Law | Data Protection Reform

Re.13 Artificial Intelligence

Authors: Winfried Veil

Last update: 2026-08-20 07:53:35 | By: Winfried Veil

Created at: 2026-08-20 07:44:04

AI models and systems (particularly LLMs) require vast amounts of data. However, as the concept of personal data is very broad, the processing of personal data during AI training and operation is almost inevitable. Whether, and to what extent, AI training and operation involving personal data are permissible under data protection law is, however, a matter of debate.

All stages of the AI lifecycle are relevant and problematic from a data protection perspective:

(1) Collection and processing of training data, including personal data

(2) Training the AI model using personal data

(3) Use of the AI model trained with personal data

(4) Input of personal data into the AI model via prompting

(5) Output of personal data by the LLM

(6) Enrichment of the AI model's outputs by incorporating external personal data (RAG = 'retrieval-augmented generation')

(7) External persistent user profile

It is impossible to say with certainty to what extent AI providers and operators actually use personal data. There have been very few instances of regulatory intervention or legal proceedings (as at August 2026). By contrast, there are numerous confusing statements from data protection supervisory authorities. Legal uncertainty can be inferred from the comments made by stakeholders in the AI sector. It is reported that AI training is therefore often not carried out within the EU. The mere fact that legal uncertainty exists is likely to deter some from using AI. For public bodies in particular, the use of AI in a legal grey area is out of the question. Furthermore, the legally questionable use of 'shadow AI' is likely to occur on a significant scale. In its 'Data Union Strategy' dated 19 November 2025, the European Commission describes "data scarcity" as a structural bottleneck to innovation.

In their "Stuttgart Proposals for the Modernisation of Data Protection" dated 19 June 2026, the German data protection supervisory authorities call for "guidelines for AI":

> "The use of AI requires a secure framework. Alongside the guidelines already under discussion for the appropriate training of AI using personal data, there must be a guarantee

of the effective enforcement of data subjects' rights. When using AI systems, technical, organisational and legal measures must be put in place to ensure the effective implementation of the fundamental principles of data protection."