

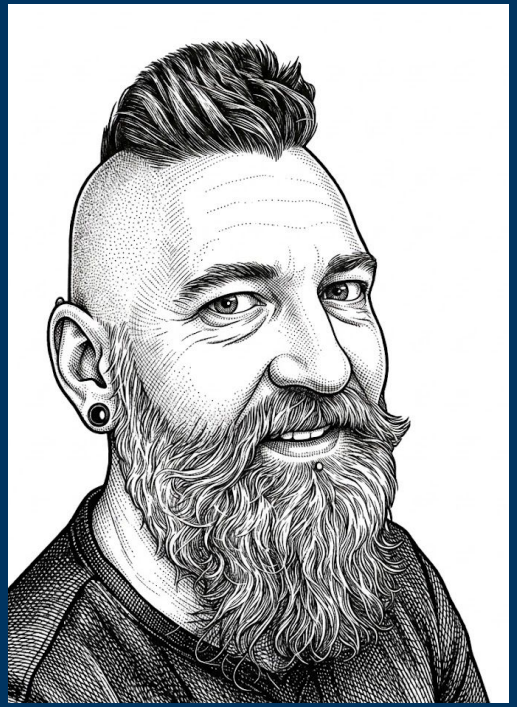
Kill your Tests

Wie wir den E2E-Sumpf mit KI-Agenten endlich trockenlegen

NÜRNBERGER VERSICHERUNG

© felix tensing

NÜRNBERGER
VERSICHERUNG



Felix Tensing

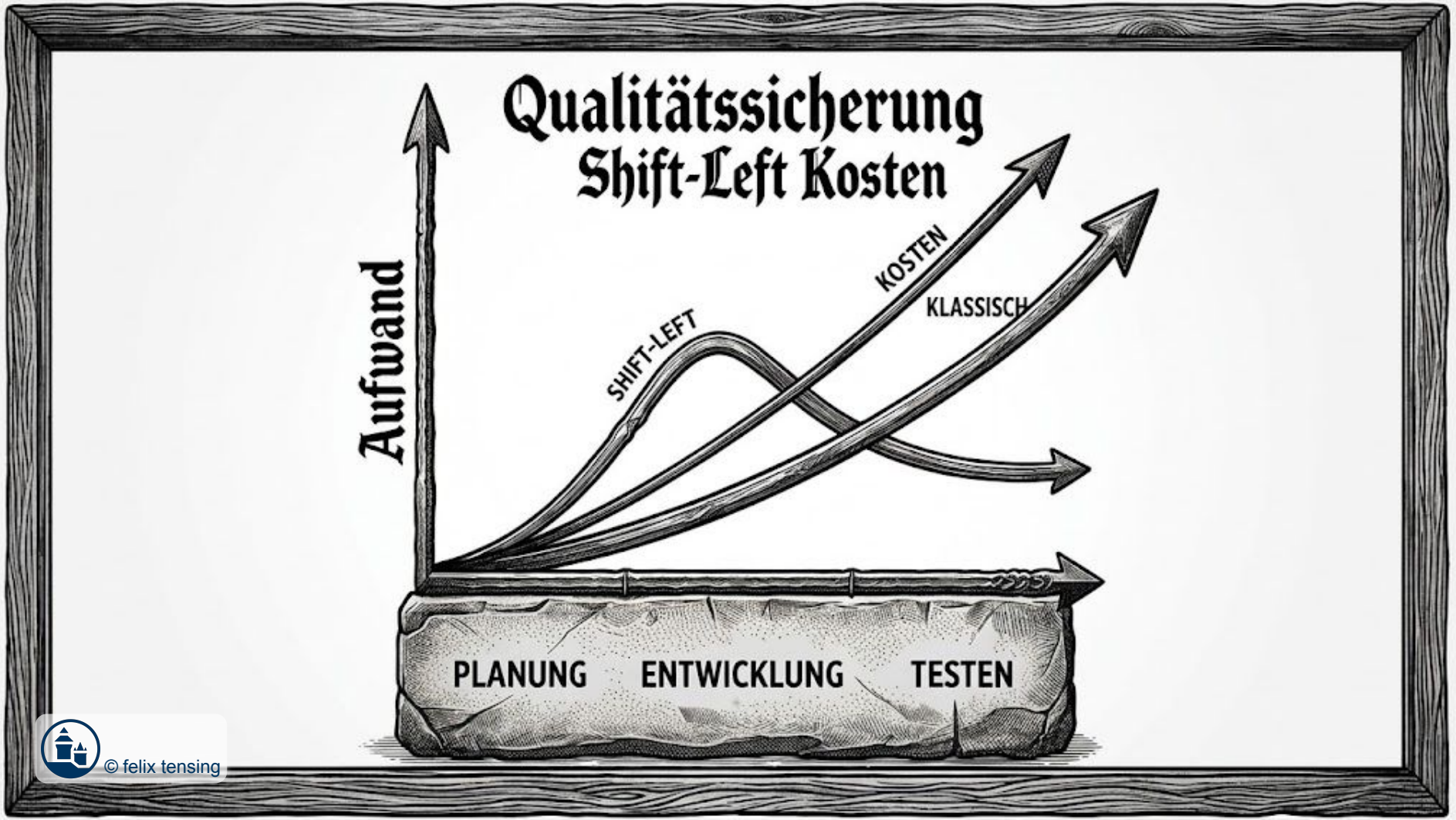
Softwareingenieur und Drachentöter im Zeichen der Burg

felix-spricht.de

Shift Left



© felix tensing



© felix tensing

Shift-Left - Mehr als Tools

Konzept	Fokus	Werkzeuge (u.a.)
Doing things right Stabilität / Effizienz	● Technische Korrektheit ● Effizienz ● Syntax ● Performance	● Unit Tests ● **Testautomatisierung** ● CI/CD ● SonarQube
Doing the right thing Nutzen / Relevanz	● Sinnhaftigkeit ● Usability ● Business Value ● Logikfehler	● Domain Driven Design ● Exploratives Testen ● Pair Programming ● Reviews



© felix tensing

Das BDD Drama



© felix tensing

Was ist BDD - Behaviour Driven Development

„E

```
@keyword("Click Login Button")
def click_login_button(self):
    login_btn = self.driver.find_element(By.ID, "login-btn")
    login_btn.click()
```

cess
e

North, Introducing BDD

Scenario: Account is in credit

Given the account is in credit
And the card is valid
And the dispenser contains cash
When the customer requests cash
Then ensure the account is debited
And ensure cash is dispensed
And ensure the card is returned



© felix tensing

Das Versprechen

- Testfälle werden in natürlicher Sprache formuliert
- BDD ist ein gesamter Prozess (nicht nur Test)
- "Nicht-Techies" definieren das Fachverhalten, und der Test ist gleichzeitig die Anforderung.
- Ziel: Living Documentation, die jeder lesen kann
- Dokumentation ist nicht nur Outcome, es ist auch Input.

"If you're using Cucumber solely for test automation, you're doing it wrong."
Aslak Hellesøy (Erfinder von Cucumber)



© felix tensing

Problem 1 - "non-technical stakeholders can understand"

- Gherkin ist eine DSL, keine natürliche Sprache
  a. Arbeiten benötigt extra Tools/IDEs nötig
  b. Formulierungen
     i. sind kompliziert (welche Actionswords gibt es schon)
     ii. sind Fehleranfällig (Actionword Reihenfolge)
- in ca. 80–90 % aller Projekte schreiben Entwickler oder QA-Automatisierer Gherkin

⇒ Es entsteht ein enormer Aufwand für ein Feature, das die Zielgruppe gar nicht nutzt.



© felix tensing

Problem 2: Cluecode of Hell

- Der Verbindungscode (Glue-Code/Step-Definitions) ist aufwendig zu Pflegen
- Bei Testfehlern sucht man erste Fehler im Gherkin, dann Fehler in der Implementierung
- Debug wird erschwert durch Indirektion
- Faktor Mensch: Arbeit die aufwendig/mühsam ist wird gescheut, minimalistisch erledigt

„Cucumber is just an extra layer of bullshit that doubles the work for QAs.“
Reddit r/QualityAssurance:



© felix tensing

Problem 3: Testsschnitt-Dilemma

- Gherkin-Beschreibungen sind Teil der Story
- Alle Gherkin-Beschreibungen werden als Test umgesetzt
- Betreibt man kein aktives Testmanagement

  => schnell steigende Testbasis

  => Wiedersprüchliche "Living Dokumentation"

  => Inversion der Testpyramide



© felix tensing

Die Utopie



© felix tensing

Der perfekte BDD Neuansatz

1. Echte natürliche Sprache für Testbeschreibung (Living Dokumentation)
2. k(l)eine technische Hürde (Jeder soll sich beteiligen können)
3. Testmanagement (keine Dopplungen)
4. Qualitätsprüfung (z.B. Testpyramide, Testslop verhindern)



© felix tensing

AI mein Retter



© felix tensing

Der Game-Changer - KI als Compiler

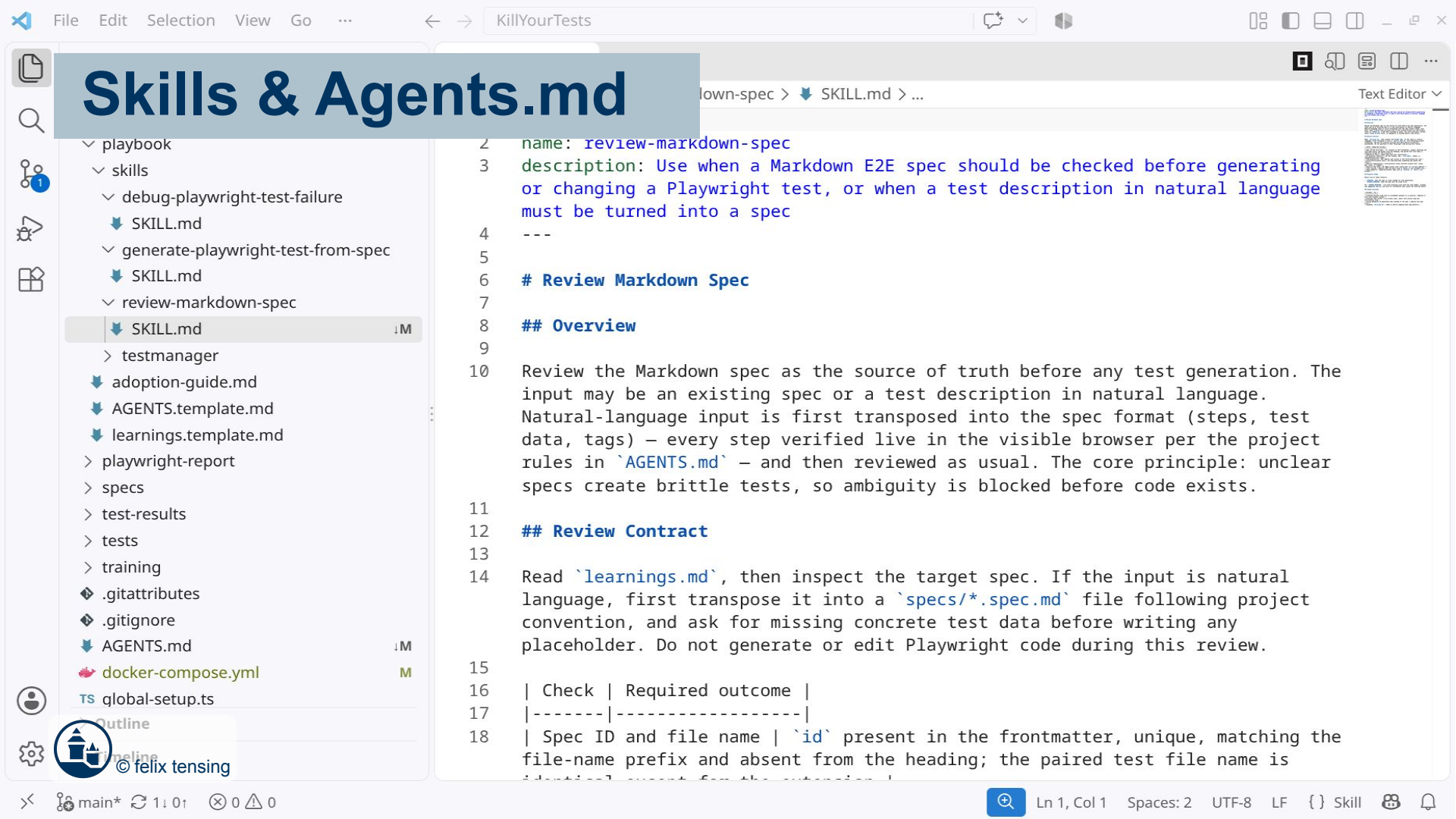
- AI hat ein Verständnis natürlicher Sprache
  a. Synonyme
  b. Nuancen
  c. Kontext
- Actionwordimplementierung entfällt
- Fachlichkeit übersteht ein Refactoring
- KI "kompiliert" in fast jeder Sprache/Framework (PlayWright, RestAssured, Postman …)

Filtere die Tabelle nach grünen Autos



© felix tensing

Skills & Agents.md

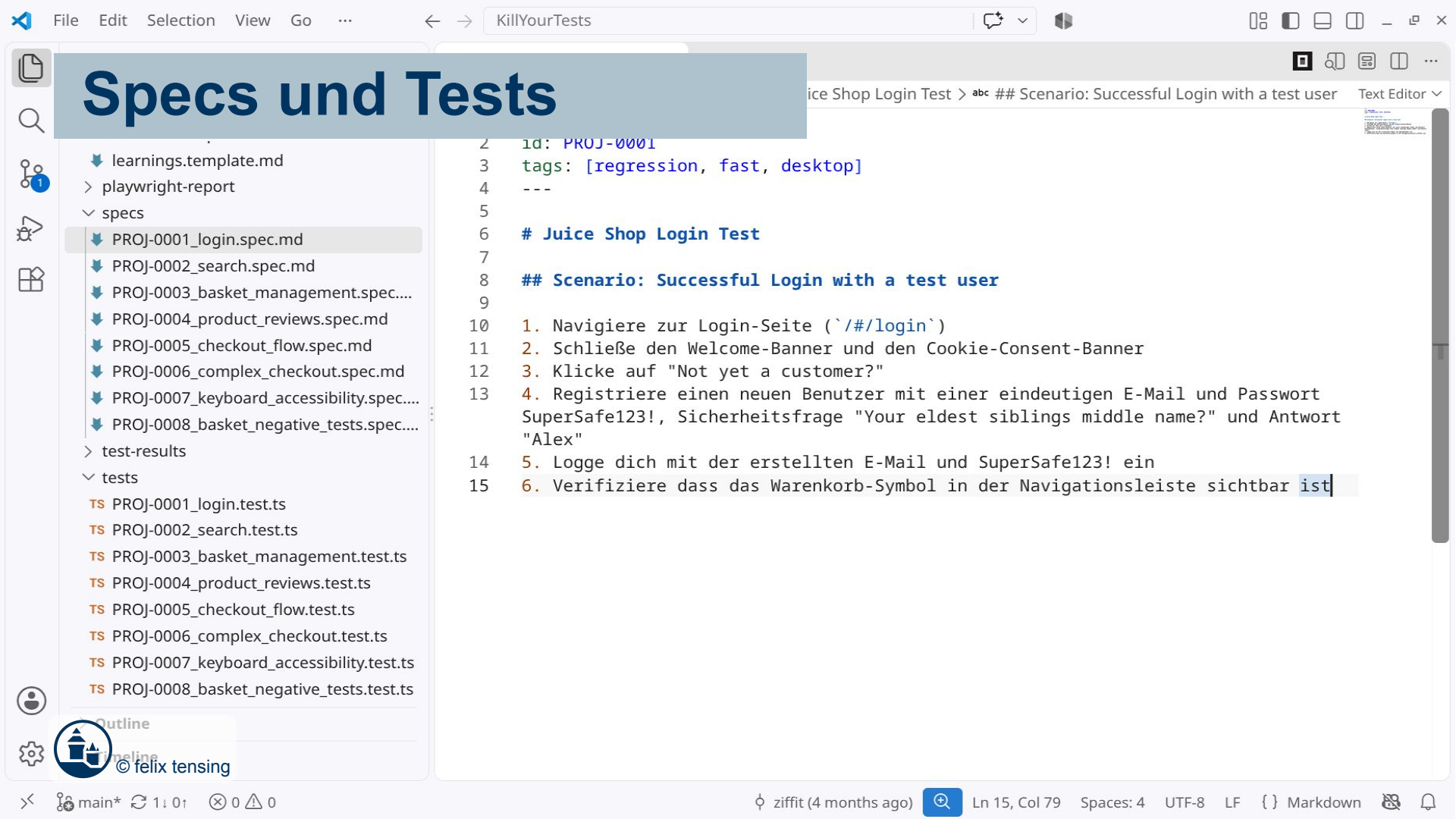


Learnings



© felix tensing

Specs und Tests

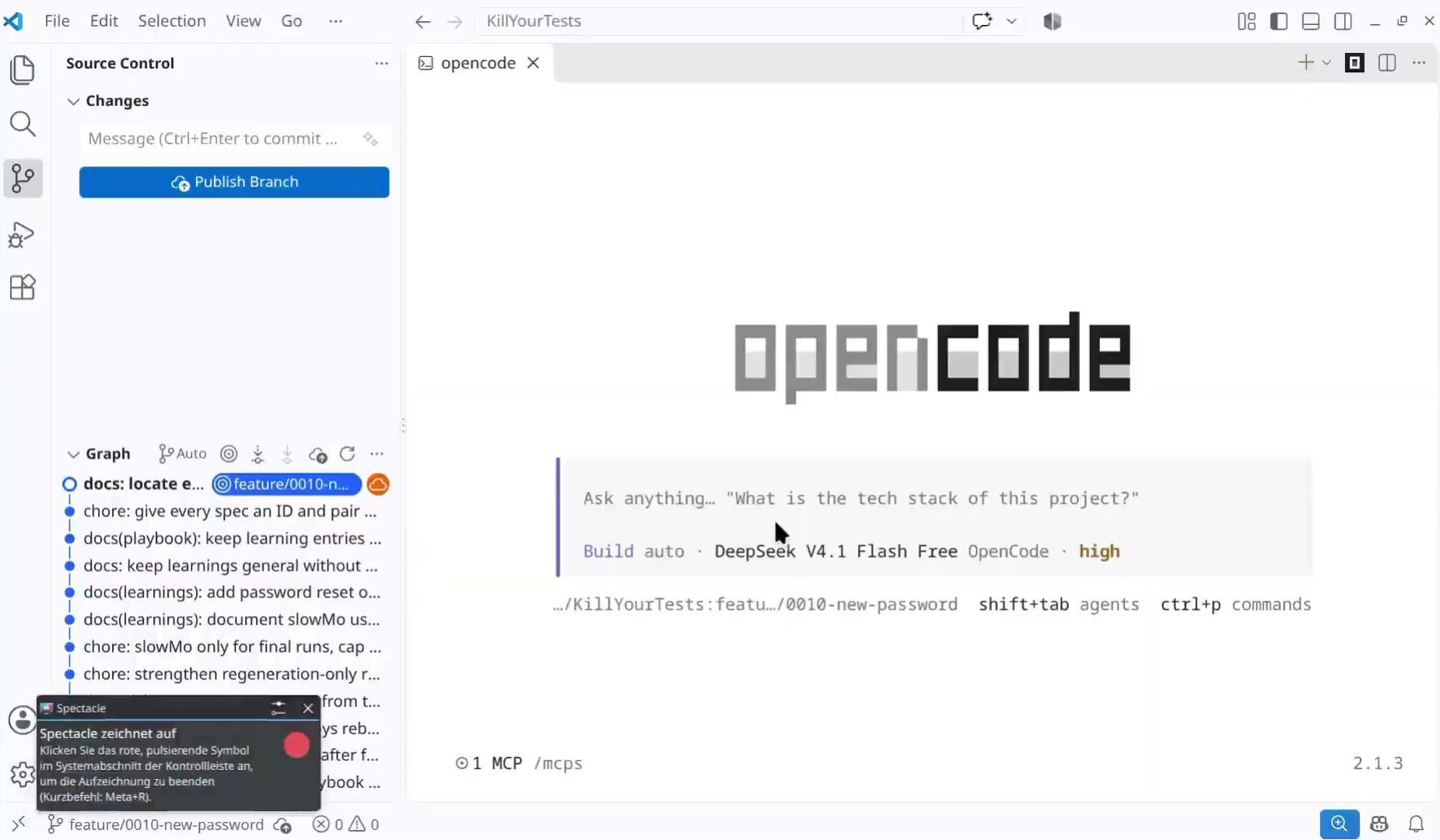


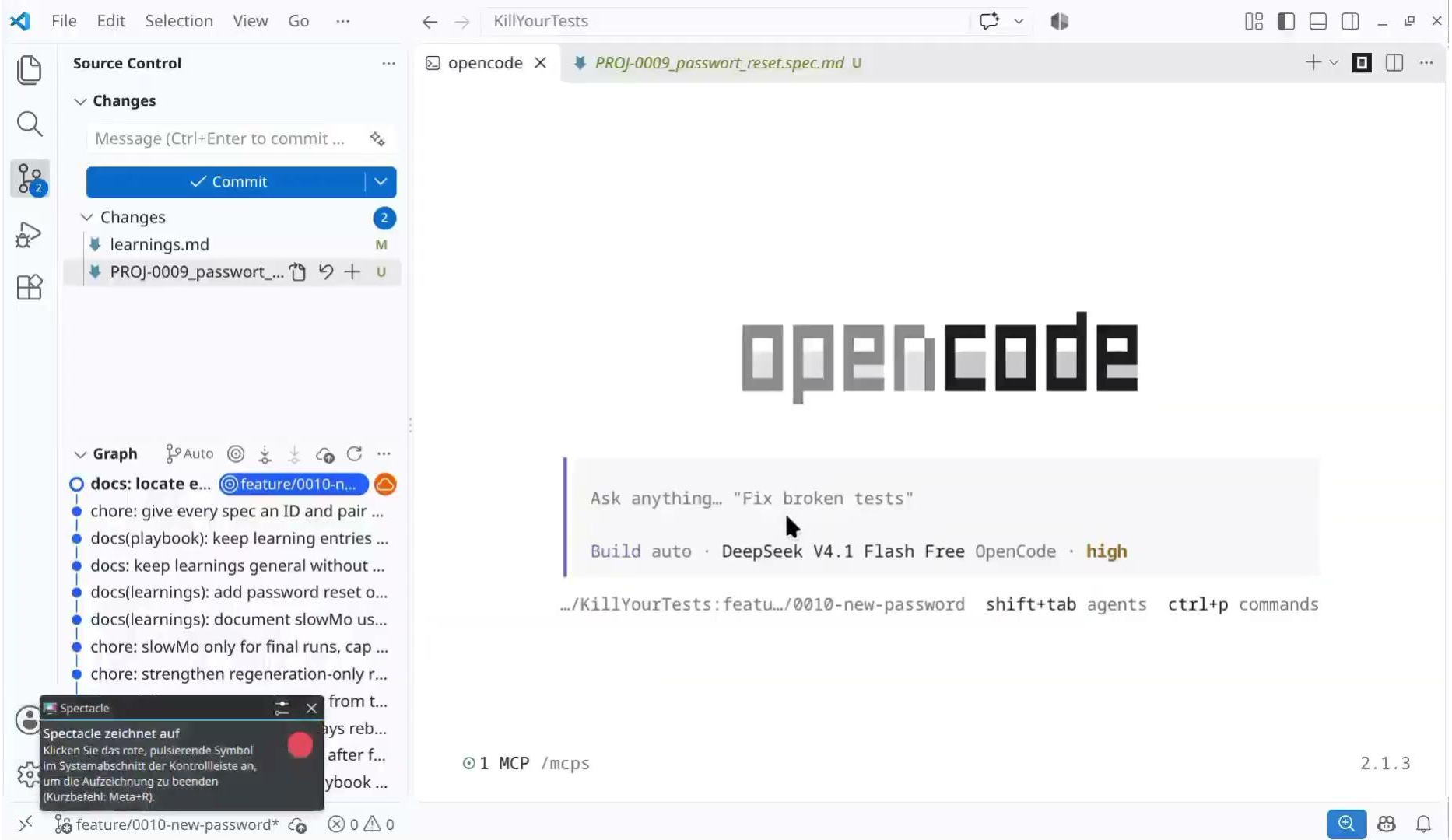
© felix tensing

Praxis



© felix tensing


File Edit Selection View Go
KillYourTests
Source Control
Changes
Message (Ctrl+Enter to commit ...
Publish Branch
opencode
Graph
Auto
docs: locate e...
feature/0010-n...
chore: give every spec an ID and pair ...
docs(playbook): keep learning entries ...
docs: keep learnings general without ...
docs(learnings): add password reset o...
docs(learnings): document slowMo us...
chore: slowMo only for final runs, cap ...
chore: strengthen regeneration-only r...
Spectacle
Spectacle zeichnet auf
Klicken Sie das rote, pulsierende Symbol im Systemabschnitt der Kontrollleiste an, um die Aufzeichnung zu beenden (Kurzbefehl: Meta+R).
opencode
Ask anything... "What is the tech stack of this project?"
Build auto · DeepSeek V4.1 Flash Free OpenCode · high
…/KillYourTests:featu…/0010-new-password shift+tab agents ctrl+p commands
1 MCP /mcps
2.1.3
feature/0010-new-password
0 0


File Edit Selection View Go
KillYourTests
Source Control
Changes
Message (Ctrl+Enter to commit ...
Commit
Changes 2
learnings.md M
PROJ-0009_passwort_... U
opencode
PROJ-0009_passwort_reset.spec.md U
opencode
Ask anything... "Fix broken tests"
Build auto · DeepSeek V4.1 Flash Free OpenCode · high
.../KillYourTests:featu.../0010-new-password shift+tab agents ctrl+p commands
Graph Auto
docs: locate e... feature/0010-n...
chore: give every spec an ID and pair ...
docs(playbook): keep learning entries ...
docs: keep learnings general without ...
docs(learnings): add password reset o...
docs(learnings): document slowMo us...
chore: slowMo only for final runs, cap ...
chore: strengthen regeneration-only r...
Spectacle
Spectacle zeichnet auf
Klicken Sie das rote, pulsierende Symbol im Systemabschnitt der Kontrollleiste an, um die Aufzeichnung zu beenden (Kurzbefehl: Meta+R).
1 MCP /mcps
2.1.3
feature/0010-new-password*
0 0

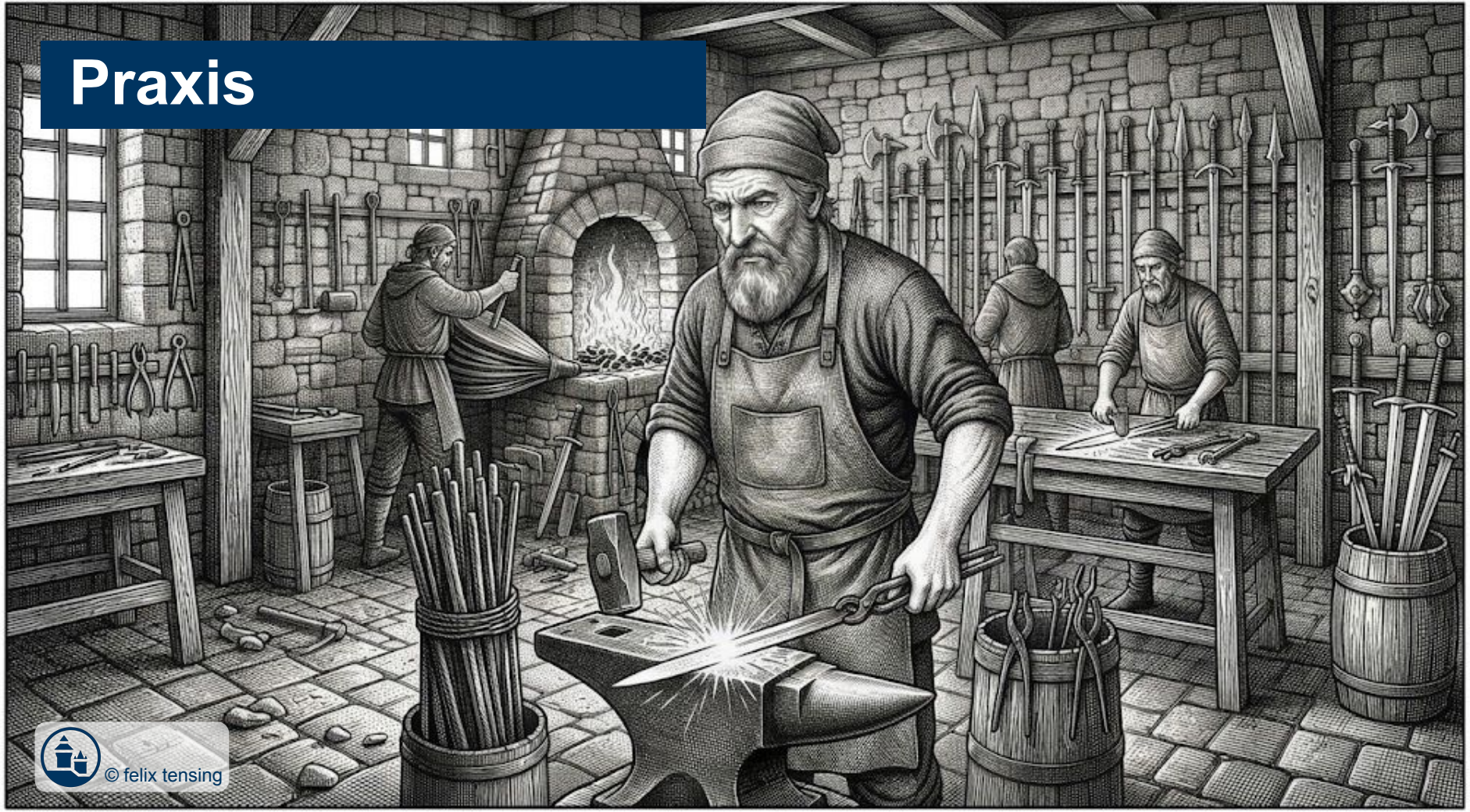
Warum "Töten" wir den Test?

1. Verhindern von "Drift"
    a. Die "Wahrheit" steht im Code
    b. Die Spec ist die fachliche Wahrheit (Aber nicht unveränderlich)
    c. ändern von Code ohne ändern der Spec verfälscht die Vereinfachung Spec = Code
2. Verifizieren der "Learnings"
3. Kein Einlesen/Debuggen in den Code



© felix tensing

Praxis



© felix tensing

Einsatz in der Praxis

1. **DoR:** Userstory beinhaltet Testbeschreibung
2. **Entwicklung:** Entwickler prüft mit KI bestehende Testbasis, ob ein Test erweitert werden kann/soll oder ein neuer entsteht
3. **DoD:** Geänderter / neue TestID wird in Story hinterlegt (Rückverfolgbarkeit), Code-Review auf Test
4. **Abnahme:** Spec ist Teil der Prüfung
5. **Dokumentation:** Spec, Testergebnis und Version werden ins Testreporting (aqua) übertragen (Szenario ist langlebiger als ein Test)

Nicht jede Story MUSS einen E2E Test haben (Testpyramide)



© felix tensing

Recap



© felix tensing

Vorteile

- Echte natürliche Sprache (BDD-Problem 1)
- Trennung fachliche Beschreibung und Testcode ohne "Glue Code" (BDD-Problem 2)
- KI unterstützt beim Testmanagement (BDD-Problem 3)
- Keine Abhängigkeiten zu proprietären (BDD) Frameworks
- Technologie Zweitrangig
- Kaum technisches Know-How nötig (Code-Review)
- KI erstellt Tests, aber Testcode ist statisch
- KI kann über Fixtures externe Systeme Abfragen, Testdaten erstellen usw.
- Wartungskosten viel geringer
- …



© felix tensing

und …

- Vision-Modelle (Statt Pixel-Vergleich)
- Testmanager Skill (keine versteckten Unittests, Abdeckung, Doppelungen)
- Automatisierte Behebung von kaputten Tests (wie dependency-bot)
- Fachliche Dokumentation mit Tests verbinden

⚠️ **Mißbrauchsgefahr! Richtigen Einsatz beachten!**



© felix tensing

Danke fürs Zuhören und erfolgreiches Drachentöten!

linkedin.com/in/felix-tensing-nbg



felix-spricht.de



Ich freue mich auf Gespräche und Feedback!

© felix tensing

Test Promt PROJ-0009

Vorbedingung: Die Sprache der Anwendung ist auf Deutsch eingestellt.
Auflösung: Mobile
Szenario:

- Ein neuer Nutzer registriert sich mit einer eindeutigen E-Mail-Adresse, dem Passwort SuperSafe123!, der Sicherheitsfrage „Zweitname deines ältesten Geschwisters?" und der Antwort „Alex".
- Nach der Registrierung öffnet er die Anmeldeseite, klickt auf „Passwort vergessen?" und gibt seine E-Mail-Adresse ein.
- Er beantwortet die Sicherheitsfrage mit „Alex" und legt das neue Passwort NewPass456! fest.
- Daraufhin wird ihm die Meldung „Dein Passwort wurde erfolgreich geändert." angezeigt.
- Jetzt meldet er sich mit dem neuen Passwort an und sieht das Warenkorb-Symbol oben rechts in der Navigationsleiste.
- Zum Schluss meldet er sich wieder ab und versucht, sich mit dem alten Passwort SuperSafe123! anzumelden. Das gelingt nicht und es erscheint die Fehlermeldung „Email-Adresse oder Passwort ungültig."

© felix tensing