SPITCH 1 YEARS

Whitepaper über die Auswirkungen des EU-KI-Gesetzes auf die Nutzung von KI zu Geschäftszwecken

www.spitch.ai

März 2024

Relevanz und potenzielle Auswirkungen des EU-KI-Gesetzes für Unternehmen

Einige der führenden Unternehmen nutzen bereits die enormen Möglichkeiten, die sich durch GenAI[1] und LLMs[2] ergeben. Sie wollen sie mit den Datenschutz- und Sicherheitsanforderungen in Einklang bringen, die auch von Endkunden als grundlegend angesehen werden. Konversationsfähige KI-Lösungen für den Kundenservice, wie z. B. Virtual Assistants (Text- und Sprach-Chatbots), Voice Biometrics und Speech Analytics, die Spitch unter anderem anbietet, machen einen signifikanten Unterschied, was Unternehmensleistung und Datensicherheit angeht. Das Gleichgewicht zwischen den geschäftlichen Vorteilen innovativer Technologien, einer wirksamen Regulierung zur Risikokontrolle und der sozialen Verantwortung scheint für eine nachhaltige Einführung von KI entscheidend zu sein.

Die Einführung des KI-Gesetzes der EU bietet den ersten umfassenden Rechtsrahmen für die KI-Governance. Es wird sich zwangsläufig sowohl auf die Anbieter als auch auf die Anwender einiger Formen von GenAI und biometrischer Gesichtserkennung auswirken, die bei unregulierter und unkontrollierter Verwendung potenzielle Risiken bergen können. Wir glauben, dass die Auswirkungen der neuen EU-Vorschriften für Anwendungsfälle in verschiedenen Branchen und im öffentlichen Sektor weitgehend positiv sein werden, mit einem relativ geringen Einfluss auf die Anbieter von geschäftsfeldspezifischer konversationsfähiger KI. Dennoch müssen alle Anbieter und Betreiber von KI-Systemen einige neue Anforderungen beachten.

Es wird erwartet, dass diese Anforderungen von den europäischen Normungsgremien weiterentwickelt werden. Folglich werden die nationalen Behörden sicherstellen müssen, dass die Unternehmen die neuen KI-Governance- und Risikomanagement-Anforderungen sowie –die dazugehörigen Standards einhalten, und gleichzeitig prüfen, inwieweit detailliertere sektorale Leitlinien erforderlich sind.

Zusammenfassung des KI-Gesetzes der EU*

Das KI-Gesetz stuft KI nach dem Risiko ein:

- Unannehmbare Risiken sind verboten (z. B. soziale Bewertungssysteme und manipulative KI).
- Der grösste Teil des Textes befasst sich mit KI-Systemen mit hohem Risiko, die reguliert sind.
- Ein kleinerer Abschnitt befasst sich mit KI-Systemen mit begrenztem Risiko, die geringeren Transparenzpflichten unterliegen: Entwickler und Betreiber müssen sicherstellen, dass die Endnutzer wissen, dass sie mit KI interagieren (Chatbots und Deepfakes).
- Das minimale Risiko ist nicht reguliert (einschliesslich der meisten KI-Anwendungen, die derzeit auf dem EU-Binnenmarkt erhältlich sind, wie z. B. KI-fähige Videospiele und Spam-Filter – zumindest im Jahr 2021; das ändert sich mit generativer KI).

Die meisten Verpflichtungen treffen die Anbieter (Entwickler) von risikoreichen KI-Systemen.

- Diejenigen, die beabsichtigen, hochriskante KI-Systeme in der EU in Verkehr zu bringen oder in Betrieb zu nehmen, unabhängig davon, ob sie in der EU oder in einem Drittland ansässig sind.
- Und auch Anbieter aus Drittländern, bei denen der Output des Hochrisiko-KI-Systems in der EU verwendet wird.

Benutzer sind natürliche oder juristische Personen, die ein KI-System beruflich einsetzen, nicht aber betroffene Endbenutzer.

- **Die Benutzer (Anwender) von risikoreichen KI-Systemen** haben einige Verpflichtungen, wenn auch weniger als die Anbieter (Entwickler).
- Das gilt für Benutzer in der EU und für Benutzer in Drittländern, wenn der Output des KI-Systems in der EU verwendet wird.

* High-Level-Zusammenfassung des KI-Gesetzes

Spitch ist der Ansicht, dass es für bestehende und potenzielle Kunden und Partner nützlich sein kann, die Antworten auf einige der am häufigsten gestellten Fragen zu den neuen Anforderungen zu lesen. Diese haben wir im Folgenden zusammengefasst.

[1] *GenAI – Generative künstliche Intelligenz*

[2] *LLMs – Grosse Sprachmodelle*

Fragen und Antworten

1. Was sind die verbotenen KI-Systeme nach dem KI-Gesetz der EU? Fällt eines der konversationsfähigen KI-Systeme für den Kundenservice unter die Kategorie „verboten“?

Die folgenden Arten von KI-Systemen sind nach dem KI-Gesetz „verboten“ (Titel II, Art. 5)[3]:

- Einsatz **unterschwelliger, manipulativer oder trügerischer Techniken** zur Verzerrung des Verhaltens und zur Beeinträchtigung einer fundierten Entscheidungsfindung, wodurch erheblicher Schaden verursacht wird.
- **Ausnutzung von Schwachstellen** im Zusammenhang mit Alter, Behinderung oder sozioökonomischen Umständen, um das Verhalten zu verzerren und erheblichen Schaden zu verursachen.
- **Biometrische Kategorisierungssysteme** Rückschlüsse auf sensible Attribute (Herkunft, politische Meinungen, Gewerkschaftszugehörigkeit, religiöse oder philosophische Überzeugungen, Sexualleben oder sexuelle Orientierung), mit Ausnahme der Kennzeichnung oder Filterung rechtmässig erworbener biometrischer Datensätze oder wenn Strafverfolgungsbehörden biometrische Daten kategorisieren.
- **Soziales Scoring**, d. h. die Bewertung oder Klassifizierung von Einzelpersonen oder Gruppen auf der Grundlage von sozialem Verhalten oder persönlichen Merkmalen, was zu einer nachteiligen oder ungünstigen Behandlung dieser Personen führt.
- **Bewertung des Risikos, dass eine Person Straftaten begeht** ausschliesslich auf der Grundlage von Profilen oder Persönlichkeitsmerkmalen, es sei denn, sie wird zur Ergänzung menschlicher Einschätzungen verwendet, die auf objektiven, überprüfbaren Fakten beruhen, die in direktem Zusammenhang mit kriminellen Aktivitäten stehen.
- **Aufbau von Gesichtserkennungsdatenbanken** durch ungezieltes Auslesen von Gesichtsbildern aus dem Internet oder CCTV-Aufnahmen.
- **Aufzeigen von Emotionen an Arbeitsplätzen oder in Bildungseinrichtungen**, ausser aus medizinischen oder Sicherheitsgründen.
- **Biometrische Fernidentifikation in „Echtzeit“ (RBI) in öffentlich zugänglichen Räumen für die Strafverfolgung**, ausser:
  - für die Suche nach vermissten Personen, Entführungsopfern und Menschen, die Opfer von Menschenhandel oder sexueller Ausbeutung geworden sind;
  - für die Vermeidung einer erheblichen und unmittelbaren Bedrohung des Lebens oder eines vorhersehbaren Terroranschlags; oder
  - für die Identifizierung von Verdächtigen bei schweren Straftaten (z. B. Vergewaltigung, Mord, bewaffneter Raubüberfall, Drogen- und illegaler Waffenhandel, organisierte Kriminalität und Umweltkriminalität usw.).

Keines der allgemein akzeptierten konversationsfähigen KI-Systeme, die üblicherweise für geschäftliche Zwecke im Kundendienst eingesetzt werden, fällt unter diese Kategorien.

Zusammenfassung des KI-Gesetzes der EU*

Allzweck-KI (GPAI):

- Alle Anbieter von GPAI-Modellen müssen technische Unterlagen und Gebrauchsanweisungen bereitstellen, die Urheberrechtsrichtlinie einhalten und eine Zusammenfassung der für die Ausbildung verwendeten Inhalte veröffentlichen.
- Anbieter von GPAI-Modellen mit freien und offenen Lizenzen müssen nur das Urheberrecht einhalten und die Zusammenfassung der Trainingsdaten veröffentlichen, sofern sie kein systemisches Risiko darstellen.
- Alle Anbieter von GPAI-Modellen, die ein systemisches Risiko darstellen – ob offen oder geschlossen –, müssen auch Modellbewertungen und Gegentests durchführen, schwerwiegende Vorfälle verfolgen und melden und Cybersicherheitsschutzmassnahmen gewährleisten.

* High-Level-Zusammenfassung des KI-Gesetzes

* High-Level-Zusammenfassung des KI-Gesetzes

[3] Quelle: https://artificialintelligenceact.eu/high-level-summary/

Viele Unternehmen, insbesondere Banken, haben damit begonnen, die Echtzeitform der stimmbiometrischen Identitätsprüfung für ihre Kunden einzusetzen. In den meisten Fällen dient es dazu, einen von Menschen gesteuerten Prozess der Identitätsüberprüfung während des Live-Gesprächs mit einem Kundendienstmitarbeiter zu ergänzen. Während des Anrufs wird die Stimme des Kunden mit dem zuvor erstellten Stimmprofil (eine mathematische Darstellung der Stimmmerkmale des Benutzers) verglichen, um sicherzustellen, dass die Stimme des Anrufers übereinstimmt.

Es gibt auch „hybride" stimmbiometrische Verifizierungsverfahren, bei denen der Anrufer eine zufällig generierte Zahl oder ein Wort wiederholt und die Live-Stimme mit dem Stimmabdruck abgeglichen wird. Diese Methode trägt dazu bei, dass die Überprüfung schneller erfolgt.

Sofern sie nicht von Strafverfolgungsbehörden im öffentlichen Raum eingesetzt werden, fallen solche ferngesteuerten stimmbiometrischen Authentifizierungsmethoden nicht unter die Kategorie der verbotenen KI-Systeme.

Wenn jedoch konversationsfähige KI-Systeme die Erkennung und Ableitung von Emotionen nutzen, um automatisierte Entscheidungen in einer Arbeitsumgebung zu treffen oder Anrufer auf der Grundlage der erkannten Emotionen zu kategorisieren, können solche Systeme unter die Kategorie der verbotenen KI-Systeme gemäss dem KI-Gesetz der EU fallen.

Wird die Stimmungsanalyse jedoch als Teil eines umfassenderen KI-Systems für den Kundenservice eingesetzt, um das Kundenfeedback zu verstehen und die Servicequalität zu verbessern, ohne dass automatisierte Entscheidungen getroffen werden, die sich direkt auf Einzelpersonen auswirken, würde dies nicht als verbotener Anwendungsfall gelten. Es ist wichtig anzumerken, dass das KI-Gesetz der EU nicht die Verwendung von Stimmungsanalysen im Allgemeinen verbietet, sondern nur bestimmte Anwendungen, die als risikoreich oder unethisch angesehen werden. Unternehmen sollten sorgfältig prüfen, wie sie die Stimmungsanalyse einsetzen wollen und sicherstellen, dass sie nicht unter die im Gesetz genannten verbotenen Anwendungsfälle fällt.

2. Welche KI-Systeme werden im Rahmen des EU-KI-Gesetzes als „hochriskant" eingestuft? Besteht die Möglichkeit, dass die Nutzung von KI-Systemen aufgrund dieser Klassifizierung und der Notwendigkeit, zusätzliche Anforderungen zu erfüllen, Risiken für Unternehmen mit sich bringt?

Einige KI-Systeme gelten nach dem KI-Gesetz (**Titel III**)[4] als „hochriskant". Für die Anbieter dieser Systeme gelten zusätzliche Anforderungen. KI-Systeme mit hohem Risiko (**Art. 6**) sind jene, die als Sicherheitsbauteil oder Produkt verwendet werden, das unter EU-Rechtsvorschriften in **Anhang II** **und** einer Konformitätsbewertung durch einen Dritten gemäss diesen Anhang-II-Rechtsvorschriften unterzogen werden muss; **ODER** unter **Anhang III** Anwendungsfällen (siehe Kasten unten), ausser wenn:

- das KI-System eine enge prozedurale Aufgabe ausführt;
- es das Ergebnis einer zuvor durchgeführten menschlichen Tätigkeit verbessert;
- es Entscheidungsmuster oder Abweichungen von früheren Entscheidungsmustern ermittelt und nicht dazu gedacht ist, die zuvor durchgeführte menschliche Bewertung ohne ordnungsgemässe menschliche Überprüfung zu ersetzen oder zu beeinflussen; oder
- es eine vorbereitende Aufgabe für eine Bewertung durchführt, die für die in Anhang III aufgeführten Anwendungsfälle relevant ist.

Wichtig ist, dass das KI-System immer dann als risikoreich gilt, wenn es **Profile von Personen** erstellt, d. h. personenbezogene Daten automatisiert verarbeitet, um verschiedene Aspekte des Lebens einer Person zu bewerten, wie z. B. Arbeitsleistung, wirtschaftliche Situation, Gesundheit, Vorlieben, Interessen, Zuverlässigkeit, Verhalten, Standort oder Bewegung.

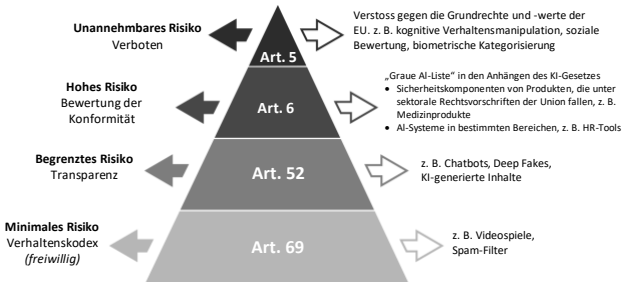
Anbieter, die der Meinung sind, dass ihr KI-System, das unter **Anhang III** fällt, kein hohes Risiko darstellt, **müssen eine solche Bewertung dokumentieren**, bevor sie es in Verkehr bringen oder in Betrieb nehmen.

Für Anbieter von KI-Systemen mit hohem Risiko gibt es eine Reihe von Anforderungen, die in Artikel 8–25 des Gesetzes aufgeführt sind. Dazu gehören die Verpflichtung zur Einrichtung eines **Risikomanagementsystems** während des gesamten Lebenszyklus des KI-Systems mit hohem Risiko; die Durchführung einer angemessenen **Datenverwaltung**, die Erstellung einer **technischen Dokumentation** zum Nachweis der Konformität und die Bereitstellung von Informationen für die Behörden zur Bewertung dieser Konformität; die Gewährleistung, dass das KI-System mit hohem Risiko eine automatische **Aufzeichnung** von Ereignissen vorsieht, die für die Identifizierung von Risiken auf nationaler Ebene und von wesentlichen Änderungen während des gesamten Lebenszyklus des Systems relevant sind.

Zu den weiteren Anforderungen gehören die Entwicklung von **Anweisungen für die Verwendung** durch nachgelagerte Bereitsteller, um diesen die Einhaltung der Vorschriften zu ermöglichen und die Umsetzung einer **menschlichen Aufsicht** zu ermöglichen; die Gewährleistung eines angemessenen Niveaus an **Genauigkeit, Robustheit und Cybersicherheit** sowie die Einrichtung eines **Qualitätsmanagementsystems** zur Gewährleistung der Einhaltung der Vorschriften.

[4] Quelle: https://artificialintelligenceact.eu/high-level-summary/

Einige nicht verbotene biometrische Systeme, wie z. B. die von Spitch bereitgestellten Voice-Biometrics-Systeme, können unter die im Anhang III aufgeführten Systeme fallen. Andere konversationsfähige KI-Systeme wie Chatbots, Speech Analytics, Knowledge Bases usw. können in die Kategorien „begrenztes Risiko“ oder „minimales Risiko“ fallen, mit minimalen oder keinen zusätzlichen Anforderungen, wie das nachstehende Diagramm[5] zeigt:



3. Wie sieht der Zeitplan für die Einhaltung des KI-Gesetzes der EU aus?

Spitch bietet seinen Kunden Beratungsdienste an, um die vollständige und rechtzeitige Einhaltung der gesetzlichen Bestimmungen in Übereinstimmung mit dem KI-Gesetz der EU und anderen gesetzlichen Rahmenbedingungen in einem bestimmten Land oder einer bestimmten Region, einschliesslich der Datenschutzbestimmungen, zu gewährleisten.

Unsere Anwälte helfen den Kunden bei der Formulierung von Kundenverträgen auf der Grundlage der in den einzelnen Ländern geltenden rechtlichen Anforderungen, die den Einsatz von KI-Systemen regeln. Diese sollten vor der Einführung des Projekts von den Juristen auf der Kundenseite genehmigt werden.

Die nachstehende Grafik[5] veranschaulicht den Zeitplan für die Umsetzung gemäss dem Amtsblatt der EU. Für die meisten konversationsfähigen KI-Lösungen, z. B. nicht verbotene biometrische Daten und Chatbots, müssen die zusätzlichen Anforderungen innerhalb von 12 bis 36 Monaten ab Mai 2024, wenn das KI-Gesetz der EU in Kraft tritt, erfüllt werden..

Anwendungsfälle Anhang III

Nicht verbotene biometrische Daten: Biometrische Fernidentifikationssysteme, ausgenommen biometrische Überprüfungen, die bestätigen, dass eine Person diejenige ist, die sie vorgibt zu sein. Biometrische Kategorisierungssysteme, die auf sensible oder geschützte Attribute oder Merkmale schliessen lassen. Systeme zur Erkennung von Emotionen.

Kritische Infrastrukturen: Sicherheitskomponenten bei der Verwaltung und dem Betrieb kritischer digitaler Infrastrukturen, im Strassenverkehr und bei der Versorgung mit Wasser, Gas, Wärme und Strom.

Bildung und Berufsbildung: KI-Systeme, die den Zugang, die Zulassung oder die Zuweisung zu Bildungs- und Berufsbildungseinrichtungen auf allen Ebenen bestimmen. Bewertung von Lernergebnissen, einschliesslich derer, die zur Steuerung des Lernprozesses der Studierenden verwendet werden. Beurteilung des angemessenen Bildungsniveaus für eine Person. Überwachung und Erkennung von unzulässigem Schülerverhalten während der Prüfungen.

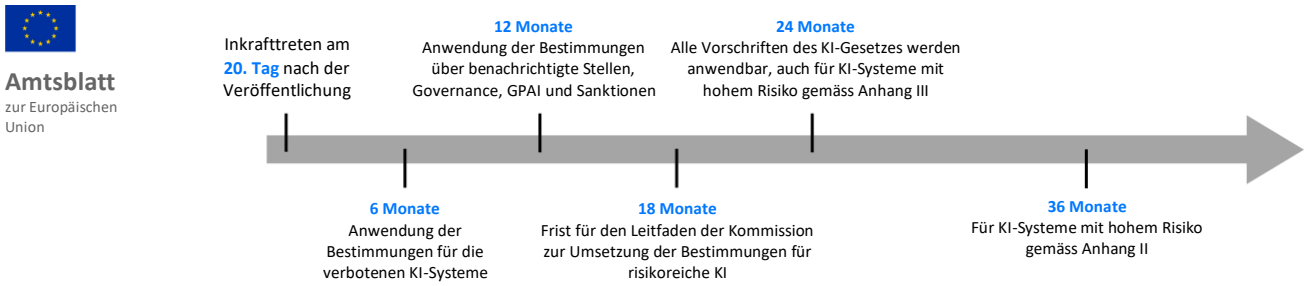
Beschäftigung, Arbeitskräftemanagement und Zugang zur Selbstständigkeit: KI-Systeme, die für die Personalbeschaffung oder -auswahl eingesetzt werden, insbesondere für gezielte Stellenanzeigen, die Analyse und Filterung von Bewerbungen und die Bewertung von Kandidaten. Beförderung und Beendigung von Verträgen, Zuweisung von Aufgaben auf der Grundlage von Persönlichkeitsmerkmalen und Verhalten sowie Überwachung und Bewertung der Leistung.

Zugang zu und Inanspruchnahme von wesentlichen öffentlichen und privaten Dienstleistungen: KI-Systeme, die von öffentlichen Behörden zur Beurteilung der Anspruchsberechtigung auf Leistungen und Dienste, einschliesslich deren Zuweisung, Kürzung, Entzug oder Rückforderung, verwendet werden. Bewertung der Kreditwürdigkeit, ausser bei der Aufdeckung von Finanzbetrug. Bewertung und Klassifizierung von Notrufen, einschliesslich der Festlegung von Prioritäten für den Einsatz von Polizei, Feuerwehr, medizinischer Hilfe und dringenden Patiententriage-Diensten. Risikobewertung und Tarifierung in der Kranken- und Lebensversicherung.

Gesetzesverfolgung: KI-Systeme zur Bewertung des Risikos einer Person, Opfer einer Straftat zu werden. Polygraphen. Bewertung der Zuverlässigkeit von Beweisen bei strafrechtlichen Ermittlungen oder Strafverfolgungen. Bewertung des Risikos einer Person, straffällig zu werden oder erneut straffällig zu werden, nicht nur auf der Grundlage eines Profils oder der Bewertung von Persönlichkeitsmerkmalen oder früherem kriminellen Verhalten. Profiling bei der Ermittlung, Untersuchung oder Verfolgung von Straftaten.

Migration, Asyl und Grenzkontrolle: Polygraphen. Bewertungen der irregulären Migration oder der Gesundheitsrisiken. Prüfung von Anträgen auf Asyl, Visum und Aufenthaltstitel und damit verbundene Beschwerden im Zusammenhang mit der Anspruchsberechtigung. Aufspüren, Erkennen oder Identifizieren von Personen, mit Ausnahme der Überprüfung von Reisedokumenten.

Justizverwaltung und demokratische Prozesse: KI-Systeme, die bei der Erforschung und Auslegung von Fakten und der Anwendung des Rechts auf konkrete Sachverhalte oder bei der alternativen Streitbeilegung eingesetzt werden. Beeinflussung der Ergebnisse von Wahlen und Referenden oder des Abstimmungsverhaltens, mit Ausnahme von Ergebnissen, die nicht direkt mit den Menschen interagieren, wie z. B. Instrumente zur Organisation, Optimierung und Strukturierung politischer Kampagnen.



[5] *Quelle:* *https://www.engage.hoganlovells.com/knowledgeservices/news/the-eu-ai-act-an-impact-analysis-part-1*

4. Gibt es irgendwelche Bestimmungen des KI-Gesetzes der EU, die es unmöglich oder zu riskant machen würden, die von Spitch angebotenen KI-Lösungen zu nutzen?

Kurz gesagt: Nein.

Bestimmte Bestimmungen sehen vor, dass der Anbieter der KI-Systeme, d. h. Spitch, zusätzliche Anforderungen erfüllen muss, nicht aber die Kunden. Spitch stellt sicher, dass alle Lösungen, die Speech Biometrics in Echtzeit, Stimmungsanalyse und Emotionserkennung nutzen, die auf GPT-4 oder andere KI-Modelle für allgemeine Zwecke zur Zusammenfassung und Kategorisierung zurückgreifen, sowie die im KI-Gesetz der EU als „mit begrenztem Risiko“ eingestuften Systeme, wie Virtual Assistants (KI-Chatbots), innerhalb der vorgeschriebenen Fristen vollständig konform sind.

5. Was sind die neuen Verpflichtungen der Kunden, die die von Spitch bereitgestellten KI-Systeme einsetzen, gemäss dem KI-Gesetz der EU?

Gemäss Artikel 52 des Gesetzes müssen die Betreiber eines Systems zur Erkennung von Emotionen oder eines biometrischen Kategorisierungssystems die natürlichen Personen, die dem System ausgesetzt sind, über den Betrieb des Systems informieren und die personenbezogenen Daten im Einklang mit der Verordnung (EU) 2016/679, der Verordnung (EU) 2016/1725 und der Richtlinie (EU) 2016/280, soweit anwendbar, verarbeiten.

Ausserdem müssen die Betreiber eines KI-Systems, das Text erzeugt oder manipuliert, der zur Information der Öffentlichkeit über Angelegenheiten von öffentlichem Interesse veröffentlicht wird, offenlegen, dass der Text künstlich erzeugt oder manipuliert wurde.

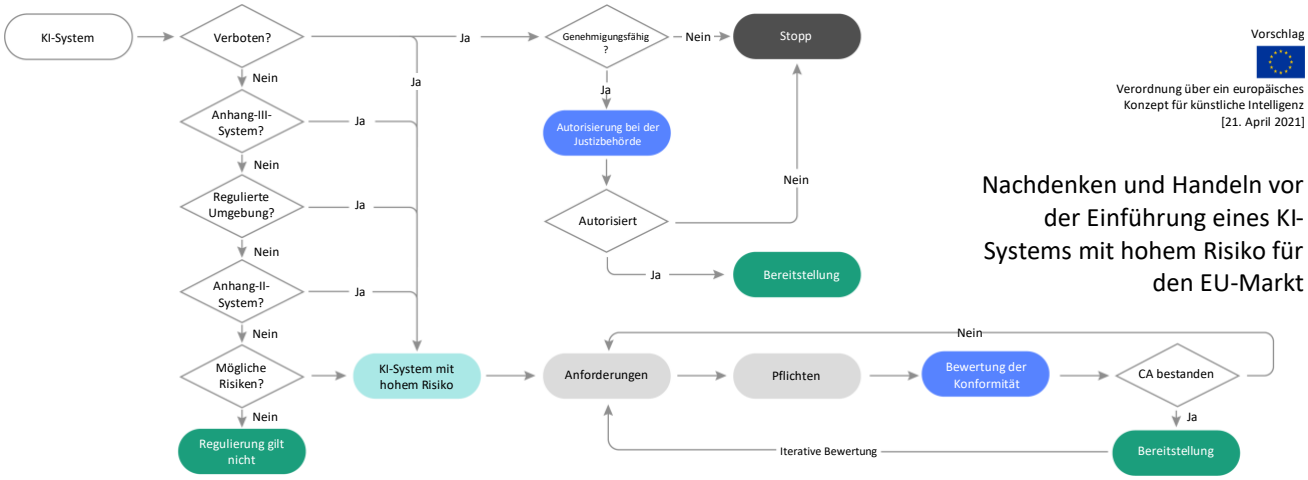
Die oben genannten Informationen müssen den betroffenen natürlichen Personen spätestens zum Zeitpunkt der ersten Interaktion oder Exposition in klarer und erkennbarer Weise zur Verfügung gestellt werden. Die Informationen müssen den geltenden Zugänglichkeitsanforderungen entsprechen.

Wie die Anforderungen umgesetzt werden, kann je nach Lösung von Spitch variieren. Unser Professional Services Team unterstützt Sie bei der Implementierung, um sicherzustellen, dass Sie die neuesten Vorschriften einhalten.

6. Ist die Gewinnung von Leistungsdaten in der Arbeitsumgebung für die Zwecke des Qualitätsmanagements nach dem Gesetz zulässig? Welche Vorgehensweise stellt Gesetzeskonformität sicher?

Die Extraktion solcher Daten durch KI-Systeme ist zulässig, es sei denn, sie dient dazu, automatisierte Entscheidungen ohne die Beteiligung von Menschen zu treffen, auf Emotionen am Arbeitsplatz zu schliessen, auf sensible Eigenschaften zu schliessen (Herkunft), politische Meinungen, Gewerkschaftszugehörigkeit, religiöse oder philosophische Überzeugungen, Sexualleben oder sexuelle Orientierung) oder soziale Bewertungen vorzunehmen, d. h. Einzelpersonen oder Gruppen auf der Grundlage von sozialem Verhalten oder persönlichen Merkmalen zu bewerten oder zu klassifizieren, was zu einer nachteiligen oder ungünstigen Behandlung dieser Personen führt.

Um festzustellen, ob die Bestimmungen der neuen Verordnung in Ihrem speziellen Fall anwendbar sind oder ob eines der KI-Systeme unter die Kategorie „hohes Risiko“ gemäss dem KI-Gesetz der EU fällt, wird das folgende Entscheidungsverfahren[6] empfohlen, um die Einhaltung der Bestimmungen sicherzustellen:



Vorschlag

Verordnung über ein europäisches Konzept für künstliche Intelligenz
[21. April 2021]

Nachdenken und Handeln vor der Einführung eines KI-Systems mit hohem Risiko für den EU-Markt

[6] *Quelle: https://futurium.ec.europa.eu/en/european-ai-alliance/open-discussion/more-visual-guide-proposed-eu-artificial-intelligence-act?language=da*