Prüfen Sie Ihr GenKI-Projekt auf Risiken.

Basierend auf "GAIRA Light"

Sie wollen eine Anwendung auf Basis von generativer KI umsetzen? Wenn ja, dann beachten Sie die folgenden Punkte.

Projekt: _____

Datum: _____

1. Gibt es mutmasslich hohe Risiken für das Unternehmen?

Können Sie nicht alle folgenden Punkte bestätigen, dann birgt Ihr Projekt vermutlich hohe Risiken und Sie sollten eine vertiefte Risikobewertung durchführen (z.B. GAIRA Comprehensive). Sonst zu Schritt 2.

- Wir erstellen/trainieren das von uns benutzte KI-Modell nicht selbst ("RAG" nicht gemeint) ☐
- Die Anwendung trifft keine Entscheide über andere Menschen, die diese für wichtig halten ☐
- Die Anwendung interagiert nicht mit vielen Menschen in Bezug auf sensible Themen ☐
- Wir würden keine Schritte prüfen, selbst wo so was gegen uns/mit unseren Daten genutzt würde ☐
- Die Anwendung hat kein hohes Potenzial negativer Schlagzeilen in den Medien ("Shitstorm") ☐
- Was wir planen, ist weder verboten nach dem EU AI Act noch ein "Hoch-Risiko"-KI-System ☐
- Die Anwendung wird nur für eigene Zwecke genutzt und nicht auch Dritten angeboten ☐
- Die Anwendung erfordert keine grosse Investition und ist nicht von strategischer Bedeutung ☐

2. Haben Sie die typischen Risiken bei generativer KI im Griff?

Können Sie jeden der folgenden Punkte angesichts der geplanten oder getroffenen Risikomassnahmen bestätigen? Falls nicht, sollten Sie die Rechts- und Reputationsrisiken des Vorhabens intern besprechen.

- Wir nutzen vertrauliche oder geschützte Daten Dritter im Einklang mit unseren Vertragspflichten ☐
- Unsere Verträge mit Dienstleistern entsprechen dem Datenschutz und unseren Vertragspflichten ☐
- Unser Input/Output wird nicht von den Dienstleistern überwacht oder wir sind einverstanden ☐
- Die Dienstleister nutzen unseren Input/Output nicht für sich selbst oder wir sind einverstanden ☐
- Die Anwendung kann keine vertraulichen oder personenbezogenen Daten an andere preisgeben ☐
- Es gibt kein systematisches Bearbeiten sensibler personenbezogener Daten oder Profiling ☐
- Wir haben eine Rechtsgrundlage für die Bearbeitung von personenbezogenen Daten (wo nötig) ☐
- Wir verwenden personenbezogene Daten nur für ihren ursprünglichen (oder erwarteten) Zweck ☐
- Wir erheben/nutzen nur die für den Zweck nötigen personenbezogenen Daten ☐
- Wir bewahren personenbezogene Daten in Bezug auf die Anwendung nur solange wie nötig auf ☐
- Wir können Betroffenenbegehren (z.B. Auskunft, Berichtigung, Widerspruch, Löschung) erfüllen ☐
- Wir haben angemessene Massnahmen zur Informationssicherheit und Geschäftsfortführung ☐
- Wir informieren andere über den Einsatz von KI, wo dies für ihre Interaktion mit uns relevant ist ☐
- Öffentlichkeit und Betroffene werden unseren Einsatz von KI weder unfair noch unpassend finden ☐
- Wir haben Massnahmen gegen fehlerhaften oder problematischen Output (z.B. Bias) ☐
- Unsere Anwendung wird vor einem Einsatz ausgiebig getestet, auch gegen böswillige Angriffe ☐
- Unser KI-Einsatz verursacht keine ungewollten negativen Folgen (z.B. Schaden, Diskriminierung) ☐
- Unser KI-Einsatz kann nicht als Ausnutzen von Schwächen der betroffenen Personen gelten ☐
- Wo unsere KI wichtige Entscheidungen treffen oder beeinflussen könnte, wird sie beaufsichtigt ☐
- Wir verwenden ein anerkanntes, qualitativ gutes KI-Modell, dessen Verhalten wir verstehen ☐
- Unser KI-Einsatz wird niemanden täuschen oder irreführen ☐
- Wir haben Massnahmen, um KI-Missverhalten zu erkennen, aufzuzeichnen & darauf zu reagieren ☐
- Unser KI-Einsatz respektiert die Würde und das Selbstbestimmungsrecht der davon Betroffenen ☐
- Die Nutzer unserer KI werden in ihrer korrekten Nutzung unterwiesen, geschult und überwacht ☐
- Wir sehen keine weiteren unkontrollierten Probleme im Zusammenhang mit unserem KI-Einsatz ☐

Verarbeiten Sie personenbezogene Daten, dann prüfen Sie, ob auch eine Datenschutz-Folgenabschätzung, eine Anpassung der Datenschutzerklärung oder des Verzeichnisses der Verarbeitungstätigkeiten (ROPA) nötig ist. Halten Sie sich auch an die weiteren Vorgaben und Prozesse für den KI-Einsatz in Ihrem Betrieb.

Ihre Risikobewertung sauber dokumentieren? Sie können hierzu das GAIRA Excel nutzen (kostenlos: https://vischerlnk.com/gaira). Es erlaubt sowohl eine Beurteilung mit den obigen Punkten ("Light") als auch eine umfassende Risikobewertung und -dokumentation.

Autor: David Rosenthal (drosenthal@vischer.com) 15.1.24 Updates: vischerlnk.com/update

GenKI-Risikobewertungen mit GAIRA Light.

Für die Bewertung Ihres KI-Projekts mit GAIRA Light sind die folgenden neun Schritte nötig. Rechnen Sie mit 90 Minuten, es sei denn, Ihr Projekt erfordert eine umfassendere Risikobewertung (z.B. falls ein mutmasslich hohes Risiko ermittelt werden sollte). Die Fragen sollten Sie selbst beantworten können, ausser möglicherweise jene, welche die Verträge mit Ihren Dienstleistern betreffen. Fragen Sie hierzu Ihre Fachleute.

Bezeichnung der KI-Anwendung

Beschreibung der geplanten KI-Anwendung

Verbotene Aktivität oder Hoch-Risiko-KI-System?

Prüfung, ob eine umfassendere Bewertung nötig ist

→ **GAIRA Comprehensive**

Fragen zur Eingrenzung der Risikobereiche

Fragen zur Einhaltung bestimmter Compliance-Vorgaben

14-25 Risikofragen, die auf etwaige Probleme hinweisen

Kennen Sie die Antwort nicht oder erscheint eine Warnung, dann sprechen Sie mit Ihren Experten (Rechtsdienst, Datenschutzstelle, CISO etc.) über die nötigen Schritte (z.B. weitere Massnahmen)

Datenschutz-Folgenabschätzung, falls notwendig

Abschluss der Bewertung

Entscheid über die Implementierung

90 Min.

Laden Sie das GAIRA-Excel kostenlos unter https://vischerlnk.com/gaira herunter. Es enthält Arbeitsblätter für die "Light"-Version und eine umfassendere Risikobewertung. Enthalten ist auch ein Werkzeug, dass Ihnen aufzeigt, ob Ihre Anwendung unter den EU AI fällt und welche Rolle Sie haben.

VISCHER SWISS LAW AND TAX