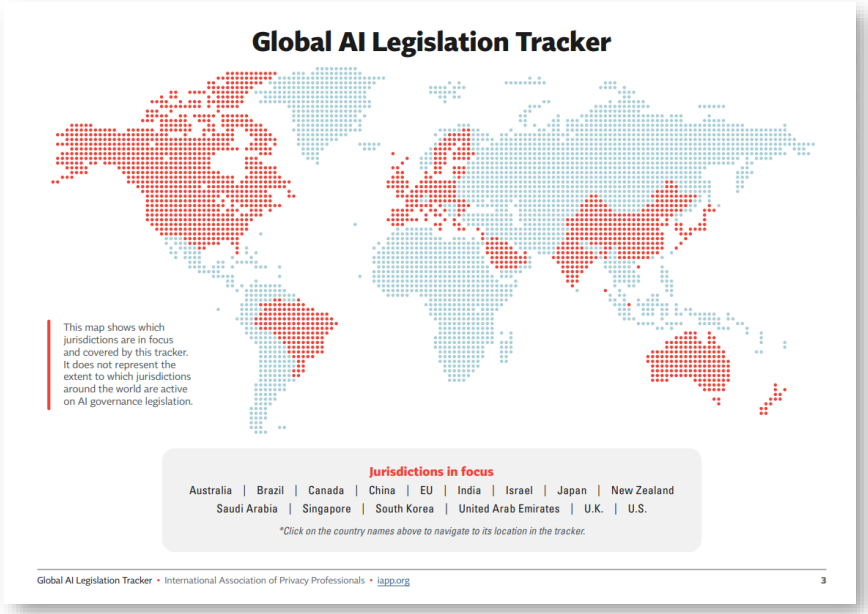
VISCHER

Generative KI.

GAIRA – ein Werkzeug zur Risikobeurteilung

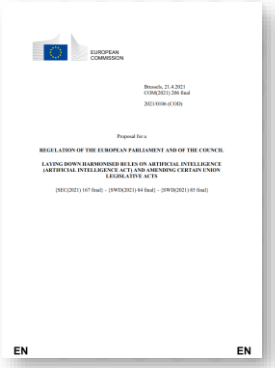
David Rosenthal, Partner, VISCHER AG
27. November 2023

Was Gesetzgeber mit KI tun



IAPP Global AI Legislation Tracker (https://iapp.org/resources/article/global-ai-legislation-tracker/)

Executive Order 14110 (USA)



AI Act (Entwurf) (EU)

Wie Betriebe mit KI umgehen tun (sollten)

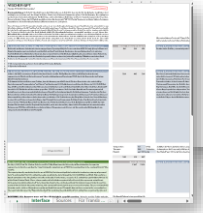
Rechtliche und ethische Vorgaben für KI-Anwendungen in Betrieben (hier: Die 11 Grundsätze für KI für Organisationen von VISCHER)

Vorgaben für einzelne Anwendungen

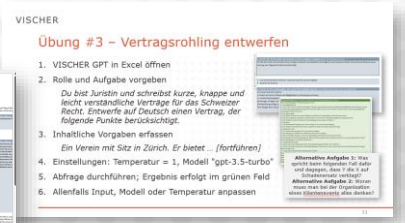
Hilfestellungen für Mitarbeitende (hier: VISCHER GenKI Cheat-Sheet)

GAIRA zur Risikobeurteilung vor KI-Projekten

Verzeichnis: Records of AI Activities (ROAIA)

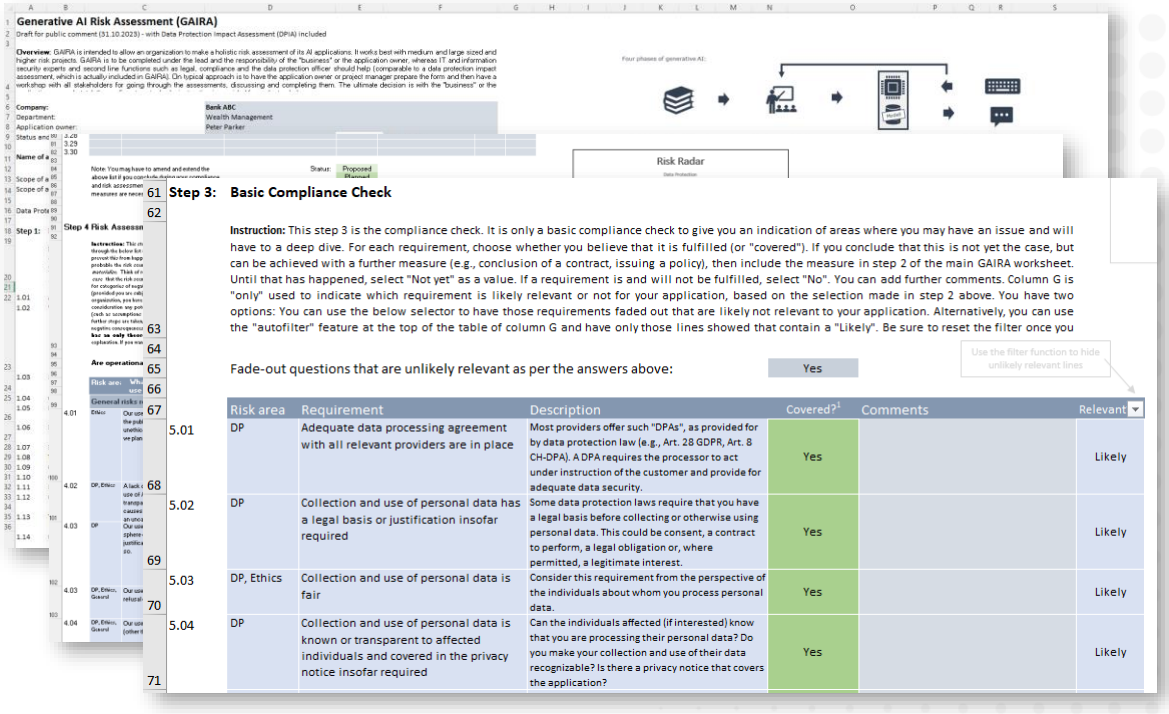


Geprüfte Tools



Schulung der Mitarbeitenden im sicheren und nutzbringenden Einsatz von GenKI

Tool für ein GenAI Risk Assessment (GAIRA)



Generative AI Risk Assessment (GAIRA)

Draft for public comment (31.10.2023) - with Data Protection Impact Assessment (DPIA) included

Overview: GAIRA is intended to allow an organization to make a holistic risk assessment of its AI applications. It works best with medium and large sized and higher risk projects. GAIRA is to be completed under the lead and the responsibility of the "business" or the application owner, whereas IT and information security experts and second line functions such as legal, compliance and the data protection officer should help (comparable to a data protection impact assessment, which is actually included in GAIRA). On typical approach is to have the application owner or project manager prepare the form and then have a workshop with all stakeholders for going through the assessments, discussing and completing them. The ultimate decision is with the "business" or the

Company: Bank ABC
Department: Wealth Management
Application owner: Peter Parker

Four phases of generative AI:

Risk Radar

Status: Proposed

Step 3: Basic Compliance Check

Instruction: This step 3 is the compliance check. It is only a basic compliance check to give you an indication of areas where you may have an issue and will have to a deep dive. For each requirement, choose whether you believe that it is fulfilled (or "covered"). If you conclude that this is not yet the case, but can be achieved with a further measure (e.g., conclusion of a contract, issuing a policy), then include the measure in step 2 of the main GAIRA worksheet. Until that has happened, select "Not yet" as a value. If a requirement is and will not be fulfilled, select "No". You can add further comments. Column G is "only" used to indicate which requirement is likely relevant or not for your application, based on the selection made in step 2 above. You have two options: You can use the below selector to have those requirements faded out that are likely not relevant to your application. Alternatively, you can use the "autofilter" feature at the top of the table of column G and have only those lines showed that contain a "Likely". Be sure to reset the filter once you

Fade-out questions that are unlikely relevant as per the answers above: Yes

Use the filter function to hide unlikely relevant lines

	Risk area	Requirement	Description	Covered?[1]	Comments	Relevant
5.01	DP	Adequate data processing agreement with all relevant providers are in place	Most providers offer such "DPAs", as provided for by data protection law (e.g., Art. 28 GDPR, Art. 8 CH-DPA). A DPA requires the processor to act under instruction of the customer and provide for adequate data security.	Yes		Likely
5.02	DP	Collection and use of personal data has a legal basis or justification insofar required	Some data protection laws require that you have a legal basis before collecting or otherwise using personal data. This could be consent, a contract to perform, a legal obligation or, where permitted, a legitimate interest.	Yes		Likely
5.03	DP, Ethics	Collection and use of personal data is fair	Consider this requirement from the perspective of the individuals about whom you process personal data.	Yes		Likely
5.04	DP	Collection and use of personal data is known or transparent to affected individuals and covered in the privacy notice insofar required	Can the individuals affected (if interested) know that you are processing their personal data? Do you make your collection and use of their data recognizable? Is there a privacy notice that covers the application?	Yes		Likely

- Risiko-Assessment beinhaltet zugleich eine Datenschutz-Folgenabschätzung (DSFA)
- Folgt hinsichtlich der Methodik dem Prinzip einer DSFA
- Separate Compliance-Checkliste
- ROAIA-Vorlage

Kostenlos abrufbar unter https://bit.ly/gaira

Tool für eine Datenschutz-Folgenabschätzung

Datenschutz-Folgenabschätzung (DSFA)
Version 25.9.2023 for public comment - Private CH-DSG/DSGV

Hinweis: Eine Anleitung zum Ausfüllen dieser DSFA und zur KI-gestützten Ausfüllhilfe (optional, nur in der Version des Excels mit Makros) findet sich am Ende dieses Arbeitsblatt

Unternehmen (Verantwortlicher): Musterfirma AG

Abteilung:
Verantwortlich intern:
Status der DSFA:
Name des Vorhabens:
Aktivität gemäss Bearbeitu

1. Beschreibung de
1.01 In welchem Bereic bzw. welcher Gesc
1.02 Was vorgesehen i
1.03 Welche Interesser
1.04 Welche Mittel und
1.05 Welche Dritten an
1.06 Welche Daten bea
1.07 Wessen Daten bei
1.08 Wo überall Daten
1.09 Wann die Datenb
1.10 Weitere Besonder

2. Erforderlichkeit /
2.01 Warum die Daten Sicht der betroffer eingreifende Vorg
2.02 Warum die Datenbearbeitung datensparsam, zeitlich auf das nötige begrenzt und auch sonst verhältnismässig ist:

Die Datenbearbeitung ist datensparsam, da nur der Stimmabdruck, die ID der Person und Tonaufnahmen gespeichert werden, die für die Identitätsverifizierung notwendig sind. Die Datenbearbeitung ist zeitlich begrenzt, da der Stimmabdruck bei jedem Anruf neu erstellt und nicht länger als nötig gespeichert wird. Die Datenbearbeitung ist verhältnismässig, da sie zur Sicherheit der Anrufer im Call-Center beiträgt und die einzigen Daten bearbeitet werden, die dafür erforderlich sind.

4. Risiken von negativen Folgen für die betroffenen Personen, die trotz der obigen Massnahmen verbleiben

10 Risiken vorschlagen (überschreibt bisherige Werte)

Hinweis: Falls die ermittelten Risiken als zu hoch erscheinen oder sich zeigt, dass es noch weitere Massnahmen zur Minimierung gibt, sollten diese oben unter Ziff. 3 eingetragen werden und bei der Risikobeurteilung hier berücksichtigt werden.

	Mögliche unerwünschte negative Folgen Weiteres Risiko vorschlagen*	Was wir dagegen tun Massnahmen vorschlagen* / Aus obigen formulieren*	Wie wir das Restrisiko einschätzen Risikobeurteilung vorschlagen*	Mögliche Folgen für die Person	Eintrittswahrscheinlichkeit (alles in allem)	Risiko (1-16)
4.01	Personendaten des Vorhabens gelangen wegen eines Fehlers oder absichtlich an **unbefugte Dritte**. Diese missbrauchen sie zum Schaden der betroffenen Personen.	- Berechtigungskonzept: Da wir nur autorisierten Personen Zugriff auf die Personendaten geben, wird das Risiko von unbefugtem Zugriff und Missbrauch reduziert. - Schulung: Durch Schulungen stellen wir sicher, dass die Mitarbeitenden die Lösung korrekt und sicher nutzen, was das Risiko von Fehlern und Missbrauch verringert. - Zugriffskontrolle: Durch die Beschränkung des Zugriffs auf autorisierte Personen können wir Missbräuche und unbefugte Nutzung der Personendaten verhindern. - Verschlüsselung "at rest": Die Verschlüsselung der Personendaten in unserem System schützt vor unautorisiertem Zugriff, falls jemand physischen Zugriff auf die Speichermedien erhält. - Datenlöschungsfunktionen: Durch die Möglichkeit, nicht mehr benötigte Personendaten zu löschen oder zu anonymisieren, minimieren wir das Risiko eines unbefugten Zugriffs auf diese Daten.	Das konkrete Restrisiko für die betroffene Person besteht darin, dass ihre Personendaten aufgrund eines Fehlers oder absichtlich an unbefugte Dritte gelangen könnten. Diese könnten die Daten dann zum Schaden der betroffenen Person nutzen, beispielsweise für Identitätsdiebstahl oder Missbrauch in sozialen Medien. Die Wahrscheinlichkeit dieses Szenarios ist jedoch insgesamt gering, da strenge Sicherheitsmassnahmen wie Zugriffskontrollen und Verschlüsselung implementiert wurden.	Substanziell	Tief	Mittel (6)
4.02	Personendaten des Vorhabens gelangen wegen eines Fehlers oder absichtlich an eine **unbefugte interne Person**.					

Datenschutz-Folgenabschätzung
vud.ch/dsfa

https://vud.ch/dsfa

VISCHER

Danke für Ihre Aufmerksamkeit!

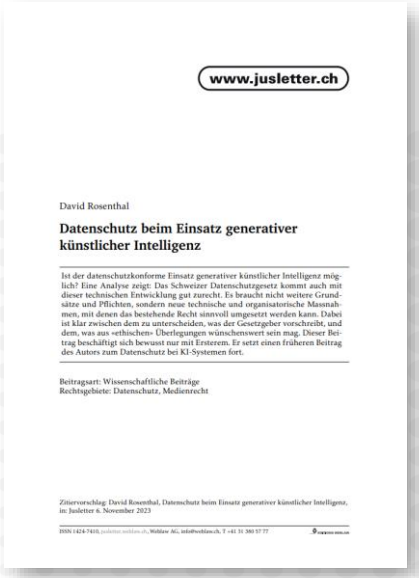
Fragen: drosenthal@vischer.com

Zürich
Schützengasse 1
Postfach
8021 Zürich, Schweiz
T +41 58 211 34 00

Basel
Aeschenvorstadt 4
Postfach
4010 Basel, Schweiz
T +41 58 211 33 00

Genf
Rue du Cloître 2-4
Postfach
1211 Genf 3, Schweiz
T +41 58 211 35 00

www.vischer.com


www.jusletter.ch

David Rosenthal

Datenschutz beim Einsatz generativer künstlicher Intelligenz

Ist der datenschutzkonforme Einsatz generativer künstlicher Intelligenz möglich? Eine Analyse zeigt: Das Schweizer Datenschutzgesetz kommt auch mit dieser technischen Entwicklung gut zurecht. Es braucht nicht weitere Grundsätze und Pflichten, sondern neue technische und organisatorische Massnahmen, mit denen das bestehende Recht sinnvoll umgesetzt werden kann. Dabei ist klar zwischen dem zu unterscheiden, was der Gesetzgeber vorschreibt, und dem, was aus «ethischen» Überlegungen wünschenswert sein mag. Dieser Beitrag beschäftigt sich bewusst nur mit Ersterem. Er setzt einen früheren Beitrag des Autors zum Datenschutz bei KI-Systemen fort.

Beitragsart: Wissenschaftliche Beiträge
Rechtsgebiete: Datenschutz, Medienrecht

Zitiervorschlag: David Rosenthal, Datenschutz beim Einsatz generativer künstlicher Intelligenz, in: Jusletter 6. November 2023

https://www.rosenthal.ch/downloads/Rosenthal-Jusletter-GenKI-Datenschutz.pdf