VISCHER

Data Protection and AI.

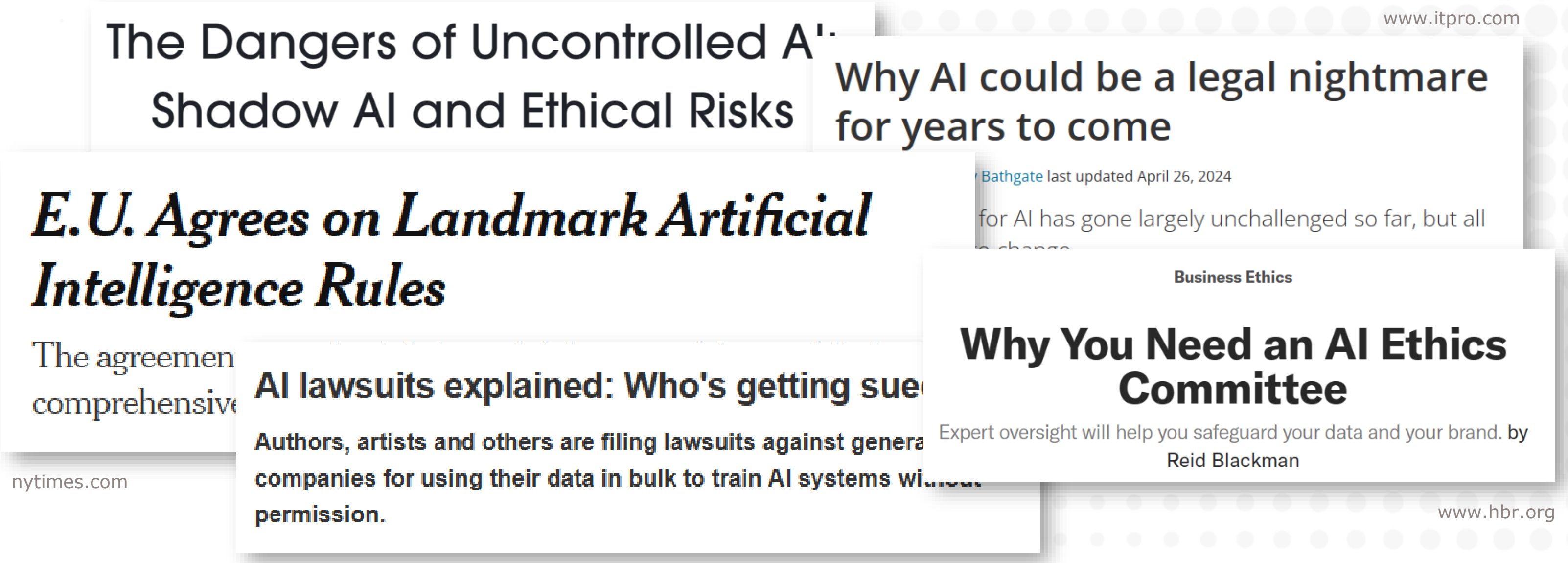
Impact on compliance and the work of the ethics & compliance teams

David Rosenthal, Partner, VISCHER Ltd.

September 27, 2024

We all know the headlines …

securiti.ai



The Dangers of Uncontrolled A

Shadow AI and Ethical Risks

www.itpro.com

Why AI could be a legal nightmare for years to come

Bathgate last updated April 26, 2024

for AI has gone largely unchallenged so far, but all

E.U. Agrees on Landmark Artificial Intelligence Rules

The agreemen

comprehensive

nytimes.com

AI lawsuits explained: Who's getting sue

Authors, artists and others are filing lawsuits against genera

companies for using their data in bulk to train AI systems wi

permission.

www.techtarget.com

Business Ethics

Why You Need an AI Ethics Committee

Expert oversight will help you safeguard your data and your brand. by Reid Blackman

www.hbr.org

What does that mean for us?

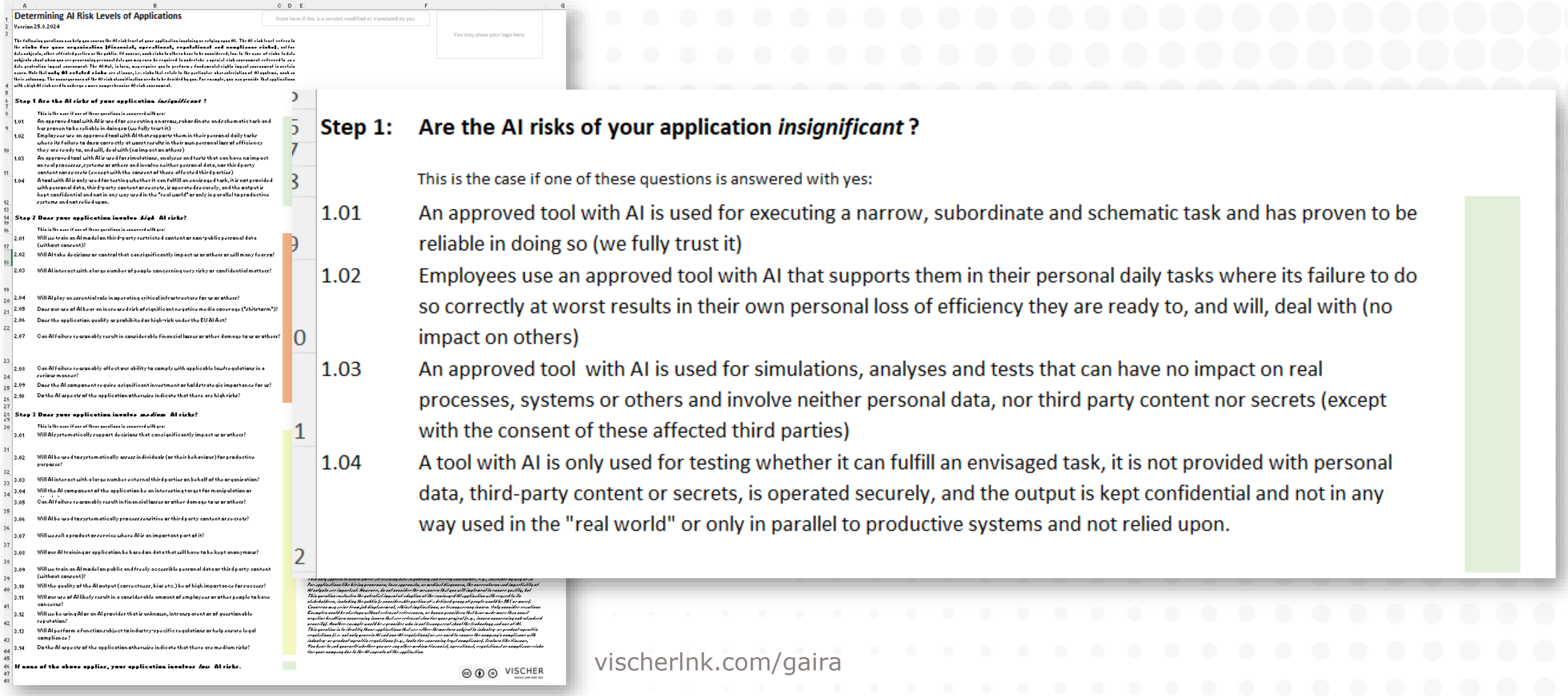
- More work
- More learning
- More concerns

- There is **good** and **bad** news for us …

More work

- The **bad** news
  - Internal pressure to move ahead with AI projects
  - Not only from IT, but also from the top and the business
  - "Tell us whether this is legally ok and what we need to do."
- The **good** news
  - The wave is already leveling off (see Gartner Hype Cycle)
  - Most projects are only "proof of concepts" that will not survive
  - Get basic infrastructure approved, so there is safe playground
  - Provide for "sandboxes" for experimental use of AI
  - Establish a risk-based approach; also define "irrelevant" AI
  - But: Going through detailed AI risks helps identifying gaps

Risk-based approach?



Determining AI Risk Levels of Applications

Version 25.8.2024

State here if this is a version modified or translated by you

You may place your logo here

The following questions can help you assess the AI risk level of your application involving or relying upon AI. The AI risk level refers to the **risks for your organization (financial, operational, reputational and compliance risks)**, not for data subjects, other affected parties or the public. Of course, such risks to others have to be considered, too. In the case of risks to data subjects about whom you are processing personal data you may even be required to undertake a special risk assessment referred to as a data protection impact assessment. The AI Act, in turn, may require you to perform a fundamental rights impact assessment in certain cases. Note that **only AI related risks** are at issue, i.e. risks that relate to the particular characteristics of AI systems, such as their autonomy. The consequences of the AI risk classification needs to be decided by you. For example, you can provide that applications with a high AI risk need to undergo a more comprehensive AI risk assessment.

Step 1: Are the AI risks of your application *insignificant* ?

This is the case if one of these questions is answered with yes:

1.01 An approved tool with AI is used for executing a narrow, subordinate and schematic task and has proven to be reliable in doing so (we fully trust it)

1.02 Employees use an approved tool with AI that supports them in their personal daily tasks where its failure to do so correctly at worst results in their own personal loss of efficiency they are ready to, and will, deal with (no impact on others)

1.03 An approved tool with AI is used for simulations, analyses and tests that can have no impact on real processes, systems or others and involve neither personal data, nor third party content nor secrets (except with the consent of these affected third parties)

1.04 A tool with AI is only used for testing whether it can fulfill an envisaged task, it is not provided with personal data, third-party content or secrets, is operated securely, and the output is kept confidential and not in any way used in the "real world" or only in parallel to productive systems and not relied upon.

Step 2 Does your application involve *high* AI risks?

This is the case if one of these questions is answered with yes:

2.01 Will we train an AI model on third-party restricted content or non-public personal data (without consent)?

2.02 Will AI take decisions or control that can significantly impact us or others or will many fear so?

2.03 Will AI interact with a large number of people concerning very risky or confidential matters?

2.04 Will AI play an essential role in operating critical infrastructure for us or others?

2.05 Does our use of AI bear an increased risk of significant negative media coverage ("shitstorm")?

2.06 Does the application qualify as prohibited or high-risk under the EU AI Act?

2.07 Can AI failure reasonably result in considerable financial losses or other damage to us or others?

2.08 Can AI failure reasonably affect our ability to comply with applicable law/regulations in a serious manner?

2.09 Does the AI component require a significant investment or hold strategic importance for us?

2.10 Do the AI aspects of the application otherwise indicate that there are high risks?

Step 3 Does your application involve *medium* AI risks?

This is the case if one of these questions is answered with yes:

3.01 Will AI systematically support decisions that can significantly impact us or others?

3.02 Will AI be used to systematically assess individuals (or their behaviour) for productive purposes?

3.03 Will AI interact with a large number external third parties on behalf of the organization?

3.04 Will the AI component of the application be an interesting target for manipulation or

3.05 Can AI failure reasonably result in financial losses or other damage to us or others?

3.06 Will AI be used to systematically process sensitive or third party content or secrets?

3.07 Will we sell a product or service where AI is an important part of it?

3.08 Will our AI training or application be based on data that will have to be kept anonymous?

3.09 Will we train an AI model on public and freely accessible personal data or third party content (without consent)?

3.10 Will the quality of the AI output (correctness, bias etc.) be of high importance for success?

3.11 Will our use of AI likely result in a considerable amount of employees or other people to have concerns?

3.12 Will we be using AI or an AI provider that is unknown, intransparent or of questionable reputation?

3.13 Will AI perform a function subject to industry-specific regulations or help ensure legal compliance ?

3.14 Do the AI aspects of the application otherwise indicate that there are medium risks?

If none of the above applies, your application involves *low* AI risks.

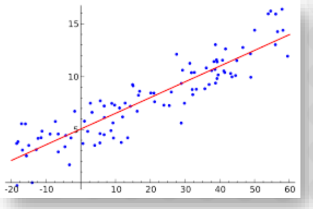
[illegible]

VISCHER

vischerlnk.com/gaira

What is AI after all?

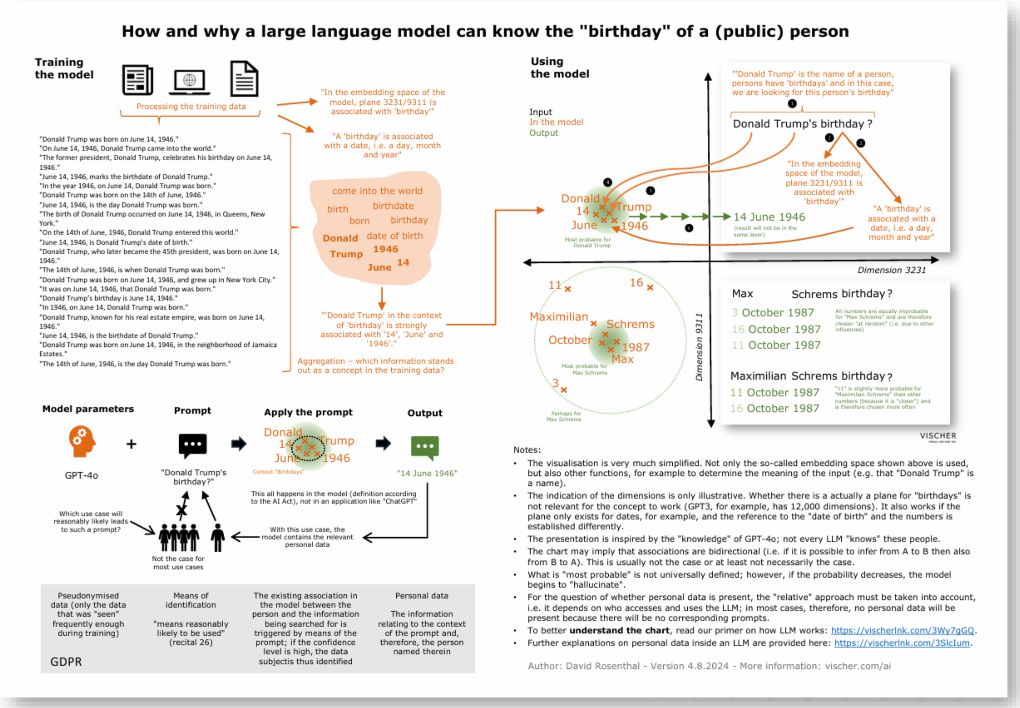
- As per the EU **AI Act** "a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments"
- The only practically relevant element is **"autonomy"**
  - In simple terms: An IT system that has been **trained** on how to decide, not only used programmed logic …
- **But**: The definition is flawed …
  - Every copy machine is AI (OCR); what about linear regression?
  - Complexity or automated decision-making should be the test



More learning

- The **bad** news
  - To assess the compliance of AI applications, we need to understand them, at least to a certain extent
  - It is not only about technology, but also mathematics …
  - AI is developing so fast, it is almost impossible to stay up-to-date
- The **good** news
  - You can (and should) push back to your team
  - The basic concepts are not all too complicated
  - There are a lot of materials to explain the "magic" …
  - And: Even the experts do not fully understand why advanced AI is able to do what it does …

Large language models?



How and why a large language model can know the "birthday" of a (public) person

Training the model

Processing the training data

"Donald Trump was born on June 14, 1946."
"On June 14, 1946, Donald Trump came into the world."
"The former president, Donald Trump, celebrates his birthday on June 14, 1946."
"June 14, 1946, marks the birthdate of Donald Trump."
"In the year 1946, on June 14, Donald Trump was born."
"Donald Trump was born on the 14th of June, 1946."
"June 14, 1946, is the day Donald Trump was born."
"The birth of Donald Trump occurred on June 14, 1946, in Queens, New York."
"On the 14th of June, 1946, Donald Trump entered this world."
"June 14, 1946, is Donald Trump's date of birth."
"Donald Trump, who later became the 45th president, was born on June 14, 1946."
"The 14th of June, 1946, is when Donald Trump was born."
"Donald Trump was born on June 14, 1946, and grew up in New York City."
"It was on June 14, 1946, that Donald Trump was born."
"Donald Trump's birthday is June 14, 1946."
"In 1946, on June 14, Donald Trump was born."
"Donald Trump, known for his real estate empire, was born on June 14, 1946."
"June 14, 1946, is the birthdate of Donald Trump."
"Donald Trump was born on June 14, 1946, in the neighborhood of Jamaica Estates."
"The 14th of June, 1946, is the day Donald Trump was born."

"In the embedding space of the model, plane 3231/9311 is associated with 'birthday'"

"A 'birthday' is associated with a date, i.e. a day, month and year"

"'Donald Trump' in the context of 'birthday' is strongly associated with '14', 'June' and '1946'."

Aggregation – which information stands out as a concept in the training data?

Using the model

Input
In the model
Output

"'Donald Trump' is the name of a person, persons have 'birthdays' and in this case, we are looking for this person's birthday"

Donald Trump's birthday ?

"In the embedding space of the model, plane 3231/9311 is associated with 'birthday'"

"A 'birthday' is associated with a date, i.e. a day, month and year"

Max Schrems birthday ?

3 October 1987
16 October 1987
11 October 1987

All numbers are equally improbable for "Max Schrems" and are therefore chosen "at random" (i.e. due to other influences)

Maximilian Schrems birthday ?

11 October 1987
16 October 1987

"11" is slightly more probable for "Maximilian Schrems" than other numbers (because it is "closer") and is therefore chosen more often

VISCHER

Model parameters + **Prompt** → **Apply the prompt** → **Output**

This all happens in the model (definition according to the AI Act), *not* in an application like "ChatGPT"

With this use case, the model contains the relevant personal data

Which use case will reasonably likely leads to such a prompt?

Not the case for most use cases

Pseudonymised data (only the data that was "seen" frequently enough during training)	Means of identification "means reasonably likely to be used" (recital 26)	The existing association in the model between the person and the information being searched for is triggered by means of the prompt; if the confidence level is high, the data subjectis thus identified	Personal data The information relating to the context of the prompt and, therefore, the person named therein

GDPR

Notes:

- The visualisation is very much simplified. Not only the so-called embedding space shown above is used, but also other functions, for example to determine the meaning of the input (e.g. that "Donald Trump" is a name).
- The indication of the dimensions is only illustrative. Whether there is a actually a plane for "birthdays" is not relevant for the concept to work (GPT3, for example, has 12,000 dimensions). It also works if the plane only exists for dates, for example, and the reference to the "date of birth" and the numbers is established differently.
- The presentation is inspired by the "knowledge" of GPT-4o; not every LLM "knows" these people.
- The chart may imply that associations are bidirectional (i.e. if it is possible to infer from A to B then also from B to A). This is usually not the case or at least not necessarily the case.
- What is "most probable" is not universally defined; however, if the probability decreases, the model begins to "hallucinate".
- For the question of whether personal data is present, the "relative" approach must be taken into account, i.e. it depends on who accesses and uses the LLM; in most cases, therefore, no personal data will be present because there will be no corresponding prompts.
- To better **understand the chart**, read our primer on how LLM works: https://vischerlnk.com/3Wy7gGQ.
- Further explanations on personal data inside an LLM are provided here: https://vischerlnk.com/3SlcIum.

Author: David Rosenthal - Version 4.8.2024 - More information: vischer.com/ai

How does a large language model work and what is really stored in it?

vischerlnk.com/3Wy7gGQ

Does a large language model contain personal data?

vischerlnk.com/3SlcIum

vischerlnk.com/3TuhLcg

More concerns

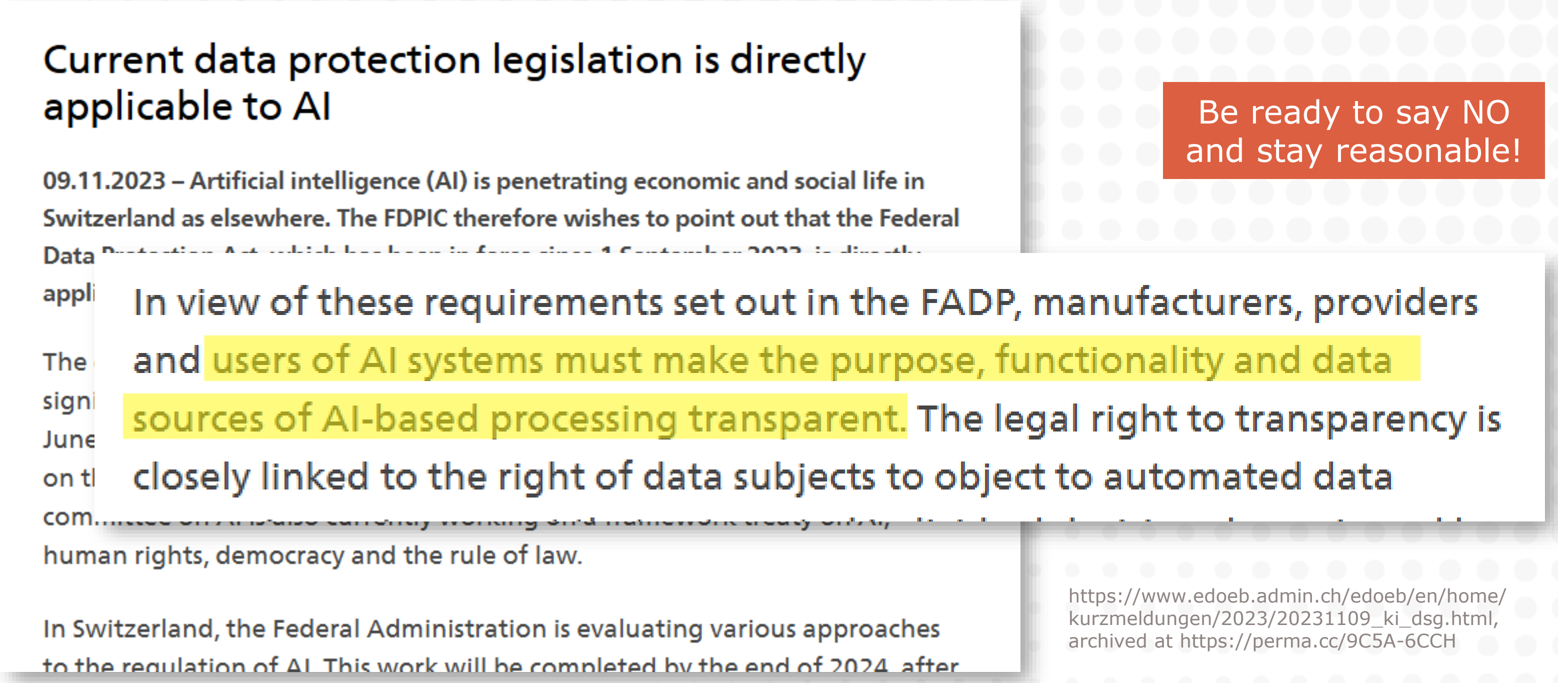
- The **bad** news
  - Since (in particular generative) AI is a black box for many of us, we do not know how to assess the risks and apply our rules
  - The supervisory authorities face the same challenges
  - We need more information or apply stricter rules than necessary
  - E.g., on automated decisions, non-discrimination, transparency
- The **good** news
  - We will get used to AI (now that we are aware of it) and treat it as any other means to do things (much like with the Internet)
  - Companies realize that "ethical" use of AI is often a buzz-word
  - The existing rules continue to apply and mostly work very well

TIME
THE STRANGE NEW WORLD OF THE
INTERNET
Battles on the frontiers of cyberspace

July 25, 1994 (time.com, Cover: James Porto)

VISCHER

FDPIC Expectations re AI



Current data protection legislation is directly applicable to AI

09.11.2023 – Artificial intelligence (AI) is penetrating economic and social life in Switzerland as elsewhere. The FDPIC therefore wishes to point out that the Federal Data Protection Act, which has been in force since 1 September 2023, is directly appli

The signi June on th com

human rights, democracy and the rule of law.

In Switzerland, the Federal Administration is evaluating various approaches to the regulation of AI. This work will be completed by the end of 2024, after

In view of these requirements set out in the FADP, manufacturers, providers and users of AI systems must make the purpose, functionality and data sources of AI-based processing transparent. The legal right to transparency is closely linked to the right of data subjects to object to automated data

Be ready to say NO and stay reasonable!

https://www.edoeb.admin.ch/edoeb/en/home/kurzmeldungen/2023/20231109_ki_dsg.html, archived at https://perma.cc/9C5A-6CCH

Data Protection Compliance #1

- **1st question:** What do we do with personal data using AI?
  - Have we informed the data subjects about this in our **privacy policy**, in particular the **purpose**?
  - Did they have to **expect** this when we received their data?
  - Is what we are doing reasonably **acceptable**? Do we remain **proportionate with regard to** the purpose? Are important decisions being made or, if asked, scrutinized by a human being?
  - Is the data that we (re-)use **correct** and complete for our purposes (to the extent that we rely on it at all)?
  - Can we guarantee the **rights of data subjects** where necessary (e.g., where information, deletion or corrections are requested)?
  - Public bodies & GDPR: Does our **legal basis cover** the use of AI, or do we have consent?

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

ISSN 1424-7410, [illegible], Weblaw AG, info@weblaw.ch, T +41 31 380 57 77

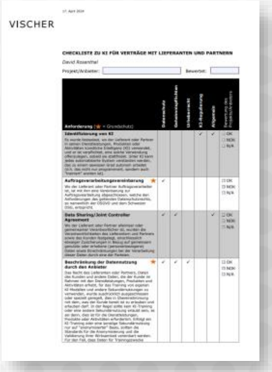
vischerlnk.com/3IdAymb

If the project could come with high risks for individuals: **DPIA**

Data Protection Compliance #2

- **2nd question:** Who do we entrust with our personal data for processing, and what does this person do with it?
  - Becomes an issue when third-party providers (Microsoft, OpenAI, Google etc.) are used
  - Check for a Data processing agreement (DPA) incl. appropriate data security, international transfer, use of your personal data for own training and abuse monitoring purposes
  - Issue #1: Lack of maturity
  - Issue #2: Lack of transparency
  - Issue #3: Constant changes
  - **Example:** Microsoft "Copilot with commercial data protection"
    - DPA finally available? Automatic activation? Abuse monitoring?

vischerlnk.com/ai-provider-check



17. April 2024

VISCHER

CHECKLISTE ZU KI FÜR VERTRÄGE MIT LIEFERANTEN UND PARTNERN

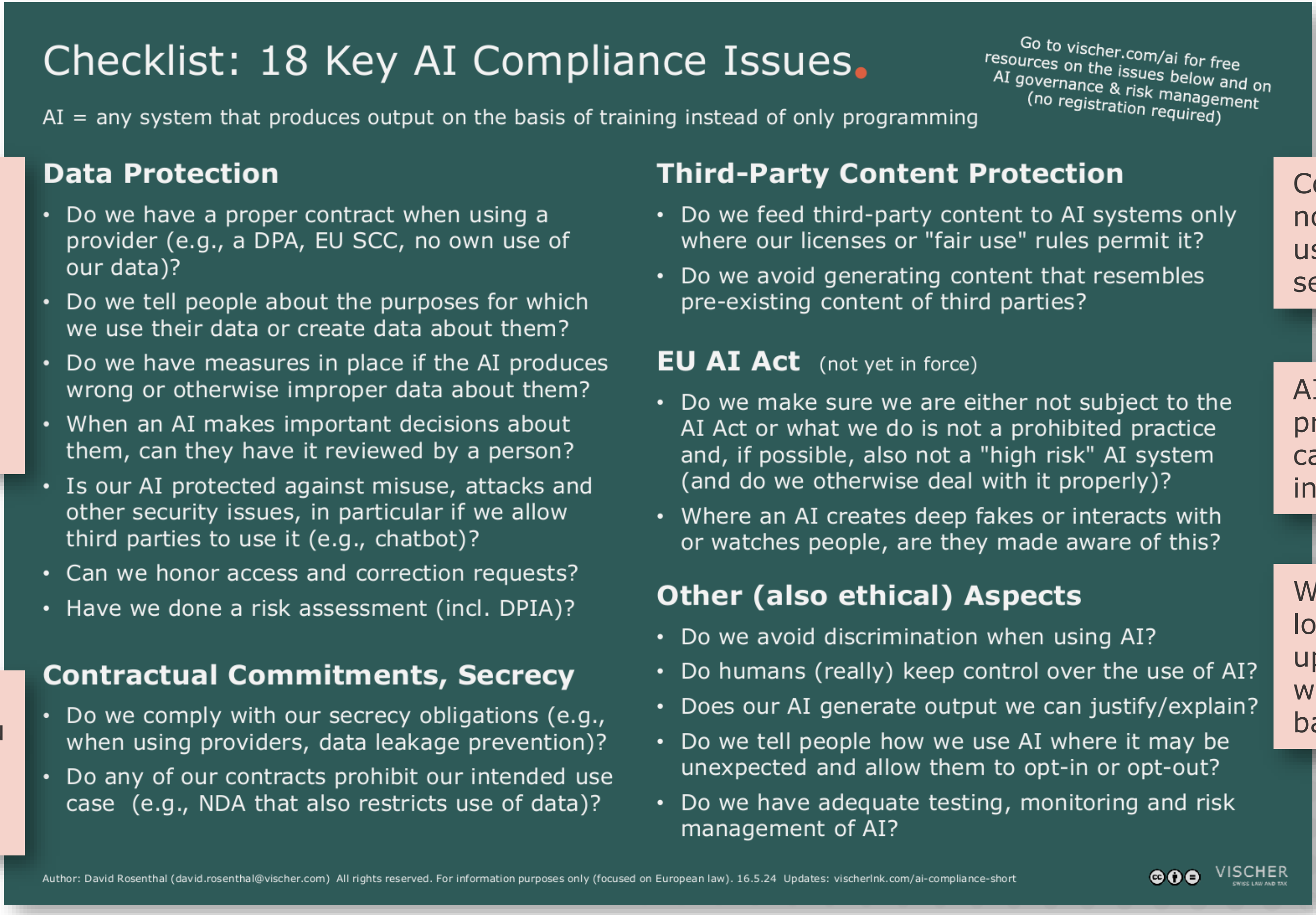
David Rosenthal

Projekt/Anbieter: Bewertet:

Anforderung (★ = Grundschutz)	Datenschutz	Geheimnispflichten	Urheberrecht	KI-Regulierung	Allgemein	Bewertung des Projekts/Anbieters
Identifizierung von KI [illegible]				✓	✓	☐ OK ☐ NOK ☐ N/A
Auftragsverarbeitungsvereinbarung ★ [illegible]	✓					☐ OK ☐ NOK ☐ N/A
Data Sharing/Joint Controller Agreement [illegible]	✓	✓			✓	☐ OK ☐ NOK ☐ N/A
Beschränkung der Datennutzung durch den Anbieter ★ [illegible]	✓	✓	✓			☐ OK ☐ NOK ☐ N/A

VISCHER

vischerlnk.com/ai-compliance-short



Checklist: 18 Key AI Compliance Issues.

Go to vischer.com/ai for free resources on the issues below and on AI governance & risk management (no registration required)

AI = any system that produces output on the basis of training instead of only programming

Data Protection

- Do we have a proper contract when using a provider (e.g., a DPA, EU SCC, no own use of our data)?
- Do we tell people about the purposes for which we use their data or create data about them?
- Do we have measures in place if the AI produces wrong or otherwise improper data about them?
- When an AI makes important decisions about them, can they have it reviewed by a person?
- Is our AI protected against misuse, attacks and other security issues, in particular if we allow third parties to use it (e.g., chatbot)?
- Can we honor access and correction requests?
- Have we done a risk assessment (incl. DPIA)?

Contractual Commitments, Secrecy

- Do we comply with our secrecy obligations (e.g., when using providers, data leakage prevention)?
- Do any of our contracts prohibit our intended use case (e.g., NDA that also restricts use of data)?

Third-Party Content Protection

- Do we feed third-party content to AI systems only where our licenses or "fair use" rules permit it?
- Do we avoid generating content that resembles pre-existing content of third parties?

EU AI Act (not yet in force)

- Do we make sure we are either not subject to the AI Act or what we do is not a prohibited practice and, if possible, also not a "high risk" AI system (and do we otherwise deal with it properly)?
- Where an AI creates deep fakes or interacts with or watches people, are they made aware of this?

Other (also ethical) Aspects

- Do we avoid discrimination when using AI?
- Do humans (really) keep control over the use of AI?
- Does our AI generate output we can justify/explain?
- Do we tell people how we use AI where it may be unexpected and allow them to opt-in or opt-out?
- Do we have adequate testing, monitoring and risk management of AI?

Author: David Rosenthal (david.rosenthal@vischer.com) All rights reserved. For information purposes only (focused on European law). 16.5.24 Updates: vischerlnk.com/ai-compliance-short

VISCHER SWISS LAW AND TAX

The usual stuff when dealing with personal data – make sure you keep control of it, in particular when using third party providers

This is critical – to whom do you disclose highly confidential customer data?

Copyright often no issue when using common sense

AI Act is about product safety; can also apply in Switzerland

What regulators love to impose upon you even without a legal basis …

You need to get AI under control? Six steps …

- A robust **data management** is the basis of all – push for it
- Regulate the tasks, authority and **responsibilities** ("AKV") in relation to AI (AI officer not needed, a committee may help)

→ See our blog at vischerlnk.com/3zjTL4R

- Decide on your AI principles and issue an **AI policy** – use it to enable users and make them "safe", not only to restrict them
- **Train** for safe, legal and responsible use of AI, and provide for AI literacy – up to the board
- **Map and track** your use of AI – and assess it (e.g., AI Act, FINMA requirements, if applicable, vischerlnk.com/3z7ZJG4)

→ See our piece on the AI Act and a cheat-sheet at vischerlnk.com/3ZkPOYh & vischerlnk.com/ai-act-uc

- Include AI in your **risk management** process – but follow a risk-based approach (most use cases will be low or medium risk, but you need to understand the unique risks of AI)

→ See our blog and our AI risk assessment tool GAIRA (also includes AI Act Check) vischerlnk.com/4bF85CW & vischerlnk.com/gaira

VISCHER

Thank you for your attention!

Questions: david.rosenthal@vischer.com

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

www.vischer.com

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

For in-depth materials on the topic visit us at vischer.com/ai