AI Governance — take-away sheet

AI Governance for Research

Which system · under which conditions · with which safeguards · for which data

DSI Excellence Program BootCamp · 3 September 2026 · Giovanni Spitale, IBME / DSI

The decision, in three lines

The data has a **level**.

The tool has a **ceiling**.

The decision is a **match**.

The four levels

		What it covers
öffentlich	*public*	published, or explicitly cleared for publication
intern	*internal*	**the default for anything unmarked** — unpublished work, internal documents, drafts
vertraulich	*confidential*	personnel files, **student data** (exams, theses, coursework), personal research data, logs
geheim	*secret*	professional secrecy under **art. 321 / 321bis** of the Swiss Criminal Code — patient data, research involving human beings under the HRA; contractual NDAs; judicial files

Source: *Weisung zur Klassifizierung von Informationen*, UZH, in force since 01.09.2023, based on cantonal law (IDG, IVSV).

On *geheim* specifically. Art. 321 CP lists the professions bound by professional secrecy — doctors, psychologists, nurses, lawyers, notaries and others, plus their auxiliaries. **Art. 321bis is**

the one that brings researchers in: disclosing a secret learned through research involving human beings is punished the same way. Its second paragraph is a door rather than a wall, though a narrow one: such secrets **may** be disclosed for research into human disease and into the structure and function of the human body, where art. 34 of the Human Research Act is satisfied **and the competent ethics committee authorises it**. Art. 34 is the *exception* route — it covers further use where informed consent could not be obtained, on cumulative conditions. Note what that authorisation does and does not do: it lets you *use* the data for your research. It does not reclassify it, and it does not let it go to a commercial provider.

Swiss Criminal Code, SR 311.0 · Federal Act on Research involving Human Beings (HRA), SR 810.30 — both on fedlex.admin.ch, authoritative in German, French and Italian; the English versions state that they have no legal force. ***This is orientation, not legal advice****: for a real case, ask your ethics committee.*

Three rules that settle most cases

1. **An unmarked document is at least *intern*.** There is no "unclassified" state. Absence of a label is itself a label, and it is not public.
2. **Until publication, everything is at least *intern*.** Including your own work. Especially your own work: the manuscript you feel is most *yours* is the one you have least right to send out. Ownership and the right to distribute are different things.
3. **Classification is inherited across whole information holdings.** One sensitive item pulls its container up with it. You do not get to treat the folder as internal because most of it is dull.

Which tool can hold what

Verified at source on 27 August 2026 — and it will change. Check before you rely on it.

Tool	Levels permitted	Hosting
Free AI tools — ChatGPT, Claude, DeepL, Copilot…	**öffentlich only**	worldwide cloud
M365 Copilot Chat (Basic)	öffentlich, intern	EU cloud
M365 Copilot (Premium)	+ vertraulich[1]	EU cloud
DeepL Pro Advanced	+ vertraulich	EU cloud
GitHub Copilot	öffentlich, intern	worldwide cloud
ScienceCloud (GPU research)	+ vertraulich[2]	on premises
ScienceCluster (GPU research)	öffentlich, intern	on premises
Supercomputer (GPU research)	öffentlich, intern[3]	CH cloud

Tool	Levels permitted	Hosting
LM Studio + Mistral	+ vertraulich	on device

[1] requires web search disabled · [2] depends on usage context · [3] non-personalised scientific data only

The two sentences worth memorising

> **Free AI tools are for *öffentlich* data only.**
>
> **For *geheim* data there is no AI tool at all.**

When there is no tool: four exits, each with a price

Exit	What it costs
Anonymise / de-identify	time — and often *precisely* the signal you needed
Switch tool (local, or institutional)	model quality, sometimes feasibility
Do it by hand	months. For a large corpus, it means not doing it
Ask for infrastructure that does not exist yet	does not solve your problem this year

Two things about anonymisation, before you assume it is the easy exit.

It is not free. Fifty interviews of fifty pages each is a real corpus, not a thought experiment. De-identifying it by hand is weeks of work by someone who understands the material — and you cannot hand it to the AI to do for you, because handing over the transcript *is already the transfer you were trying to avoid*.

Sometimes it is impossible. In rare disease, that patient may be the only case in the canton: strip the name, the age and the hospital, and the condition alone still identifies them. Same for a small language community, a single-site cohort, an unusual occupation. **When the clinical detail *is* the identifier, de-identification and usefulness are the same quantity** — you cannot spend one without spending the other.

The question is about the flow, not the tool

A single system can operate at two levels at two different points. Published literature going in is *öffentlich*; the log of which claims you tested, in which order, is your unpublished research programme, and that is *intern* at least.

> **So never ask "what level is this tool". Ask "what level is this flow, at this point".**
> Public inputs can produce internal outputs — and the moment they do, your obligations change.

Where to check, when today's answer stops being true

- **Tool matrix (ZI):** zi.uzh.ch → Staff → Software & eLearning → AI Tools and Services
- **The directive:** *Weisung zur Klassifizierung von Informationen* (rud.uzh.ch)
- **Recommendations on generative AI:** uzh.ch/en/explore/basics/ai
- **A demo to poke at:** *Which tool, which data?* — an unofficial GPT built for this session · [link] · a snapshot of the August matrix, and wrong in interesting places
- **The law behind *geheim*:** Swiss Criminal Code, art. 321 / 321bis — **SR 311.0**, fedlex.admin.ch/eli/cc/54/757_781_799 · Human Research Act — **SR 810.30**, fedlex.admin.ch/eli/cc/2013/617
- **What is coming:** ZI blog, category *KI* — contracted ChatGPT, an institutional LLM proxy, JupyterHub and more are all on the roadmap between now and 2027

The matrix above has a date on it for a reason. What is worth carrying away is not the table — it is how to read one: **find the ceiling, check the hosting, read the footnotes, check the date on the page you are reading.**

Next

(Leveraging) AI for Research Purposes — 19–21 October 2026, SOC-E-010, 1 ECTS, pass/fail. Research *on* AI vs *with* AI · tools by research task · context management and writing · a workflow of your own, designed and defended. **No preparatory work.**

Questions, or a case you are genuinely unsure about: **giovanni.spitale@ibme.uzh.ch**