Software Practicals
Summer Semester 2026

Data Science Group
Heidelberg University
April 15, 2026

Slides Online

Scan to access course materials

Access the Slides

The slides and other relevant course information are available on our official teaching webpage.

Direct Link:

ds.ifi.uni-heidelberg.de/teaching/current/

Note: Content is updated regularly throughout the semester.

Organization

Outline

1. Topic Assignment

Overview of topics (today)

- Send application for a topic until **Monday, April 20, 1pm**
- Assignment of topics through advisors by **April 23**

2. First Milestone

End of May

- 1st Prototype
- Summary of research (literature and related systems, tools, data, …)
- Further milestones in agreement with your supervisor

3. Practical End

End July / Beginning of August

- Code submission to local Gitlab of Data Science group
- Presentation of your project (10-12 minutes)
- Report / Documentation as README.md

Application Process

How to Apply

Apply directly to supervisor(s) via email. Your email must include:

- Personal Info: Program, semester, matriculation number
- Relevant course experience and grades
- Other experience (side projects, "Anwendungsgebiet", Job/Project experience)
- Tentative schedule and milestones for the practical
- **Note**: Most projects are team projects

Pro-tip: You might want to apply for multiple topics (e.g., your top-3 list) to increase chances.

error_out
line

Binding Application

Your application is **binding** once submitted.

Don't apply if you don't intend to complete the practical!

Deadlines

Meeting Cadence

- Per default, generally meetings with supervisor every other week.
- **Come prepared for the meetings (e.g., 2-3 slides)!**

Key Dates

- **Presentations:** Last week of July 2026
- **Report/Code in GitLab:** August 4, 2026

error_outline

No Extensions

No extension possible.

Not finished = failed (grade 5,0)!

Assessment

Credit points (Leistungspunkte)

- **Beginners / Bachelor Practical (IAP, 2 CP + 4 FÜK)**
  workload: 180 h (about 1.5 days/week)
- **Advanced / Master Practical (IFP / IMP, 8 CP)**
  workload: 240 h (about 2 days/week)

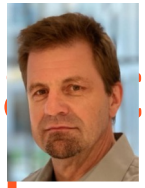
Grading Criteria

- Code (readability, structure, functionality; local GitLab)
- Documentation (README.md, GitLab comments)
- Commitment and self-reliance
- ***Cool ideas!!***

IMPORTANT: Regular communication with your advisor

(per default biweekly meetings)

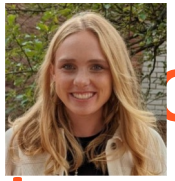
Supervisors



Michael Gertz (MG)

Professor / Supervisor

gertz@informatik.uni-heidelberg.de



Marina Walther (MW)

PhD Candidate / Advisor

walther@informatik.uni-heidelberg.de

Jonas Gann (JG)

PhD Candidate / Advisor

gann@informatik.uni-heidelberg.de

Contact your respective advisor for specific topic questions.

Project Topics

AP = Advanced Topic

Typically suited for Master level students or advanced Bachelor students.

BP = Beginners Topic

Designed specifically for BSc students starting their research journey.

Refer to the following slides for specific project descriptions.

Overview of Topics

1. Medical QA: Re-evaluating LLMs – **BP/AP** (Walther)
2. diabinfo: data extraction and RAG pipeline – **AP** (Walther)
3. Medical QA benchmark based on DDG guidelines – **AP** (Walther)
4. Debugging and Test Suite for Symbolic RAG framework – **BP/AP** (Gann)
5. Adapting a symbolic RAG framework for diverse input modalities – **BP/AP** (Gann)
6. AI-Based Multi-Agent Tutoring System – **BP/AP** (Gertz)
7. Multi-Agent Forklift Coordination in a 2D Grid Environment – **BP/AP** (Gertz)
8. From Natural Language Rules to Executable Game Logic – **BP/AP** (Gertz)

BP/AP: Medical QA: Re-evaluating LLMs (MW)

Given:

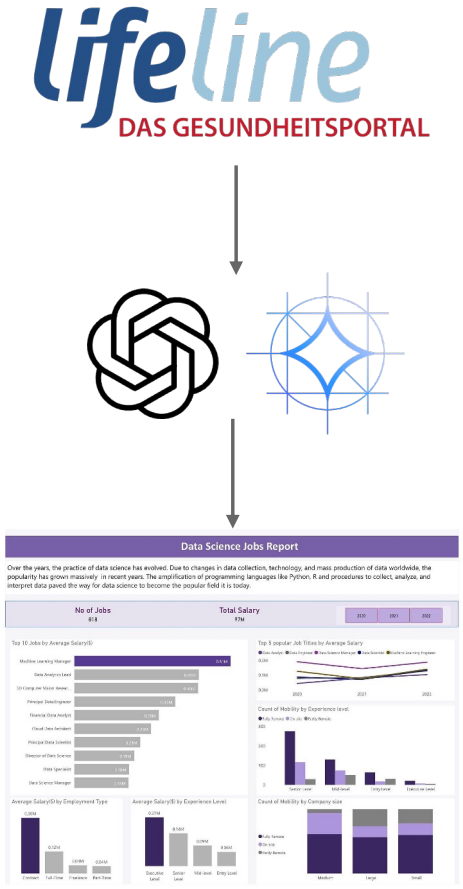
- German medical QA datasets, e.g., Lifeline Expertenrat
- Paper, prior experiments

Tasks:

- Re-evaluate experiments with recent models of different sizes
- Adapt experiments to “reasoning” capabilities of modern models

Subtasks:

- Familiarize with existing experimental setups and datasets
- Setup modular LLM inference pipeline
- Test the current best practices for prompting LLMs
- Implement dashboard to inspect the results



Tools:

Python, Ollama, Langchain/LlamaIndex, Streamlit or js-framework, public APIs (OpenAI), cloud services (GCC), matplotlib, pandas, …

AP: diabinfo: data source and RAG pipeline (MW)

Given:

- Diabetes information website for non-medical people
  diabinfo.de

Tasks:

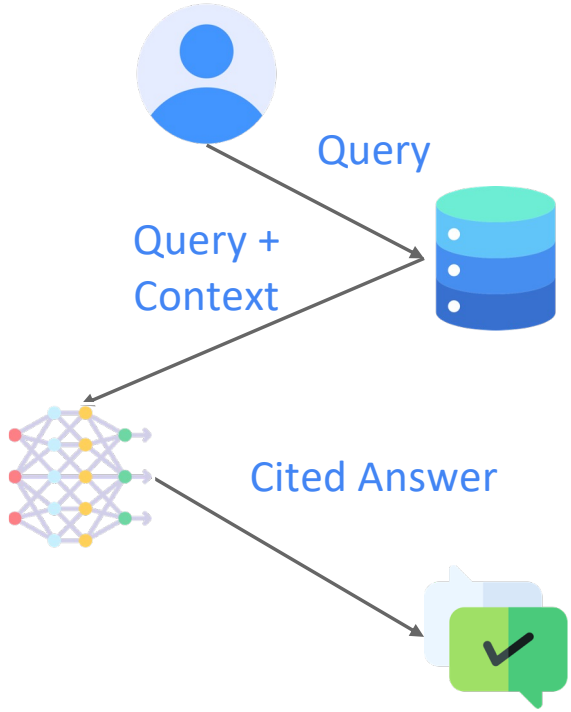
- Crawl website → transfer to database
- Setup RAG pipeline to answer user questions

Subtasks:

- Quantitative analysis of dataset: size, topics, level of professionalism
- Build chat-interface to interact with RAG pipeline

Tools:

Docker + Docker Compose, Vector-DB, Python, Ollama, Streamlit or js-framework, public APIs (OpenAI), cloud services (GCC), matplotlib, pandas, …

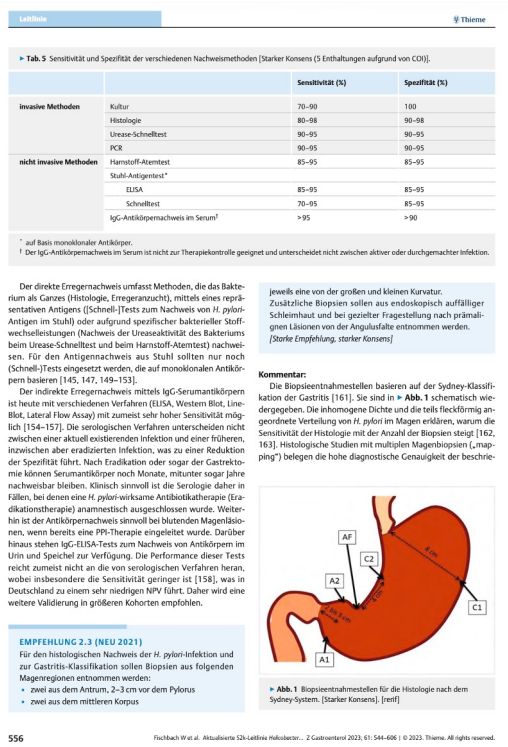


AP: Medical QA benchmark based on DDG guidelines (MW)

Given:

- PDF collection of official diabetes guidelines from Deutsche Diabetes Gesellschaft (DDG)

DDG Deutsche Diabetes Gesellschaft



Leitlinie — Thieme

Tab. 5 Sensitivität und Spezifität der verschiedenen Nachweismethoden [Starker Konsens (5 Enthaltungen aufgrund von COI)].

		Sensitivität (%)	Spezifität (%)
invasive Methoden	Kultur	70–90	100
	Histologie	80–98	90–98
	Urease-Schnelltest	90–95	90–95
	PCR	90–95	90–95
nicht invasive Methoden	Harnstoff-Atemtest	85–95	85–95
	Stuhl-Antigentest*		
	ELISA	85–95	85–95
	Schnelltest	70–95	85–95
	IgG-Antikörpernachweis im Serum†	>95	>90

* auf Basis monoklonaler Antikörper.
† Der IgG-Antikörpernachweis im Serum ist nicht zur Therapiekontrolle geeignet und unterscheidet nicht zwischen aktiver oder durchgemachter Infektion.

Der direkte Erregernachweis umfasst Methoden, die das Bakterium als Ganzes (Histologie, Erregeranzucht), mittels eines repräsentativen Antigens ([Schnell-]Tests zum Nachweis von *H. pylori*-Antigen im Stuhl) oder aufgrund spezifischer bakterieller Stoffwechselleistungen (Nachweis der Ureaseaktivität des Bakteriums beim Urease-Schnelltest und beim Harnstoff-Atemtest) nachweisen. Für den Antigennachweis aus Stuhl sollten nur noch (Schnell-)Tests eingesetzt werden, die auf monoklonalen Antikörpern basieren [145, 147, 149–153].

Der indirekte Erregernachweis mittels IgG-Serumantikörpern ist heute mit verschiedenen Verfahren (ELISA, Western Blot, Line-Blot, Lateral Flow Assay) mit zumeist sehr hoher Sensitivität möglich [154–157]. Die serologischen Verfahren unterscheiden nicht zwischen einer aktuell existierenden Infektion und einer früheren, inzwischen aber eradizierten Infektion, was zu einer Reduktion der Spezifität führt. Nach Eradikation oder sogar der Gastrektomie können Serumantikörper noch Monate, mitunter sogar Jahre nachweisbar bleiben. Klinisch sinnvoll ist die Serologie daher in Fällen, bei denen eine *H. pylori*-wirksame Antibiotikatherapie (Eradikationstherapie) anamnestisch ausgeschlossen wurde. Weiterhin ist der Antikörpernachweis sinnvoll bei blutenden Magenläsionen, wenn bereits eine PPI-Therapie eingeleitet wurde. Darüber hinaus stehen IgG-ELISA-Tests zum Nachweis von Antikörpern im Urin und Speichel zur Verfügung. Die Performance dieser Tests reicht zumeist nicht an die von serologischen Verfahren heran, wobei insbesondere die Sensitivität geringer ist [158], was in Deutschland zu einem sehr niedrigen NPV führt. Daher wird eine weitere Validierung in größeren Kohorten empfohlen.

EMPFEHLUNG 2.3 (NEU 2021)
Für den histologischen Nachweis der *H. pylori*-Infektion und zur Gastritis-Klassifikation sollen Biopsien aus folgenden Magenregionen entnommen werden:
- zwei aus dem Antrum, 2–3 cm vor dem Pylorus
- zwei aus dem mittleren Korpus

jeweils eine von der großen und kleinen Kurvatur.
Zusätzliche Biopsien sollen aus endoskopisch auffälliger Schleimhaut und bei gezielter Fragestellung nach prämalignen Läsionen von der Angulusfalte entnommen werden.
[Starke Empfehlung, starker Konsens]

Kommentar:
Die Biopsieentnahmestellen basieren auf der Sydney-Klassifikation der Gastritis [161]. Sie sind in ▸ **Abb. 1** schematisch wiedergegeben. Die inhomogene Dichte und die teils fleckförmig angeordnete Verteilung von *H. pylori* im Magen erklären, warum die Sensitivität der Histologie mit der Anzahl der Biopsien steigt [162, 163]. Histologische Studien mit multiplen Magenbiopsien („mapping") belegen die hohe diagnostische Genauigkeit der beschrie-

▸ **Abb. 1** Biopsieentnahmestellen für die Histologie nach dem Sydney-System. [Starker Konsens]. [rerif]

Fischbach W et al. Aktualisierte S2k-Leitlinie Helicobacter... Z Gastroenterol 2023; 61: 544–606 | © 2023. Thieme. All rights reserved.

Tasks:

- Extract (text)-data from PDF
- Create QA- pairs similar to https://sharc-data.github.io/

Subtasks:

- Create suitable data model for typical QA-problems: single-turn QA, attribution using document-chunks
- Implement modular pipeline: from extraction to storage
- Evaluate created benchmark(s)

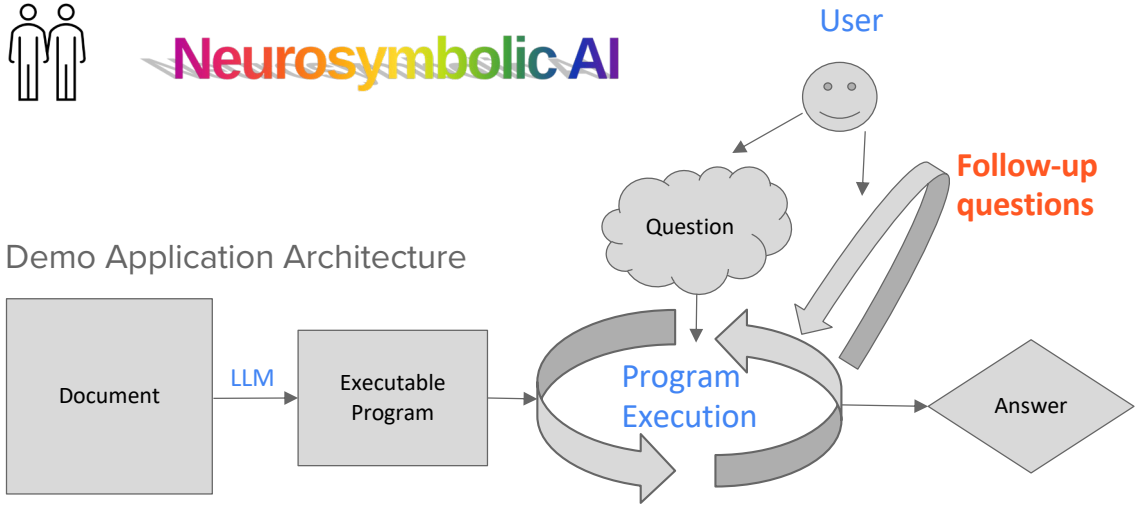
Tools:

Docker + Docker Compose, Vector-DB, Python, Ollama, public APIs (OpenAI), cloud services (GCC), matplotlib, pandas, …

AP/BP: Input Modalities (JG)

Given:

- **Demo Application**
- Sample Documents
  - e.g. Medical Guidelines, …



Task:

- Follow-up questions currently only **yes/no**. Extend the demo to use additional modalities (e.g. numbers, strings, selections, …) => What is your age? vs. Are you above 18?

Subtasks:

- **Catch Up:** Understand existing codebase and related literature
- **Logic:** Improve LLM code generation to incorporate new modalities for follow-up questions
- **UI:** Expose new modalities in the UI to the user
- **Validation:** Check the correct usage of available modalities

Languages / Tools / Platforms:

AP/BP: Testing and Debugging Suite (JG)

Neurosymbolic AI

Given:

- **Demo Application**
- Sample Documents
  - e.g. Medical Guidelines, ...

Sample Test Cases

Question	Followup Question + User Answer	Expected Final Answer
Is my blood sugar elevated?	is_blood_sugar_level_above_200mgdl? User answer: **yes**	**Yes**, your blood sugar level is elevated
Is my blood sugar elevated?	is_blood_sugar_level_above_200mgdl? User answer: **no**	**No**, your blood sugar level is not elevated

Task:

- Extend the demo application to enable **convenient testing** of generated programs

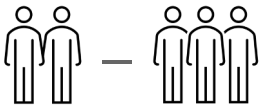
Subtasks:

- **Catch Up:** Understand existing codebase and related literature
- **UI** for manual specification and execution of system test cases
- **Automated evaluation** of system responses to test questions

Languages / Tools / Platforms:

Python, Streamlit, Ollama, OpenAI, Prolog

AP/BP: AI-Based Multi-Agent Tutoring System (MG)



Given:

- (CS) course materials such as textbooks, slides, assignments

Tasks:

- Design and implement prototype of an intelligent tutoring system based on multi-agent architecture that utilizes given materials
- System shall explain concepts, generate exercises, provide structured hints

Subtasks:

- System is based on interacting agents (Teaching Agent, Exercise Agent, Hint Agent, and – optionally – Learning Progress Agent)
- Design and implement modular and full stack system architecture, incl. frontend

Languages, Tools, and Platforms:

- Python, CrewAI, Streamlit, PostgreSQL / SQLite

AP/BP: Multi-Agent Forklift Coordination in a 2D Grid Environment (MG)

Given:

- Abstract 2D gridworld environment for forklifts, obstacles, and assets

Tasks: (in collaboration with Flowcate)

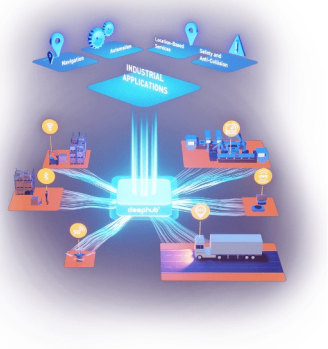
- Simulate spatial environment (simulation) & tracking integration
- Design and implement multi-agent coordination framework

Subtasks:

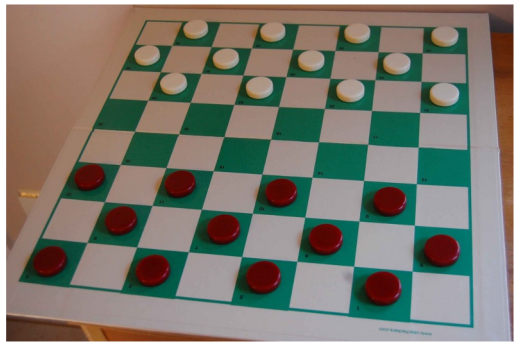
- Managing agent decomposes high-level transport orders into individual pickup-and-delivery tasks
- Individual forklift agents, one per vehicle

Languages / Tools / Platforms:

- Python, CrewAI, omlox, DeepHub, Streamlit …



AP/BP: From Natural Language Rules to Executable Game Logic (MG)



Given:

- Machine-readable rulebooks (as plain text or PDFs)

Tasks:

- LLM-to-Prolog translation pipeline
- Interactive rule refinement and query interface

Subtasks:

- Multi-game generalisation. Evaluate pipeline on at least two structurally different games
- Implement lightweight game state engine driven entirely by the generated Prolog program, capable of stepping through a game move by move.

Languages / Tools / Platforms:

- Python, LLM APIs, SWI-Prolog, Streamlit

Slides Online

Scan to access slides

Access the Slides

The slides and other relevant course information are available on our official teaching webpage.

Direct Link:

ds.ifi.uni-heidelberg.de/teaching/current/

Note: Content is updated regularly throughout the semester.