Lecture XI - Project Work I

Programming with Python

Dr. Tobias Vlćek



EP 11: The Build

The shop has a new owner

MunchCorp runs the shop now. Tobi kept his job, the investor kept her spreadsheets, and her next due-diligence visit already has a date: **Session XIII**, when your projects run in this room.

Until then, the shop rule stands: pull before you start, push when you stop, and let the commit history do the talking.

The Mark

Could anyone tell?

- Your README carries an AI-disclosure section. Suppose you skipped it: could anyone **prove** that an AI wrote part of your work?
- Since August 2026, EU rules require providers to mark AI content machine-readably, and the major ones do
- The trick behind it is simple enough for three slides

How a watermark hides in plain text

- Before each word, the model splits its vocabulary into a **green** and a **red** half, reshuffled at every step
- It then softly prefers green words. The text still reads completely normally
- By chance, a human writes about half green words. A model writes far more

You already know the mechanism

- Session VI: the model picks a **likely** next token, never *the* next token
- A watermark is that choice, steered on purpose: green wins a little more often than chance
- Same next-token choice, one hidden thumb on the scale

Who can actually detect it

- Detection is **counting**: one paragraph of marked text can push chance below 1 in 10 trillion
- But you need the secret key that decided which half was green

- In practice, only the provider can run the test

What a mark proves (and does not)

- Proves: a model **processed** this text
- Does not prove: who did the work, or whether a human verified it
- Heavy editing can wash a mark out, and short snippets may not carry one, so an absence proves nothing

Presence without disclosure, though, looks exactly like what it is.

Disclosure is part of the work

- The rules converged everywhere: AI cannot be an author, and its use gets disclosed
- Your README section says which tool, for what, and what you verified yourself
- That is **standard practice**, not a confession

The Work Menu

Today

- **Pull before you start, push when you stop.** GitHub is your backup; commit as you go
- Progress today means one working step pushed to GitHub, with commits that tell the story
- I am in the room: use me for blockers, design questions, and “is this scope realistic?”
- Use the agent where it helps, and note down what it did for your disclosure section

When you’re stuck

> **Tip**
> A red traceback you have never seen? Read the last line first, exactly as in Part I. Then, and only then, the agent.

Wrap-up

Three things to remember

- AI text carries a **watermark** now: fluent, normal-looking, detectable by the provider
- Disclosure is part of the work: which tool, what for, what you verified
- Small commits, pushed often, are the cheapest insurance any of you can buy

Next week: your agent keeps making the same mistakes. Time to **teach it once** instead of correcting it every session.

Bibliography