalte, insbesondere aus unbekannten Quellen, kritisch zu hinterfragen und ggf. durch weitere unabhängige Quellen zu validieren.

5) **Derzeit kursieren zahlreiche Vorschläge, um die regulatorischen Herausforderungen generativer KI-Anwendungen in den EU-Gesetzgebungsvorhaben für eine KI-Verordnung und eine KI-Haftungsrichtlinie passgenau zu verankern: Ist der risikobasierte Ansatz zur Regulierung generativer KI überhaupt geeignet oder braucht es z. B. eine systemische Risikoanalyse analog zum Risikoanalyse- und Minimierungsmechanismus im DSA?**

Antwort BSI:

Vorrangig sollte der risikobasierte Ansatz gewählt werden (siehe Antwort zu Frage 1 und Frage 18).

6) **Sind neue Phänomene und Fragestellungen im Hinblick auf einen negativen Einfluss von Anwendungen generativer KI auf den demokratischen Meinungsbildungsprozess zu erwarten und wie lassen sich Medienfreiheit und Meinungsvielfalt im Zeitalter generativer KI rechtlich und politisch stärken, auch – aber nicht ausschließlich – im Hinblick auf die angemessene Vergütung von Journalist:innen, Künstler:innen und Kreativen und wo sehen Sie möglichen Anpassungsbedarf etwa im Urheberrecht?**

Antwort BSI:

Es ist nicht auszuschließen, dass Anwendungen generativer KI Einfluss auf die Meinungsbildung der Bürgerinnen und Bürger im Allgemeinen hat. Neben dem Modell und weiteren Faktoren haben die Trainingsdaten von Anwendungen generativer KI einen erheblichen Einfluss auf das Ergebnis: Ihre Auswahl und Zusammenstellung - versehentlich oder beabsichtigt - beeinflusst das Ergebnis. Ihre Manipulation durch Dritte ist ein denkbares Angriffsszenario.

Je nach Anwendung entstehen weitere Risiken durch generative KI, z. B.:

- Die Kommunikation mit einem Chatbot kann wie mit einem menschlichen Gesprächspartner wirken und – sofern der Chatbot nicht als solcher erkannt werden kann – so falsches Vertrauen aufbauen.
- Große KI-Sprachmodelle sind darauf ausgelegt, jede Frage zu beantworten. Das trifft auch dann zu, wenn die KI im Training keinerlei Informationen gelernt hat, die zur Beantwortung einer Frage oder einer Anweisung genutzt werden können. Falsche Antworten sind möglich.
- Im Internet verfügbare persönliche Informationen z. B. aus den sozialen Medien können Inhalt der Trainingsdaten sein. Damit können persönlichen Informationen Teil einer Ausgabe werden.
- Eingaben und Ausgaben bei der Nutzung eines Textgenerators werden durch den Anbieter verarbeitet und gegebenenfalls weiterverwendet.

Durch diese und weitere Risiken ist es ein Anliegen im Digitalen Verbraucherschutz des BSI Verbraucherinnen und Verbraucher zu informieren und zu sensibilisieren. Aufgrund der täuschend echt wirkenden Inhalte sollten Ausgaben daher immer kritisch bewertet werden. Deren Wahrheitsgehalt ist idealerweise durch mindestens zwei weitere unabhängige Quellen zu prüfen. Sollen KI-generierte Inhalte veröffentlicht oder weiterverbreitet werden, ist eine sorgfältige Prüfung und ggf. eine manuelle Nachbearbeitung dieser Inhalte empfehlenswert.

Verbraucherinnen und Verbraucher sollen die Chancen von KI vertrauensvoll nutzen können. Dazu gehört es KI zu erkennen, die Risiken von KI zu kennen und damit umgehen zu können. Aus Sicht des Digitalen Verbraucherschutzes ist die Einbeziehung sowie Berücksichtigung der Verbraucherinnen und Verbraucher bei direkter und indirekter Verwendung von KI in Bezug auf

- die transparente Gestaltung von KI,
- die sichere Handhabung im Sinne einer Usable Security,
- klare und für alle Risikoklassen gültige (Mindest-)Anforderungen an die Gestaltung von KI-Systemen,
- notwendige Warnung und Information der Verbraucherinnen und Verbraucher sowie
- eine ausbleibende Ausnutzung von Schwächen oder Eigenschaften der Verbraucherinnen und Verbraucher durch KI

wichtig.

7) **Welche rechtlichen Ansatzpunkte gibt es im EU-Recht (z.B. KI-VO-E, Wettbewerbsrecht, Urheber-RL) und im nationalen Recht (etwa UWG, Medienstaatsvertrag), um eine Kennzeichnungspflicht für KI-generierte Inhalte (etwa Videos, Bilder oder Texte) und Entscheidungen möglichst ohne Umgehungsmöglichkeiten zu implementieren – und welche technischen Ansatzpunkte sind denkbar, um solche Pflichten effektiv in digitalen Diensten um- und durchzusetzen?**

Antwort BSI:

Transparenzpflichten für KI-generierte Inhalte, wie sie z. B. in Artikel 13 (Transparenz und Bereitstellung von Informationen für die Nutzer) und Artikel 52 (Transparenzpflichten für bestimmte KI-Systeme) des aktuellen AI Act Entwurfs stehen, bilden eine rechtliche Grundlage zur Kennzeichnung von KI-generierten Inhalten. Es besteht jedoch weiterer Forschungsbedarf zur technischen Umsetzung dieser rechtlichen Anforderungen.

Technische Aspekte werden in der Antwort zu Frage 4 erläutert.

8) **Welche technisch-organisatorischen Maßnahmen halten Sie zum Schutz Minderjähriger für geeignet – sowohl im Hinblick auf das Einfließen ihrer personenbezogenen Daten in die Trainings- und Lernumgebung generativer KI als auch bezüglich der eigentlichen Nutzung von Anwendungen, die KI-basiert Texte, Videos oder Bilder generieren?**

Antwort BSI:
- keine Antwort -

9) **Welche KI-getriebene, wirtschaftliche Entwicklung prognostizieren Sie in der kurzen, mittleren sowie langen Frist für die deutsche sowie europäische Wirtschaft angesichts ihrer jeweiligen spezifischen Struktur und gehen Sie bzgl. der Implikationen für die reale Wirtschaftsleistung dieser Ökonomien im globalen Vergleich, auch in Abhängigkeit von Regulierung, von einer positiven oder negativen Entwicklung aus?**

Antwort BSI:

Die Entwicklungen im forschungsgetrieben Feld der KI-Technologie sind komplex, hoch dynamisch und schwer prognostizierbar, wie die zurückliegenden und aktuellen Ereignisse rund um generative KI (ChatGPT, Google Bard etc.) verdeutlichen. Internationale Normen und Standards für KI sollten daher zügig, praxisnah und flexibel entwickelt werden, um so auch den Interessen der Unternehmen im europäischen Wirtschaftsraum Rechnung tragen zu können. Die Unterstützung von europäischen Unternehmen, die jetzt handeln und bei der Entwicklung von KI-Systemen im Einklang mit europäischen Werten Sicherheits- und Verbraucherschutzaspekte berücksichtigen, erscheint vor diesem Hintergrund notwendig.

10) Wie stehen Sie zu dem von vielen anerkannten KI-Expertinnen und Experten unterzeichneten Brief des Future of Life Institutes: Bis zu welchem Grad teilen Sie die darin geäußerten Bedenken und halten Sie die darin formulierten Forderungen für sinnvoll?

Antwort BSI:

Das BSI geht davon aus, dass es sich hier um den Aufruf „Pause Giant AI Experiments: An Open Letter“[2] handelt.

Das BSI unterstützt die Forderung an Forschung und Entwicklung und arbeitet selbst aktiv daran, die nachprüfbare Sicherheit, Interpretierbarkeit, Transparenz, Robustheit und Vertrauenswürdigkeit von KI-Systemen zu steigern. Auch die Forderung nach einer schnelleren Entwicklung der legislativen, prozessualen sowie technischen Aspekten von KI-Governance-Systemen wird hier positiv gesehen.

Der Forderung nach einer mindestens sechsmonatigen Pause bei der Entwicklung von KI-Systemen, die mächtiger als GPT-4 sind, schließt sich das BSI nicht an. Die Einhaltung solcher Verbote lässt sich nicht wirksam kontrollieren. Daher sollte vielmehr die Aufklärung der Gesellschaft über Chancen und Risiken der KI intensiviert werden. Außerdem sollte sichergestellt werden, dass solche mächtige KI-Systeme ohne entsprechende wirksame Prüfungen nicht in kritischen Bereichen, wie sie z. B. in dem aktuellen Entwurf des AI Act definiert sind, eingesetzt werden.

11) Zum Ausbau einer Recheninfrastruktur in Deutschland für das Training von Algorithmen sind laut KI Bundesverband Investitionen in Höhe von 300 Millionen Euro erforderlich. Solle es Ihrer Auffassung nach Aufgabe des Staates sein, mit der (Ko-) Finanzierung einer solchen Infrastruktur aktive Industriepolitik zu betreiben, um es deutschen Unternehmen zu ermöglichen, auf dem globalen Markt generativer KI zu bestehen?

Antwort BSI:
- keine Antwort -

12) Es gibt eine weitgehende Übereinstimmung, Künstliche Intelligenz so zu regulieren, dass ihr Einsatz bestimmten Wertevorstellungen folgt. Wie kann dies konkret realisiert werden und wo ist die Grenze zu ziehen hinsichtlich einer möglichen Überregulierung, bei der Künstliche Intelligenz zu Künstlicher Ideologie werden könnte?

[2] https://futureoflife.org/open-letter/pause-giant-ai-experiments/

Antwort BSI:

Für die Durchsetzung europäischer Werte in KI-Systemen ist die Entwicklung von internationalen Normen und Standards, in denen europäische Werte verankert sind, äußerst wichtig. Sie müssen so entwickelt werden, dass deutsche und europäische Unternehmen dadurch nicht benachteiligt werden. Eine angemessene Förderung der europäischen Unternehmen und vor allem der KMUs und Start-Ups für die aktive Teilnahme an Standardisierungsprozessen wäre sinnvoll, damit ihre Interessen in den derzeit entstehenden Standards berücksichtigt werden. Auch das BSI wirkt in relevanten Standardisierungsgremien aktiv mit und setzt sich dort für die Verankerung eines angemessenen Sicherheitsniveaus für KI-Systeme ein.

13) **Bislang kommen knapp drei Viertel aller großen KI-Foundation-Modelle aus den USA, weitere fünfzehn Prozent aus China. Welche Maßnahmen sollte die Politik in Deutschland und Europa vor diesem Hintergrund mit Blick auf Förderung und Stärkung des Ökosystems von Generativer KI vorrangig ergreifen, wenn wir verhindern wollen, vollständig in Abhängigkeit von außereuropäischen Foundation-Modellen zu geraten und nur noch als Einkäufer dieser Modelle am Ende der Wertschöpfungskette agieren zu können?**

Antwort BSI:
- keine Antwort -

14) **Welche Regeln braucht es aus Ihrer Sicht beim AI-Act für Generative KI, konkret was die Pflichten für Entwickler von Foundation-Modellen zur Informationsweitergabe innerhalb der Lieferkette angeht, welche Vor- und Nachteile gehen mit solchen Pflichten einher und ab welcher Schwelle sollten für Anwendungen, die auf Generativer KI basieren, die Hoch-Risiko-Regeln greifen, welche im AI-Act vorgesehen sind?**

Antwort BSI:

Moderne KI-Systeme sind in der Regel sehr komplex. Sie bestehen aus mehreren Hard- und Software-Komponenten, deren Entwickler, Produzenten, Anbieter und Nutzer über mehrere Wirtschafträume - mit potentiell unterschiedlichen gesetzlichen Vorgaben - verteilt sein können. Um beispielsweise Nachweise über die Sicherheit eines in einer Cloud betriebenen KI-Produktes zu erbringen, bedarf es entsprechender Nachweise von beteiligten Hardware- und Software-Providern, Plattform Providern, Framework Providern, Data Providern, AI Service Providern und ggf. Weiteren. Um diese Komplexität in der gesamten Lieferkette und über den gesamten Lebenszyklus eines KI-Systems zu beherrschen, sind entsprechende internationale Standards notwendig. Der Entwicklung solcher Standards sollte dementsprechend die höchste Priorität eingeräumt werden.

Aus dem AI Act folgen mittelbar auch Pflichten für Entwickler von Foundation-Modellen, wenn diese Modelle direkt oder indirekt über abgeleitete Anwendungsfälle in den Bereich der Hochrisiko-KI fallen. Durch den risikobasierten Ansatz in der Anwendung bedarf es keiner Unterscheidung zwischen generativen und nicht generativen KI-Systemen und daher auch keiner unterschiedlichen Schwellwerte.

15) **Welche Initiativen gibt es insbesondere bei großen Sprachmodellen (LLMs) für die Entwicklung europäischer Modelle und wie bewerten Sie die Möglichkeiten und Grenzen von Private Public Partnerships in diesem Bereich?**

Antwort BSI:
- keine Antwort -

16) **Welches sind nach Ihrer Einschätzung die nächsten Entwicklungsstufen von generativer KI, nach Sprach- und Videomodellen (Stichpunkte KI-Agenten, Embodied AI etc.) und wo liegen hier die größten Chancen für unsere Gesellschaft und Wirtschaft?**

Antwort BSI:
- keine Antwort -

17) **Inwiefern unterscheidet sich die Verteilung von Vor- und Nachteilen durch GPAI zwischen unterschiedlichen Bevölkerungsgruppen (sowohl innerhalb nationaler Gesellschaften als global betrachtet mit Blick auf den globalen Süden/Norden) aufgrund der nachfolgend aufgezählten Aspekte:**
    - **Unterschiedliche Zugangsmöglichkeiten zur Technologie (z.B. wegen unterschiedlicher technischer, materieller, bildungs- u.a. anderer Voraussetzungen)**
    - **Unterschiedliche Repräsentanz in Trainingsdaten (z.B. Gesundheitsdaten von Frauen vs. Männern, von Weißen vs. PoC, afrikanische Sprachen vs. Englisch etc.)**
    - **Unterschiedliche Betroffenheit durch stereotype Zuschreibungen und Diskriminierungen (z.B. aufgrund von Geschlecht oder Ethnie)**
    - **Unterschiedliche Belastung durch den von KI-Systemen verursachten Ressourcenverbrauch**

    und wie wäre eine gerechtere Verteilung der Vor- und Nachteile erreichbar?

Antwort BSI:

Fairness und Diskriminierungsfreiheit, insbesondere im Hinblick auf die zugrundeliegenden Trainingsdaten, ist ein grundsätzliches Anliegen bei der Gestaltung und Regulierung von GPAI (General Purpose AI). Kriterien und Prüfmethoden zur Sicherstellung der Fairness sind aktiver Gegenstand der Forschung. Im Rahmen des digitalen Verbraucherschutzes fördert

das BSI die technische Entwicklung in diesem Bereich. Einige Aspekte davon werden in der Antwort zu Frage 6 behandelt.

18) Sollte generative KI als Mehrzweck-KI grundsätzlich als Hochrisiko-KI im Sinne der europäischen KI-Verordnung eingestuft werden, um höhere Standards zu erfüllen und für wie sinnvoll/umsetzbar halten Sie Regulierungsoptionen für generative KI wie Transparenzpflichten zu Trainingsdaten und Trainingsprozesen, die Verpflichtung zum Risikoassessment durch Bereitsteller einer GPAI und dessen Veröffentlichung, sichtbare oder unsichtbare Kennzeichnungen von allen oder bestimmten KI-generierten Inhalten, das Recht auf Überprüfbarkeit der Diskriminierungsfreiheit und den Zugang für Forscher:innen und andere diskutierte Optionen?

Antwort BSI:

Auf die Antworten zu Frage 1 und Frage 14 wird verwiesen.

Aus Sicht des BSI ist eine risikobasierte und anwendungsspezifische Bewertung von KI-Systemen im Sinne eines horizontalen Rahmenwerks und darauf abgestimmter vertikaler Standards sinnvoll. Eine Unterscheidung zwischen generativer und nicht-generativer KI ist vor diesem Hintergrund nicht zielführend.

***