thws Technical University of Applied Sciences Würzburg-Schweinfurt

FAKULTÄT INFORMATIK UND WIRTSCHAFTSINFORMATIK

LECTURE NOTES

Data Protection I: Foundations & the GDPR

Informational self-determination, scope, principles and lawful bases

MODUL
Introduction to the Ethics and Regulation of AI

SEMESTER
WS 26/27

VERSION
1.0

Prof. Dr. Markus Oermann
Instructor
markus.oermann@thws.de

Technische Hochschule Würzburg-Schweinfurt
Stand: July 23, 2026

Contents

In the previous unit we saw how bias and discrimination can creep into AI systems through the very data on which they are trained. That observation leads naturally to a harder question: on what legal terms may those data be collected and used in the first place? Data protection law is the first field of law in this course that translates an ethical value, informational self-determination, into duties that can be enforced with the power of the state. This unit lays the foundation. It clarifies *why* data protection exists, *what* it protects, *where* it applies and by which *principles* every act of processing must be judged, so that you can recognise when the General Data Protection Regulation (GDPR) governs an AI project and what standards it imposes.

01
KAPITEL

Why data protection, and what it protects

1.1 Data protection is not data security, and not data ownership

In everyday speech, data protection and data security are used almost interchangeably. In law they are two distinct fields with different aims. **Data security**, or more precisely information security, protects the hardware, software and the data themselves against loss, destruction or misuse; its classic goals are confidentiality, integrity and availability. **Data protection**, by contrast, protects the natural person behind the data against violations of their personality rights. The two fields overlap, since data protection cannot be implemented without technical security, but they pursue different purposes. For an AI system, this means that encrypting a training set well is a security measure, whereas asking whether the people in that set may be processed at all is a data-protection question.

A second clarification matters just as much. European data protection is not about *who owns the data*. A legal concept of "data ownership" simply does not exist in the EU. Instead, data protection seeks to shield individuals in situations of structural power asymmetry, whether against the state or against powerful firms. The party that processes our data usually knows far better than we do what is stored about us and how it is used. Data protection tries to soften that imbalance of knowledge and power, a concern that grows sharper, not weaker, as AI systems infer ever more from ever less.

DEFINITION

Data security vs. data protection

Data security protects the *data* against loss and misuse; data protection protects the *person* behind the data against violations of their personality rights.

1.2 The Census judgment and informational self-determination

The intellectual core of European data protection can be traced to a landmark decision of the German Federal Constitutional Court in 1983, the so-called Census judgment. A planned national census had provoked a broad protest movement: thousands feared comprehensive statistical registration and saw in it the beginnings of a surveillance state. From human dignity in conjunction with the general right of personality, the Court derived a new fundamental right to *informational self-determination*.

The core idea is that each person may in principle decide themselves about the disclosure and use of their personal data, and thus who learns which facts about their life. The Court also held that under the conditions of automated processing there is no longer any "trivial datum": even an innocuous piece of information can, when linked with others, acquire real significance for the picture drawn of a person. This insight anticipates the logic of modern AI almost word for word, since machine-learning systems specialise precisely in recombining scattered fragments into revealing profiles.

DEFINITION

Informational self-determination

The right of the individual to decide, in principle, about the disclosure and use of their own personal data, derived by the German Federal Constitutional Court in the 1983 Census judgment from human dignity and the general right of personality.

1.3 The panopticon and the chilling effect

Why is the mere collection of data a problem for someone who has "nothing to hide"? An image from political philosophy answers the question: the **panopticon**, a prison designed by Jeremy Bentham in which a single guard at the centre could observe every inmate at any time, without the inmates ever knowing whether they are being watched. The mere *possibility* of observation disciplines behaviour. Michel Foucault later turned this model into a metaphor for modern surveillance societies, in which total visibility slides into total control.

This is exactly where the Census judgment intervenes. It guards against the so-called **chilling effect**: those who must fear that deviant behaviour will be recorded and processed at any moment will inevitably adapt their conduct to the supposed mainstream. A social order in which citizens can no longer know who knows what about them, when and on what occasion, the Court held, is incompatible with the right to informational self-determination. The argument runs on two levels at once, an *individual* one (protecting the free development of personality) and a *collective* one (self-determination as a precondition for a free, democratic community). Data protection is therefore not merely private secrecy; it protects foundations of democracy. For AI, the point is direct: a population that senses it is permanently profiled behaves differently, and that behavioural shift is itself a democratic cost.

EXERCISE · TERMS

The imagined prison in which the mere **possibility** of observation disciplines behaviour is the **panopticon**. The resulting deterrence from exercising one's rights is called the **chilling effect**.

1.4 From a national right to the GDPR

Data protection also has a long statutory history, running from the world's first data-protection act in the German Federal State of Hessen in 1970, through the European Data Protection Directive (1995) to the **General Data Protection Regulation (GDPR)**, adopted in 2016 and directly applicable in all EU Member States since 25 May 2018. As a regulation, the GDPR needs no national transposition; it replaced a patchwork of national laws with a largely uniform, directly applicable regime backed by severe sanctions, a model that has since radiated worldwide. In Germany it is supplemented by the Federal Data Protection Act (BDSG) through so-called opening clauses, but the substance that concerns an AI developer sits in the Regulation itself.

QUICK CHECK

Which fundamental right did the German Federal Constitutional Court develop in the Census judgment (1983)?

The right to data ownership
The right to informational self-determination
The right to be forgotten
The secrecy of telecommunications

02
KAPITEL
When does the GDPR apply?

2.1 The material and territorial scope

Before the sharp sword of the GDPR can be drawn at all, two entry conditions must be met. The **material scope** (Art. 2 GDPR) requires the wholly or partly automated processing of personal data; alternatively, it suffices that data are stored, or meant to be stored, non-automatically in a structured filing system. Excluded in particular is processing by a natural person in the course of a purely personal or household activity, the so-called household exemption (Art. 2(2)(c) GDPR). For AI this threshold is almost always crossed, since training, inference and analytics are automated by definition.

The **territorial scope** follows the establishment and marketplace principles (Art. 3 GDPR). The Regulation applies whenever a controller or processor has an establishment in the EU and the

processing occurs in the context of its activities (the establishment principle). It also applies, however, as soon as goods or services are offered to a person located in the EU, or that person's behaviour is monitored (the **marketplace principle**). It is therefore legally irrelevant on which server, in Silicon Valley or elsewhere, the data physically reside. An AI provider based abroad that offers a service to European users, or profiles their behaviour, is squarely within the Regulation's reach. Let us recapitulate the applicability test as a sequence of questions.

CASE STUDY

Case: glancing at a classmate's laptop

During a lecture, Anna glances at her neighbour's unlocked laptop and memorises his Instagram handle and phone number, meaning to invite him for a coffee later. Does this everyday act fall within the scope of the GDPR?
Solution. Anna undoubtedly handles personal data (name, number), but she does so neither by automated means nor by storing them in a structured filing system (Art. 2(1) GDPR). Even if one were, with much legal imagination, to assume a relevant act of processing, the household exemption of Art. 2(2)(c) GDPR would apply: Anna processes the data purely for a personal activity. The GDPR does not apply. The contrast with AI is instructive, since a system that automatically scrapes and stores the same handles at scale would cross both thresholds at once.

2.2 Personal data and the concept of processing

The pivotal criterion of every data-protection case is the presence of personal data. Under Art. 4(1) GDPR these are all information relating to an identified or, far more often, identifiable living natural person. This covers classic identifiers such as names, addresses and account numbers, but it also extends to photos of recognisable people, location data and even dynamic IP addresses, provided the user can be identified by means reasonably likely to be used (CJEU, *Breyer*, C-582/14). As the Census judgment already stressed, European law knows no "trivial datum": seemingly insignificant fragments can, through intelligent linkage, gain immense relevance for a personality profile. Purely machine-generated values with no reference to a person (say, the power consumption of a milling machine), truly anonymous data and data of deceased persons fall outside the protected scope, since no living natural person stands behind them.

The **concept of processing** in Art. 4(2) GDPR is drawn as widely as possible. It captures virtually any operation performed on data, from collection and storage through alteration and transmission to erasure. For an AI pipeline this means that every stage counts as processing: assembling a training corpus, fine-tuning a model, running inference on a prompt and logging the output. Each stage needs its own justification, a point that becomes central in the next unit. Test in the following scenario checker which kinds of data qualify as personal.

2.3 Controller and processor

Who must comply with the GDPR is no trivial preliminary question; it decides at whom duties and fines are directed. The Regulation distinguishes two central roles. The **controller** (Art. 4(7) GDPR)

is the natural or legal person who, alone or jointly with others, determines the *purposes and means* of the processing. Put simply, whoever sets the "why" and the "how" of the processing is the controller and bears the main burden of data-protection duties. The **processor** (Art. 4(8) GDPR), by contrast, processes personal data *on behalf of* the controller without deciding on purposes and means itself; the classic example is a cloud-hosting provider or, in an AI context, an external firm that runs a training job strictly to the client's instructions.

This division is secured contractually. Where several parties *jointly* decide on purposes and means, joint controllership arises (Art. 26 GDPR), and they must transparently agree who fulfils which duty, in particular towards the data subjects.

Getting the role right is highly consequential in practice, because it determines who is liable, who must satisfy information duties and against whom data-subject rights are directed. In AI value chains this classification is rarely obvious: a company that fine-tunes a foundation model on its own customer data, choosing purposes and configuration, will typically act as controller even while relying on a third party's infrastructure.

QUICK CHECK

An AI provider hosts a client's customer database on its cloud and processes the records only on the client's instructions. What role does the provider have?

Controller
Processor
Joint controller
Data subject

03

KAPITEL

The principles of processing (Art. 5 GDPR)

Once the GDPR applies, every single processing activity must satisfy the demanding requirements of Art. 5 GDPR. These principles form the normative foundation of data protection; they are the yardstick against which any processing must be measured, and later units spell them out into concrete duties and rights.

The most important principle is **lawfulness**. It holds that processing personal data is, in principle, prohibited unless consent or another permissive rule applies. This inverts the usual default of a liberal legal order, in which everything not otherwise regulated is allowed. The principles of **fairness** and **transparency** require that data subjects can understand what happens to their data and are not tricked by misleading practices or dark patterns. The strict principle of **purpose limitation** demands a specified, explicit and legitimate purpose for every processing and forbids

repurposing data for incompatible new ends, a rule that sits uneasily with the appetite of AI development for reusing whatever data lie to hand.

Two further principles curb the sheer volume of data. **Data minimisation** requires forgoing superfluous collection: only as much data may be gathered as is necessary to achieve the purpose. It is flanked by **storage limitation**, which orders erasure once the purpose has lapsed. Finally, data must be factually **accurate** and protected by appropriate technical and organisational measures against unauthorised access (**integrity and confidentiality**).

Spanning all of this is **accountability**: the burden of proving compliance always rests with the controller, who must not only achieve conformity but document and demonstrate it. Explore the principles interactively below.

QUICK CHECK

Which principle requires that data be erased once they are no longer needed for the original purpose of collection?

- Data minimisation
- **Storage limitation**
- Purpose limitation
- Integrity and confidentiality

DEEP DIVE **Deep dive: accountability as a shift of the burden of proof**

The accountability principle anchored in Art. 5(2) GDPR marks a reversal of the burden of proof. Where formerly the supervisory authority had to prove an infringement, today the controller must actively demonstrate compliance, through records of processing activities, data-protection impact assessments and documented technical and organisational measures. This duty is the hinge between the abstract principles of this unit and the concrete compliance obligations treated later. For AI systems it is unusually demanding, because opaque models make it hard to show, after the fact, exactly which data flowed where and why.

04
KAPITEL

Lawful bases: prohibition subject to permission (Art. 6 GDPR)

If lawfulness is the cardinal principle, then Art. 6 GDPR is its operative core. European data protection rests on a rule as simple as it is far-reaching: every processing of personal data is first of all prohibited. Only once the controller can point to an explicit statutory permission, a lawful basis, does the processing become admissible. Lawyers call this a “prohibition subject to permission”.

DEFINITION

Prohibition subject to permission

Every processing is in principle inadmissible unless a legal basis (under Art. 6 GDPR) expressly permits it.

The relevant permissions sit in Art. 6(1) GDPR. If just one of them applies, the data-protection traffic light turns green; the six grounds of points (a) to (f) together form the switchboard of data-protection law, and every processing must rest on at least one of them.

DEEP DIVE **The six lawful bases of Art. 6(1) GDPR at a glance**

- **Consent (a):** the data subject has agreed to the processing in the manner the law requires.
- **Performance of a contract (b):** the processing is objectively necessary to perform a contract, for instance processing an address to deliver a parcel.
- **Legal obligation (c):** the controller must comply with a law and process data to do so, such as tax retention duties.
- **Vital interests (d):** where it is literally a matter of life and death, for instance transmitting a blood type to emergency services after an accident.
- **Public task (e):** where processing is necessary to perform a statutory public task, important for authorities and public universities.
- **Legitimate interests (f):** the processing is necessary for the legitimate interests of a body and the interests of the data subject do not override them. Authorities may not rely on this ground for their public tasks.

Two grounds deserve special attention in an AI context. **Performance of a contract (b)** carries only so far as the processing is *objectively necessary* for the contract; mere convenience or an added advertising purpose is not enough. **Legitimate interests (f)** is the most flexible but also

the most demanding basis. It requires a three-step test: a legitimate interest of the controller, the necessity of the processing to pursue that interest, and a balancing exercise in which the interests, fundamental rights and freedoms of the data subject must not prevail. Precisely because point (f) offers no fixed catalogue, clean documentation of the balancing is decisive in practice, and it is the ground on which providers increasingly try to justify training AI models on publicly available data. The Court of Justice has begun to police its outer limits: in *Schrems v. Meta* (C-446/21) it held that data a person has manifestly made public do not become a blank cheque for unlimited advertising processing.

EXERCISE · TERMS

Data protection follows the rule of a prohibition subject to **permission**: processing is lawful only if at least **one** condition of Art. 6 GDPR is met. A company may rely on its **legitimate interest**, provided the rights of the data subjects do not **prevail**.

4.1 Consent and its strict conditions

Practice tends to misread consent (Art. 6(1)(a) GDPR) as a data-protection blank cheque. In truth, consent, defined in Art. 4(11) and shaped by Art. 7 GDPR, is tied to strict conditions of validity. It must be **freely given**, which means a genuine, unpressured choice; where a power imbalance constrains that choice, as between employer and employee, very strict standards apply. Consent must also be an **active act** (opt-in): silence, inactivity or pre-ticked boxes do not meet the requirements, which is why tracking analytics may only start after an affirmative click. It must be **informed and specific**, given after clear, plain-language explanation of who processes what and above all why. And it must be **revocable** at any time, as easily as it was given, though the withdrawal operates only for the future: processing carried out before withdrawal remains lawful. For AI, this makes consent a fragile basis for large training sets, since informed, specific and revocable agreement is hard to obtain at scale.

DEFINITION

Consent

Any freely given, specific, informed and unambiguous indication of the data subject's wishes, by a statement or a clear affirmative action, that they agree to the processing of their personal data.

05
KAPITEL

The highest alert level: special categories of data (Art. 9)

Some information is so sensitive that the ordinary bases of Art. 6 and the general duties of the GDPR do not suffice to protect it. Art. 9 GDPR therefore raises a distinctly stricter shield. It distinguishes ordinary personal data from special categories and, in its first paragraph, sets out a rigorous, principled prohibition on processing the latter.

DEEP DIVE Special categories of personal data (Art. 9(1) GDPR)

These specially protected data reveal:

- racial or ethnic origin;
- political opinions, religious or philosophical beliefs;
- trade-union membership;
- genetic data and biometric data used for unique identification;
- data concerning health, sex life or sexual orientation.

Exceptions to the prohibition exist, but they are drawn far more narrowly than under Art. 6. Processing is permitted only where a specific, still stricter ground of Art. 9(2) GDPR applies, such as **explicit**, separately given consent for exactly these sensitive data, or a legal duty in employment and social-security law. The heightened standard, “explicit” rather than merely “unambiguous”, makes the point clear: whoever wants to base sensitive-data processing on consent must obtain it separately, clearly delimited and targeted at precisely that category. This matters acutely for AI, because models routinely infer sensitive attributes (health, ethnicity, sexual orientation) from innocuous inputs, and such inference can itself amount to processing special-category data. Note also the interface with the EU AI Act, whose Art. 10(5) carves out a narrow permission to process special categories for the purpose of detecting and correcting bias in high-risk systems, a bridge back to the previous unit.

CASE STUDY

Case: the tattoo studio’s consent form

A tattoo studio routinely has clients sign a form before the session. Buried in a long block of text on page two is the clause: “By this declaration we collect health data from you (e.g. on infections) in order to tattoo you safely. We also take before-and-after photos of the work for our Instagram channel.” Is this valid?

Solution. The explicit inquiry into health conditions collects health data and thus falls within Art. 9 GDPR. The blanket consent, hidden in continuous text without a clear visual and substantive separation of two entirely different purposes (medical preparation versus social-media advertising), fails the basic requirements of informed and specific consent. An explicit, separately given agreement targeted at exactly these sensitive health data is missing. In this undifferentiated form the consent is almost certainly wholly invalid, and the two purposes would each need their own, clearly delimited legal footing.

EXERCISE · TERMS

The processing of sensitive data, such as **health** data, is subject to a principled prohibition under **Art. 9** GDPR. Exceptions are possible only under strict conditions, for example an **explicit** consent that is clearly set apart from the surrounding text.

With this, the foundation is laid: we now know *why* data protection exists, *what* it safeguards as the value of informational self-determination, *when* the GDPR applies, on which *principles* and *lawful bases* every processing rests, and how special categories are shielded. The next unit builds directly on this footing to confront the questions that arise when AI moves from ordinary data handling to profiling and automated decision-making: the special regime for automated decisions and profiling (Art. 22), Privacy by Design, the data protection of AI models and training data, and the international transfer problem crystallised in *Schrems II*.

06
KAPITEL
References

6.1 Literature

- Foucault, M. (1977): *Discipline and Punish: The Birth of the Prison* (translated by Alan Sheridan). Pantheon Books, New York.

6.2 Norms & Standards

- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data (General Data Protection Regulation, GDPR), in particular Art. 2, 3, 4(1), 4(2), 4(7), 4(8), 4(11), 5, 6, 7, 9, 26. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679
- Regulation (EU) 2024/1689 (Artificial Intelligence Act), Art. 10(5). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689

6.3 Case law

- German Federal Constitutional Court, judgment of 15 December 1983, 1 BvR 209/83 and others (Census / *Volkszählung*), BVerfGE 65, 1.
- CJEU, judgment of 19 October 2016, C-582/14 (*Breyer*, dynamic IP address as personal data); continued by German Federal Court of Justice, judgment of 16 May 2017, VI ZR 135/13.
- CJEU, judgment of 4 July 2023, C-446/21 (*Schrems v. Meta Platforms Ireland*, limits of legitimate interests and of the exception for manifestly public sensitive data).