AI & Data Protection Series

AI and GDPR purpose limitation requirements

June 2025

Dr. Tina Gausling, Digital Business
Matteo Susta, Digital Business

Our speakers

Dr. Tina Gausling
Counsel
Germany

Matteo Susta
Supervising Associate
Italy

AI and GDPR purpose limitation requirements

Today's topics

- Foundations of Purpose Limitation
- AI vs. GDPR Purpose Limitation
- EDPB Guidance on AI and Purpose Limitation
- National Perspectives: Supervisory Authorities & Case Law
- Balancing AI and Purpose Limitation
- Key Takeaways & Recommendations

Purpose limitation principle

Art. 5 GDPR

- Personal data shall be collected for specified, explicit and legitimate purposes
- Not further processed in a manner that is incompatible with those purposes.

Purpose limitation principle

Critical role of this principle in the GDPR

Substantive compatibility assessment

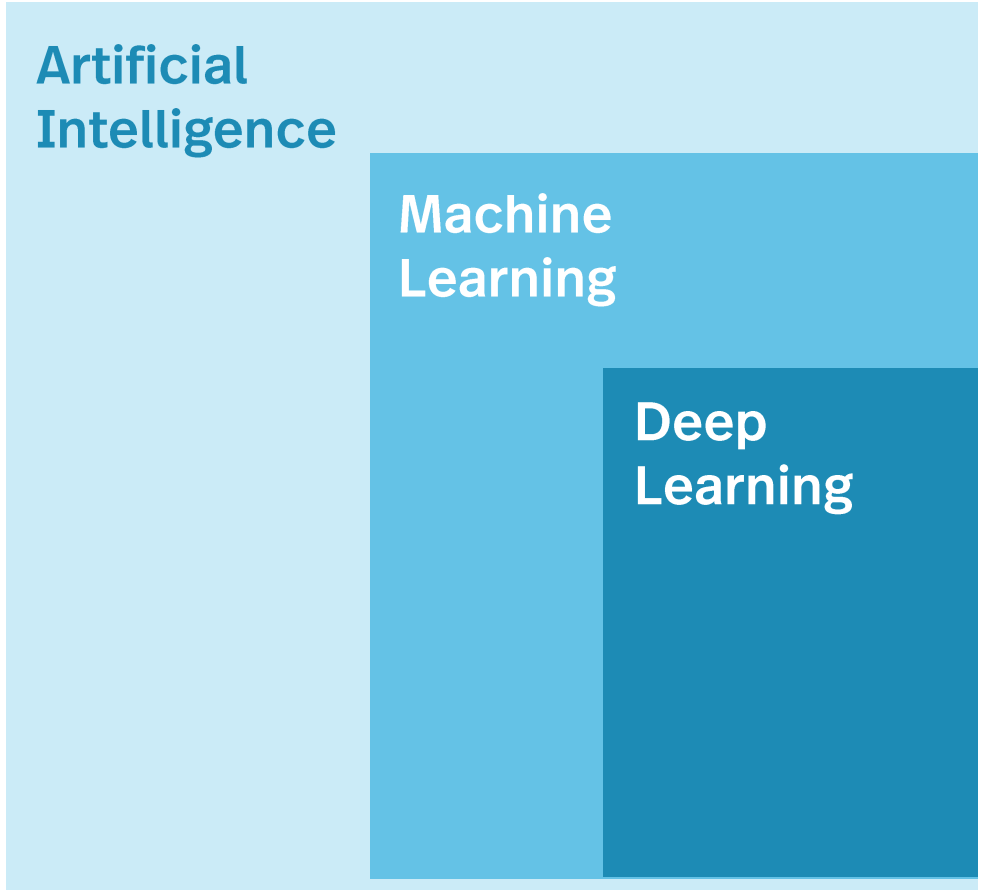
Art. 6 GDPR

Where the processing for a purpose other than that for which the personal data have been collected is not based on the data subject's consent or on a Union or Member State law, in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected, the controller must **take into account**, *inter alia*:

- any link between the purposes for which the personal data have been collected and the purposes of the intended further processing;
- the context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller;
- the nature of the personal data, in particular whether special categories or data related to criminal convictions and offences are processed;
- the possible consequences of the intended further processing for data subjects;
- the existence of appropriate safeguards, which may include encryption or pseudonymisation.

AI vs. GDPR Purpose Limitation

Technical Background | Machine Learning and Deep Learning



AI includes many different approaches

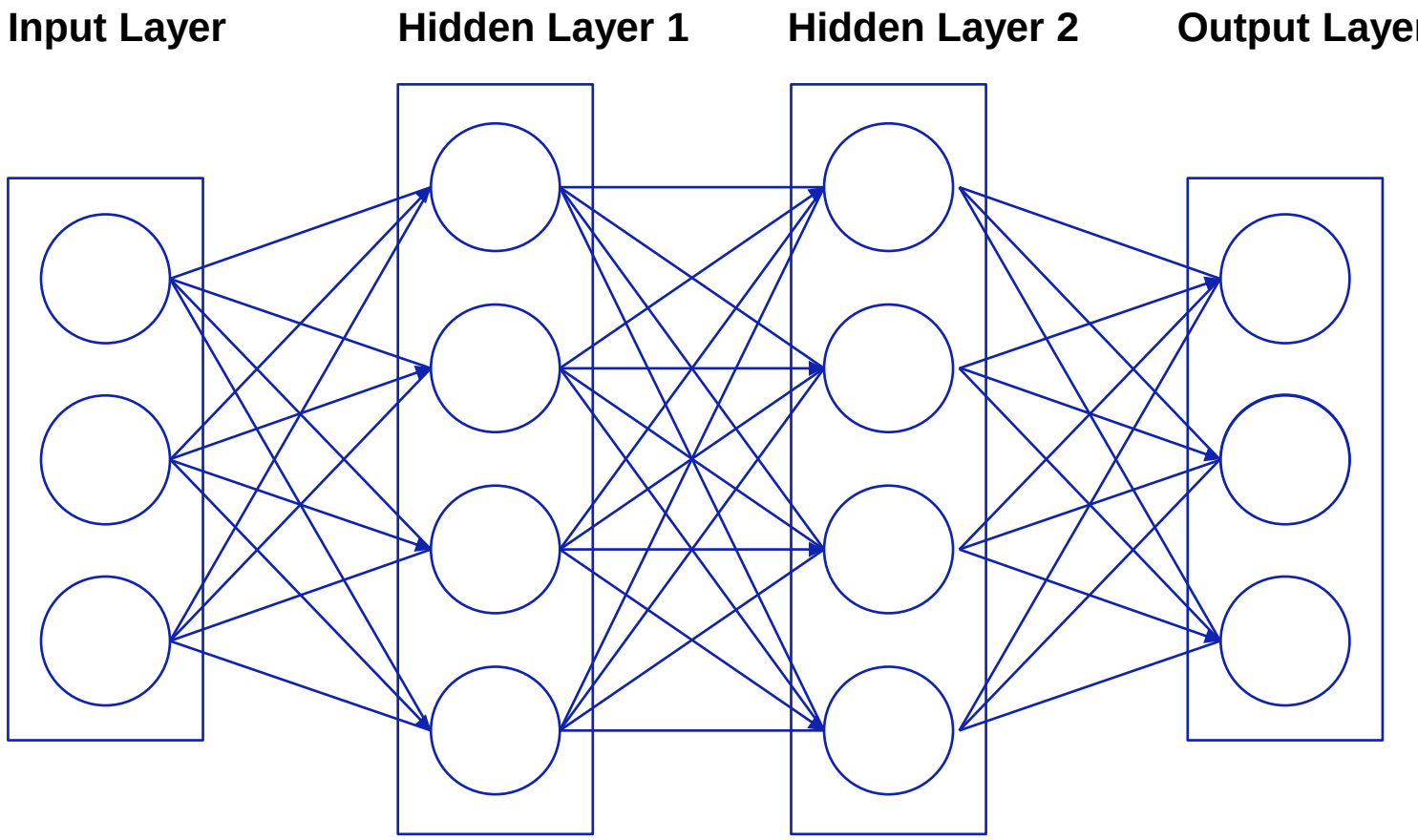
Machine Learning (ML) is one of the most important methods – mostly known under the term "neural networks"

Deep Learning as a special ML technique in which AI learns directly & through itself from examples

- basically the same as ML, but more layers (e.g. 150 instead of 2-3)
- significantly higher demand on data (classified data) and computing power
- most important application techniques today e.g. in the areas of autonomous driving or voice recognition

AI vs. GDPR Purpose Limitation

Blackbox of AI



AI vs. GDPR Purpose Limitation

Legal Challenges and Risks

Black box: Full transparency of algorithms technically not possible

- no traceability of processing and decision-making in detail
- Black Box tinkering: Input of various cases and comparison of output

Challenge for purpose limitation (Art. 5(1)(b) GDPR)

- If data is used for AI model training, it may later be repurposed for uses beyond the original intent
- Difficult to ensure that new uses remain within the scope of the original purpose or a compatible one
- Risk of “function creep” when AI learns patterns and applies them beyond initially communicated objectives

AI vs. GDPR Purpose Limitation

Legal Challenges and Risks

Appropriate legal basis for the development and use of AI?

- Consent, Art. 6 (1) (a) GDPR (e.g., for the analysis of each player's performance)?
  - Granular, informed consent is often not feasible in complex AI systems
  - Can personal data subject to self-learning algorithms continue to be used after consent is withdrawn?
- Legitimate Interest, Art. 6 (1) (f) GDPR (e.g., in order to prevent fraud/match fixing)?
  - Processing of personal data "strictly necessary" for the purposes of preventing fraud?
  - High threshold: interests of controller vs. rights and freedoms of data subjects
  - Purpose Limitation Concern
    - Using historical data for unforeseen detection patterns may go beyond what data subjects expected
    - Data reuse must be demonstrably linked to the original or a compatible purpose

EDPB Guidance on AI and Purpose Limitation

Guidance

European Data Protection Board (EDPB)

- EU body responsible for the consistent application of the General Data Protection Regulation (GDPR) across the EU/EEA
- Composed of representatives from all EU national data protection authorities (DPAs) and the European Data Protection Supervisor (EDPS)
- EDPB issues guidelines, opinions, binding decisions as well as consistency mechanisms to ensure unified interpretation of GDPR
- Guidance on AI and Purpose Limitation:
  - ChatGPT Taskforce Report (5/2024)
  - Opinion 28/2024 related to the processing of personal data in the context of AI models (12/24)
  - LLM Risks & Mitigation Strategies (4/2025)

EDPB Guidance on AI and Purpose Limitation

ChatGPT Taskforce Report (May 2024)

Aspect	Taskforce
Objective	Coordinated effort by EU supervisory authorities (via the EDPB) examining compliance of ChatGPT/OpenAI with GDPR – focusing on data collection, training, transparency, fairness, accuracy, and data-subject rights
Legal Basis for Training Data Collection	• OpenAI claimed legitimate interest (Art. 6(1)(f) GDPR) for scraping and using publicly available data. • Taskforce emphasized strict necessity, balancing test, and aligning with data subject expectation.
Transparency & Training Disclosure	• Users must be "clearly and demonstrably informed" that their prompts/input may be used for training. • Enables reasonable expectation, key to purpose limitation under Art. 5(1)(b) GDPR.
Safeguards to Reinforce Purpose Boundaries	Taskforce suggests • Filtering out sensitive categories at collection stage • Deletion or anonymization of personal data prior to training These measures help keep training within compatible, intended uses.

EDPB Guidance on AI and Purpose Limitation

Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models

Adopted on 17 December 2024

EDPB Guidance on AI and Purpose Limitation

EDPB Opinion 28/2024 – Purpose Limitation & AI (Dec 2024)

Aspect	EDPB Guidance	Opportunities for Companies
Legal Basis (Art. 6(1)(f) GDPR)	Legitimate interest may serve as legal basis for AI training if 3-step test is met: (1) legitimate interest exists, (2) processing is necessary, (3) interests of data subject do not override.	Companies may rely on legitimate interest if they can carefully justify necessity, balance interests, and implement safeguards.
Balancing Test Criteria	Factors include: Public availability of data, source and relationship with data subject, further use of model, transparency, reasonable expectations (e.g., self-disclosed data), data subjects' awareness of processing	Publicly available or self-disclosed data may strengthen company position if expectations are managed transparently.
Mitigation Measures	Additional safeguards can favour companies: Early and clear information, opt-out options, proactive transparency measures (e.g. public information campaigns)	Proactive communication and opt-outs increase compliance and support balancing test in favor of companies.
Purpose Limitation (Art. 5(1)(b) GDPR)	Purpose must be specific, explicit, and legitimate. Purpose must be clearly identified at each stage of AI development. Deployment context informs purpose.	Early documentation of AI purpose, use case, and deployment context helps demonstrate compliance.
Data Minimisation (Art. 5(1)(c) GDPR)	Use of anonymized/synthetic data is preferred where feasible.	Anonymization reduces GDPR exposure.
Compatibility Assessment (Art. 6(4) GDPR)	Opinion does not fully address whether compatibility assessment allows re-use of data collected for different purposes.	Uncertainty remains; companies should conduct careful case-by-case compatibility assessments.

Privacy Risks & Mitigations – Large Language Models (LLMs)

- Project launched by the EDPB in the context of the Support Pool of Experts programme
- Focus on LLM privacy risks including the “unlawful repurpose” (data used for a different purpose as infringing Art.5(1)(b) GDPR.
- This underscores that repurposing data beyond its original collection purpose (e.g. re-using conversation logs to refine the model) violates GDPR unless clearly consented or covered by original purposes.
- Ensure compliance with Article 5(1)(c) GDPR by
    - clearly limiting personal data processing to what is necessary for specific, well-defined purposes; and
    - avoiding overly broad purposes like "developing and improving an AI system." Instead, specify the type of AI system (e.g., large language model, generative AI for images) and its technically feasible functionalities and capabilities.

National Perspectives | Italy

Garante enforcement

Two of the most significant measures:

1. **Clearview**: database of over 10 billion facial images collected from public web sources through web scraping, such as news sites, social media, and online videos. These images were used to offer an advanced search service that, thanks to AI, allowed the creation of profiles based on biometric data, potentially enriched with other related information, such as the title and geolocation of the photo, and the web page of publication.

2. **OpenAI** : further processing of data uploaded by users of ChatGPT service for the purpose of training the system deemed incompatible with the provision of the service – need to a specific legal basis and transparency on the training activities on tha basis of PI

National Perspectives | Germany

German DSK Guidance

- German Data Protection Conference (DSK)
  - Body of Germany's independent data protection supervisory authorities (federal & state)
- AI & Data Protection Guidance (May 2024)
  - AI processing requires clearly defined purposes upfront (Art. 5(1)(b) GDPR).
  - AI training with personal data qualifies as separate processing – requires specific legal basis.
  - DSK approach: strict and purpose-driven, limited room for interpretation.
  - Companies face narrow flexibility; extensive legal assessment required for each processing step.

National Perspectives | Germany

OLG Cologne – Decision dated 23 May 2025

Facts of the case

- Meta (Facebook & Instagram) announced it would use publicly available user data (from adult users) to train AI models, providing an opt-out mechanism
- Verbraucherzentrale NRW (consumer association) filed for an injunction, arguing the use violates the GDPR, particularly purpose limitation and consent requirements.

Key legal Question

- AI Training: Can user-generated content be processed based on legitimate interest (Art. 6(1)(f) GDPR) without opt-in?

Outcome

- OLG Cologne rejected the injunction in summary proceedings
- AI training may qualify as a “further compatible use” and is permissible under Art. 6(1)(f) GDPR, provided safeguards are in place

National Perspectives | Germany

OLG Cologne – Decision dated 23 May 2025

Aspect	OLG Cologne	Criticism
Original Purpose	User communication and self-expression on social media	AI training for model development = different function and purpose
Purpose change allowed?	Not expressly mentioned in press release	No – AI training goes beyond original expectations
Legal Basis	Legitimate interest under Art. 6(1)(f) GDPR; opt-out and limited to public adult data	GDPR requires opt-in for incompatible secondary use
Impact on users	Sufficient transparency and ability to object (opt-out)	Users cannot foresee how their data will be reused for model training; no meaningful control or awareness

Balancing AI and GDPR Purpose Limitation

Revisiting Purpose Limitation

- EDPB leaves limited room for AI training based on legitimate interest, subject to strict purpose definition, balancing test, strong safeguards and transparency.
- Use of anonymized/synthetic data where possible
  - e.g. by means of Generative Adversarial Networks
  - Deep Learning method: two competing neural networks working against each other; trained with identical data
  - Enables artificial generation of data → GDPR not applicable, no consent necessary
- Initiatives to broaden purpose limitation requirement
  - EHDS: introduces legally defined secondary use purposes (incl. AI development) for health data beyond original collection purpose
  - Political initiatives in Germany to relax purpose limitation (e.g. Bitkom: German digital industry association representing tech companies; calls for more flexible purpose limitation for AI training and research, in particular for health data)

Key takeways

1. Define all intended purposes for data use upfront, including training – privacy by design – and verify the lawfulness on the processing on dataset purchased from other controllers.
2. Take purpose limitation requirement into account in daily data processing activities and document any compatibility assessments.
3. Anticipate misuse risks and build safeguards (like encryption and pseudonymization) from the start.
4. Provide sufficient granularity and detail within the privacy notices regarding the processing activities intended to be carried out.
5. Keep record of data processing up to date as use cases evolve.

Useful resources

CtrlTransfer

Digital Regulation Tracker

AI Literacy Programme

Next session > 24 June: AI-related Data Protection Impact Assessments (DPIAs)