



FernUniversität
in Hagen



Automatisierungspotenziale in der Datenbanklehre

Aufgaben, Bewertung, Feedback, Personalisierung

Paul Christ

AUTOMATISIERUNGSPOTENZIALE IN DER DATENBANK-LEHRE

FernUniversität in Hagen
Paul Christ





Vortragsstruktur

1. Motivation aus drei Perspektiven
2. Gegenstand der Automatisierung
3. Anwendungsfälle
4. (Zwischen)Fazit
5. Technical (Deep) Dive am Beispiel von SQL-Aufgaben
6. Fragen & Diskussion



1. Interventionen im Lehr-Lern-Prozess und ihre Auswirkung auf den Lerneffekt

	<i>Visible Learning</i> (2009)	<i>Visible Learning for Teachers</i> (2013)	<i>Visible Learning Insights</i> (2019)	<i>Visible Learning: The Sequel</i> (2023)
Number of meta-analyses	816	931	1,412	2,313
Number of primary studies	52,469	60,167	82,955	132,389
Number of students	approx. 200 million	approx. 240 million	approx. 300 million	approx. 400 million

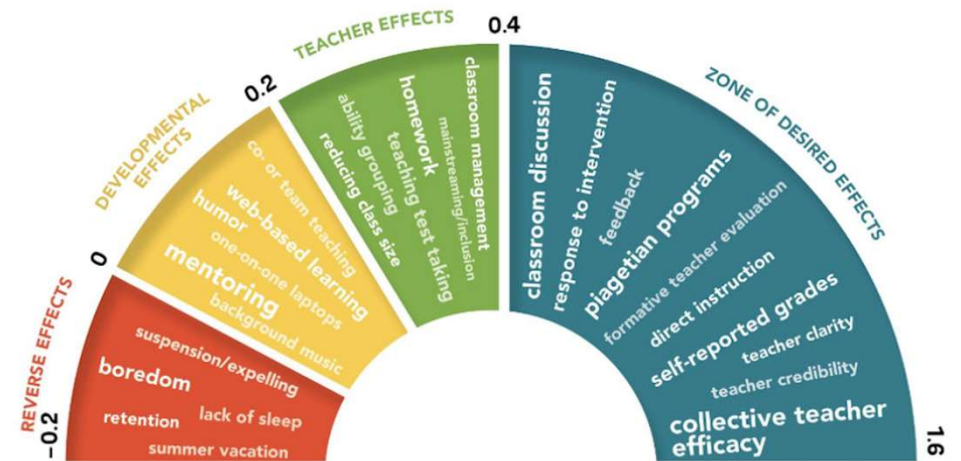
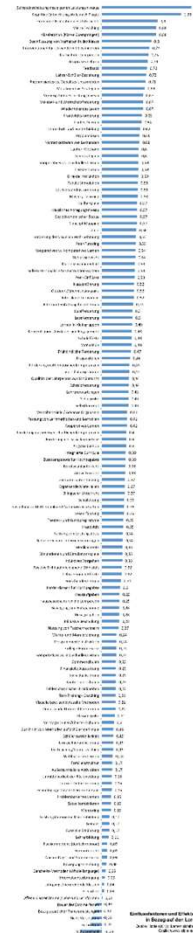
Hattie, J., & Zierer, K. (2025). *Visible learning: Lesson planning: An evidence-based guide for successful teaching*. Routledge.



1. Interventionen im Lehr-Lern-Prozess und ihre Auswirkung auf den Lerneffekt

	Visible Learning (2009)	Visible Learning for Teachers (2013)	Visible Learning Insights (2019)	Visible Learning: The Sequel (2023)
Number of meta- analyses	816	931	1,412	2,313
Number of primary studies	52,469	60,167	82,955	132,389
Number of students	approx. 200 million	approx. 240 million	approx. 300 million	approx. 400 million

Hattie, J., & Zierer, K. (2025). *Visible learning: Lesson planning: An evidence-based guide for successful teaching*. Routledge.

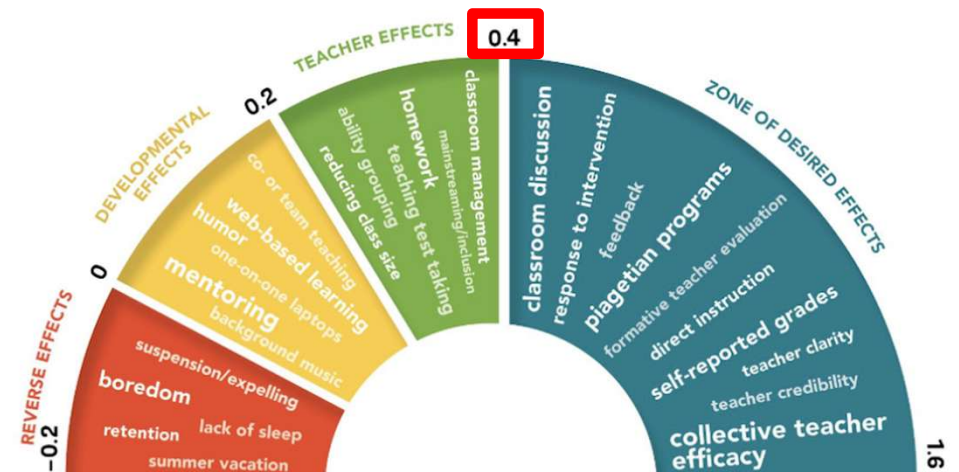
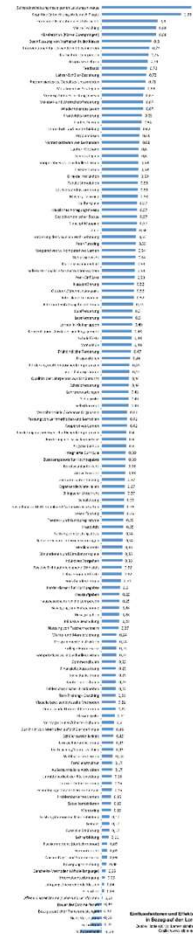




1. Interventionen im Lehr-Lern-Prozess und ihre Auswirkung auf den Lerneffekt

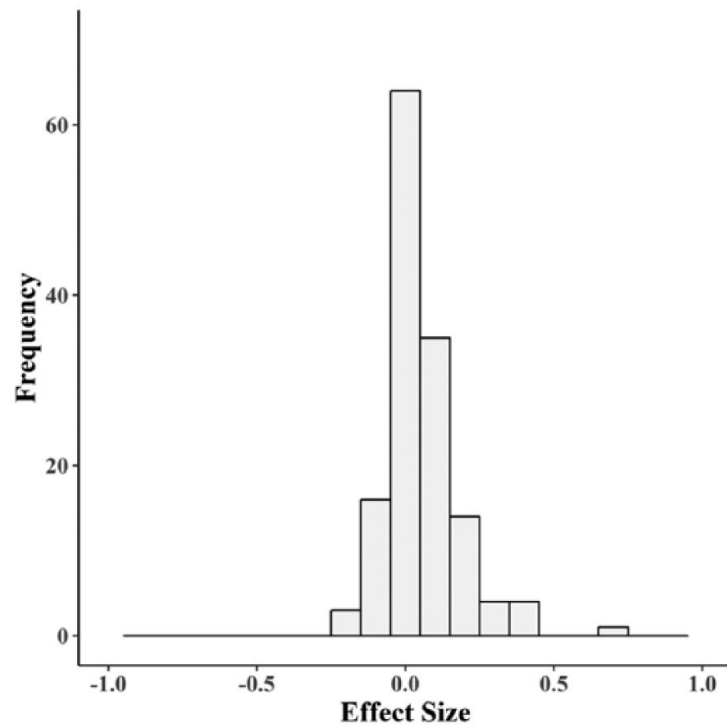
	Visible Learning (2009)	Visible Learning for Teachers (2013)	Visible Learning Insights (2019)	Visible Learning: The Sequel (2023)
Number of meta- analyses	816	931	1,412	2,313
Number of primary studies	52,469	60,167	82,955	132,389
Number of students	approx. 200 million	approx. 240 million	approx. 300 million	approx. 400 million

Hattie, J., & Zierer, K. (2025). *Visible learning: Lesson planning: An evidence-based guide for successful teaching*. Routledge.





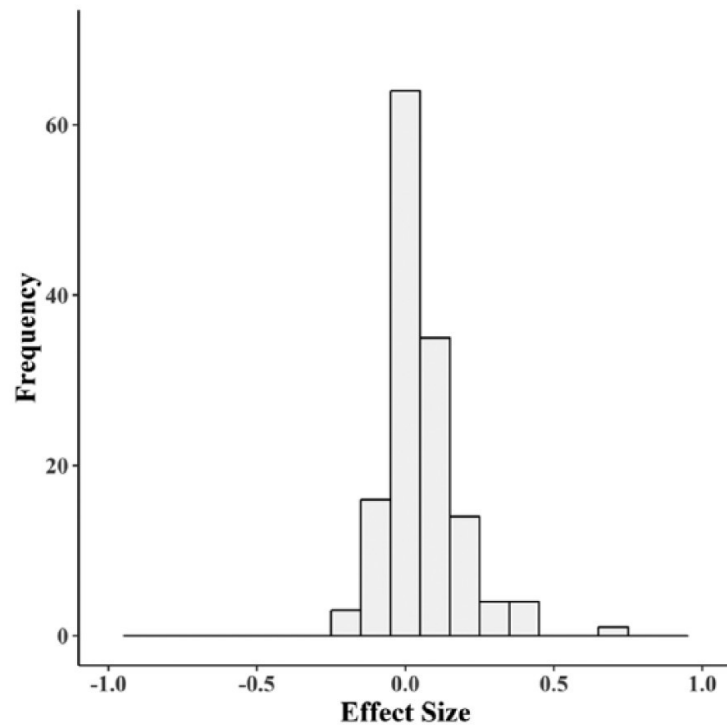
1. Bildungsforschung ist komplex und belastbare Zahlen rar



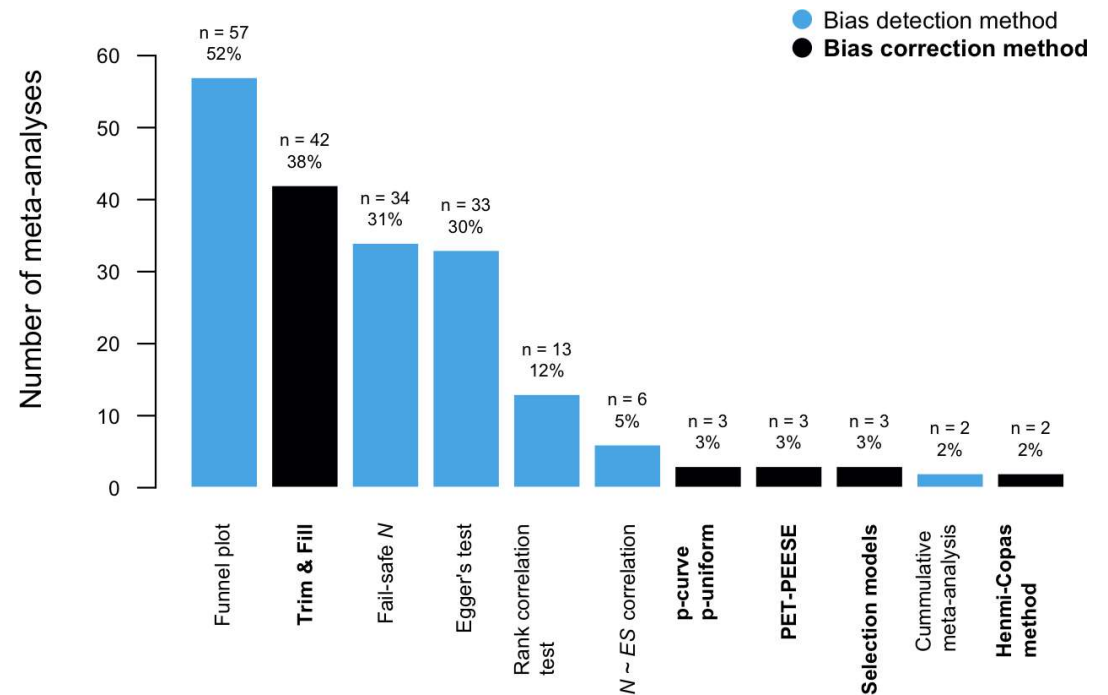
Lortie-Forgues, H., & Inglis, M. (2019). Rigorous Large-Scale Educational RCTs Are Often Uninformative: Should We Be Concerned? *Educational Researcher*, 48(3), 158-166. <https://doi.org/10.3102/0013189X19832850>



1. Bildungsforschung ist komplex und belastbare Zahlen rar



Lortie-Forgues, H., & Inglis, M. (2019). Rigorous Large-Scale Educational RCTs Are Often Uninformative: Should We Be Concerned? *Educational Researcher*, 48(3), 158-166. <https://doi.org/10.3102/0013189X19832850>



Neglect of publication bias compromises meta-analyses of educational research

Ropovik I, Adamkovic M, Greger D (2021) Neglect of publication bias compromises meta-analyses of educational research. *PLOS ONE* 16(6): e0252415. <https://doi.org/10.1371/journal.pone.0252415>



1. Motivation zur Automatisierung

Schlussfolgerung aus der Empirie:



1. Motivation zur Automatisierung

Schlussfolgerung aus der Empirie:

Wie man sieht, sieht man nichts... _(ツ)_/



1. Motivation zur Automatisierung

Schlussfolgerung aus der Empirie:

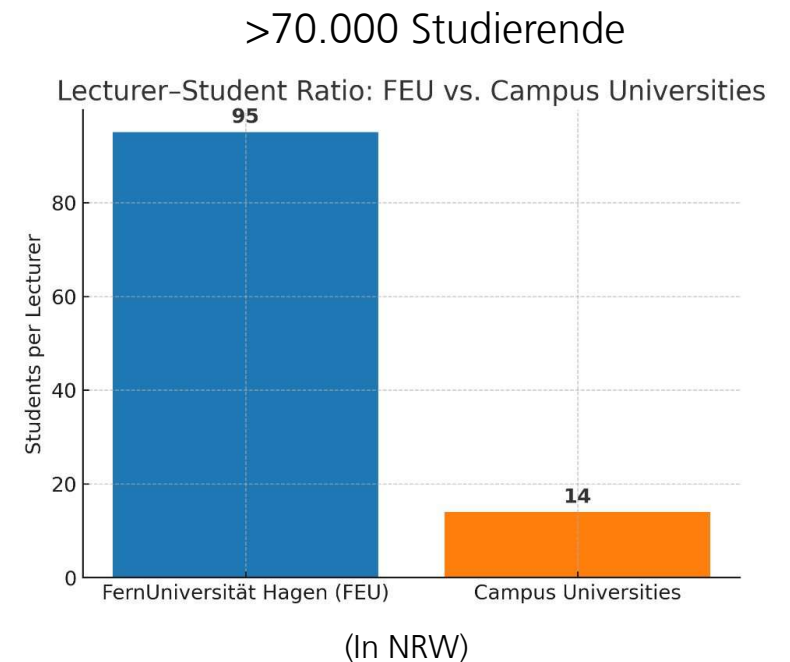
Anderer Blickwinkel – weg vom Lerneffekt?

Wie man sieht, sieht man nichts... _(ツ)_/



1. Aus der Praxis resultierende Motivation an der FernUni

- Ungünstiges Verhältnis zwischen Lehrkräften und Studierenden an der FeU
- Mangel an personalisierten Übungen
- Verzögerte Rückmeldung für „Einsendeaufgaben“
- Eingeschränkte Kommunikation zwischen Studierenden und Dozenten



Pressestelle I. und T. NRW. (2023, November 27). Zahl der Studierenden an NRW-Hochschulen weiter rückläufig. Landesbetrieb IT.NRW. <https://www.it.nrw/zahl-der-studierenden-nrw-hochschulen-weiter-rueckklaeufig-125789>



1. Automatisierungspotenziale des Lernens?!

... oder was Studis manchmal denken unter Lernen zu verstehen.

➔ Prüfung bestanden = Ich habe „es“ gelernt!



„Informatik ist cool!“

...



„Mit dem Wisch
zahlt mir mein AG
das doppelte!“

1. Automatisierungspotenziale des Lernens?!

... oder was Studis manchmal denken unter Lernen zu verstehen.

➔ Prüfung bestanden = Ich habe „es“ gelernt!

Katz-und-Maus-Jagd

Stetig neue Übungen und Prüfungen, um die Studis auf Trab zu halten



„Informatik ist cool!“

...



„Mit dem Wisch
zahlt mir mein AG
das doppelte!“



1. Automatisierungspotenziale des Lernens?!

... oder was Studis manchmal denken unter Lernen zu verstehen.

➔ Prüfung bestanden = Ich habe „es“ gelernt!

Katz-und-Maus-Jagd

Stetig neue Übungen und Prüfungen, um die Studis auf Trab zu halten

Wer jagt eigentlich wen?
(und nicht erst seit ChatGPT und Co.)

*Tatsächliches Lernen lässt sich natürlich noch nicht automatisieren
(zumindest bis ein Gehirn-Computer-Schnittstellen „Direkteinspielung“ erlauben)*



„Informatik ist cool!“

...

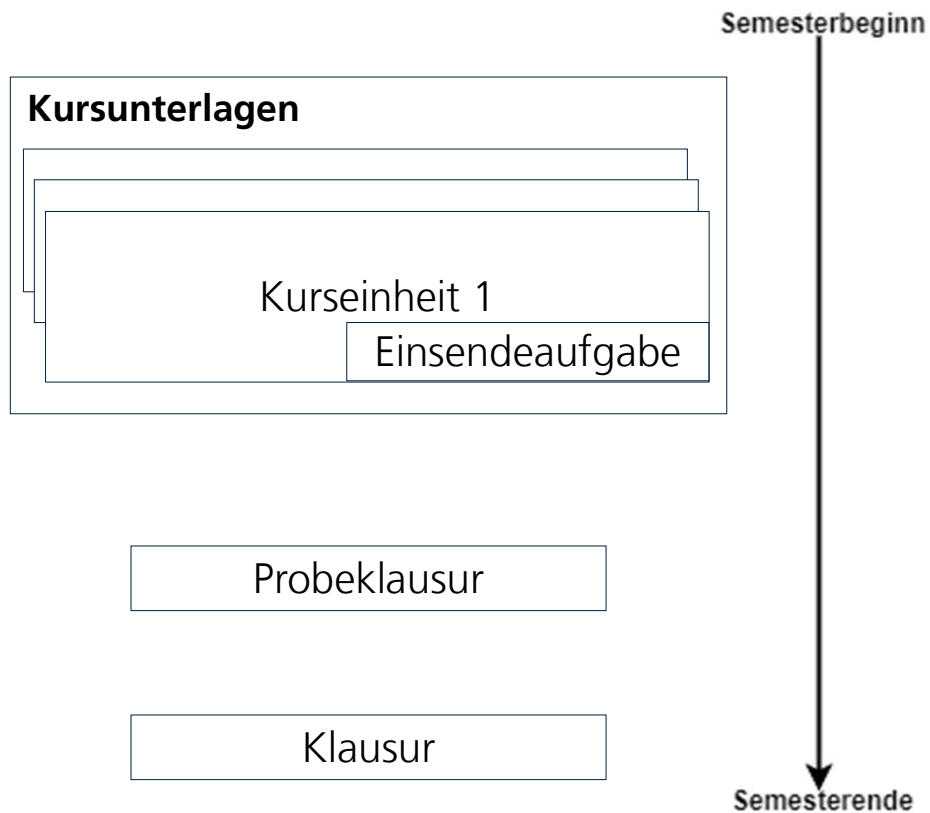


„Mit dem Wisch
zahlt mir mein AG
das doppelte!“



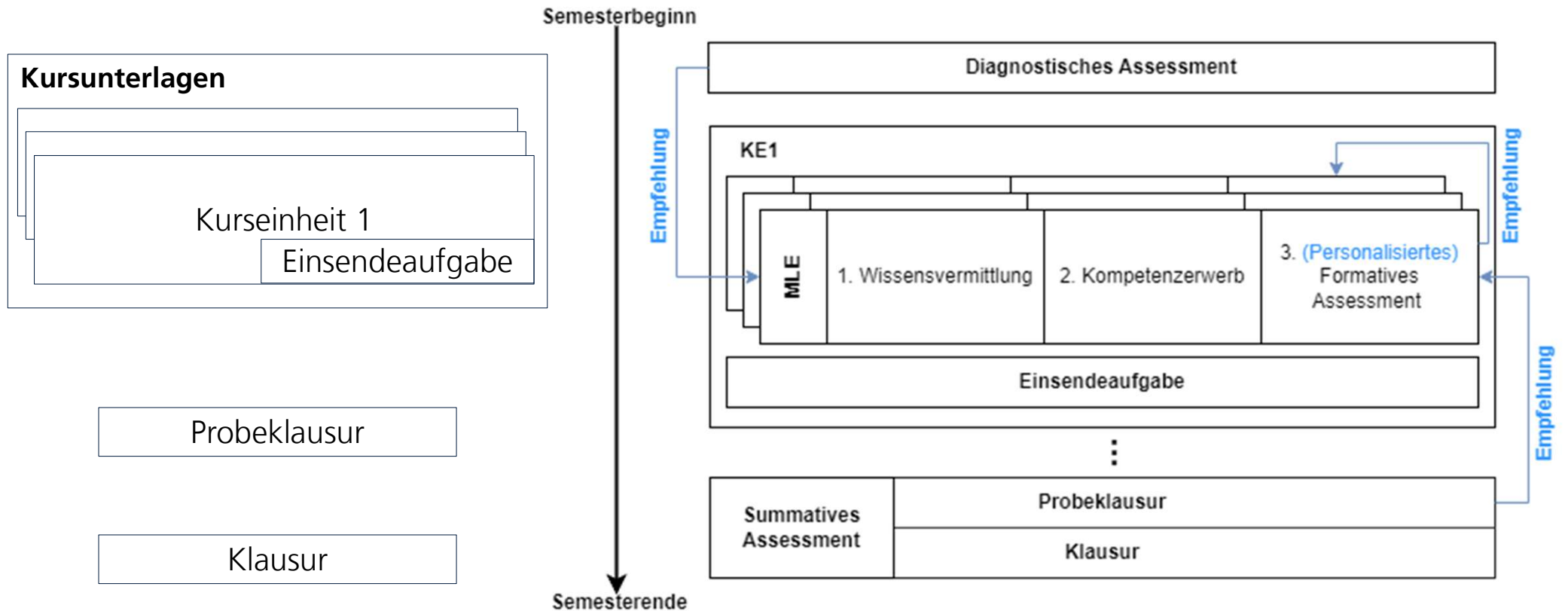


2. Aufgaben, Bewertung, Feedback, Personalisierung





2. Aufgaben, Bewertung, Feedback, Personalisierung



Lehr- und Forschungsprojekt METALADIN

<https://www.fernuni-hagen.de/forschung/schwerpunkte/catalpa/forschung/projekte/METALADIN.shtml>



2. Klassifikation von Kompetenzen und Aufgabentypen

Christ, Paul & Munkelt, Torsten & Schulz, Tim. (2025). Simulation der dynamischen Anpassung des Schwierigkeitsgrades von Übungsaufgaben zu formalen Sprachen an die Kompetenzniveaus von Lernenden. 10.13140/RG.2.2.36650.27845.

		Kognitive Dimension					
		Erinnern		Verstehen	Anwenden	Analysieren	Erzeugen
Inhaltliche Dimension	Sprachverständnis und -interpretation	Theoretische Fragen		Erklären	Lösung eines Problems aufgrund eines Sprachkonstrukts	Überprüfen formaler Eigenschaften	
		Wissen	Verständnis			Finden von Fehlern	
						Vergleich	
						Eignung	
	Sprachmanipulation				Modifikation	Evaluation eines Konstruktionsprozess	Selbständige Konstruktion
					Geführte Konstruktion		

Einfach

Automatisierung

Schwer

Beispiele im Bereich Datenbanken:

- SQL
- ER-Modellierung
- Relationale Algebra
- ...



3. Anwendungsfall 1: Aufgabengenerierung

Automatische Generierung von Aufgaben für Übungen oder Prüfungen

Für Studis:

- Können beliebig viel Üben
- Idealerweise *personalisierte* Übungen oder zumindest große Auswahl an Übungen
- Keine Abhängigkeit zum Lehrenden
 - Asynchrones Üben
- *(Unmittelbares Feedback)*

Für Lehrende:

- Können beliebig viele Varianten einer Aufgabe erzeugen
 - erhöht Prüfungssicherheit (jeder bekommt eigene Aufgabeninstanz)
 - erhöht Prüfungsvalidität (jedes Jahr neue Aufgabeninstanzen)
- *(Reduzierter/Kein Korrekturaufwand)*



3. Anwendungsfall 1: Aufgabengenerierung

Automatische Generierung von Aufgaben für Übungen oder Prüfungen

Für Studis:

- Können beliebig viel Üben
- Idealerweise *personalisierte* Übungen oder zumindest große Auswahl an Übungen !
- Keine Abhängigkeit zum Lehrenden
 - Asynchrones Üben
- (*Unmittelbares Feedback*)

Für Lehrende:

- Können beliebig viele Varianten einer Aufgabe erzeugen
- ! • erhöht Prüfungssicherheit (jeder bekommt eigene Aufgabeninstanz)
- ! • erhöht Prüfungsvalidität (jedes Jahr neue Aufgabeninstanzen)
- (*Reduzierter/Kein Korrekturaufwand*)

! Voraussetzung:

- ! Erfordert Parametrisierbarkeit des Generators zur gezielten Steuerung der Komplexität/Schwierigkeit einer Aufgabe
(Dazu mehr bei AF3)



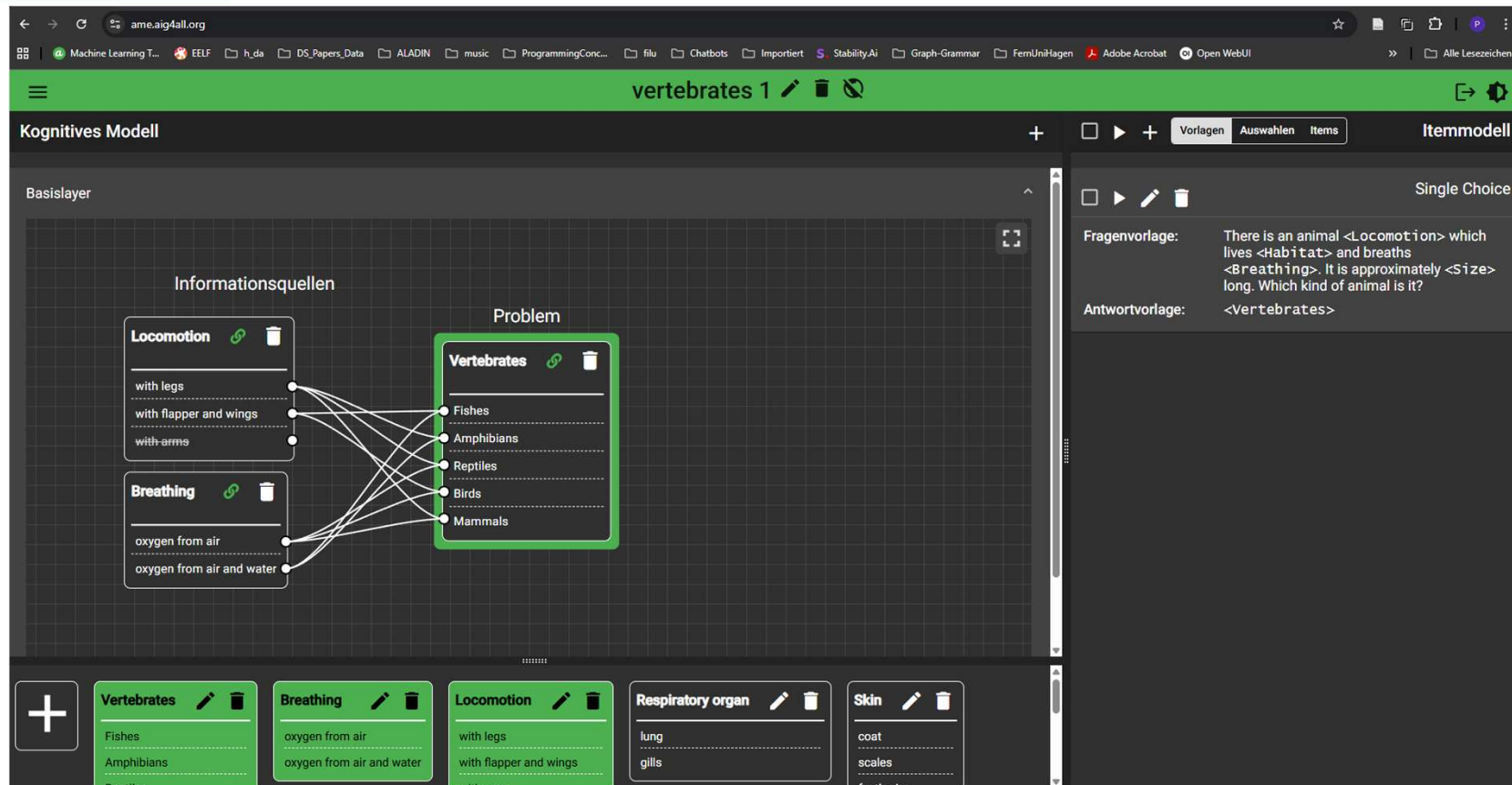
3. AF1: Generische Aufgabengenerierung

Erinnern

Verstehen

Anwenden

<https://ame.aig4all.org/>



Kucharski, S., Damnik, G., Stahr, F. and Braun, I. (2023). Revision of the AIG Software Toolkit: A Contribute to More User Friendliness and Algorithmic Efficiency. In Proceedings of the 15th International Conference on Computer Supported Education - Volume 2: CSEDU; ISBN 978-989-758-641-5; ISSN 2184-5026, SciTePress, pages 410-417. DOI: 10.5220/0011982000003470



3. AF1: Spezifische Aufgabengenerierung (Bsp. SQL)

Anwenden

Analysieren

Erzeugen

Input:

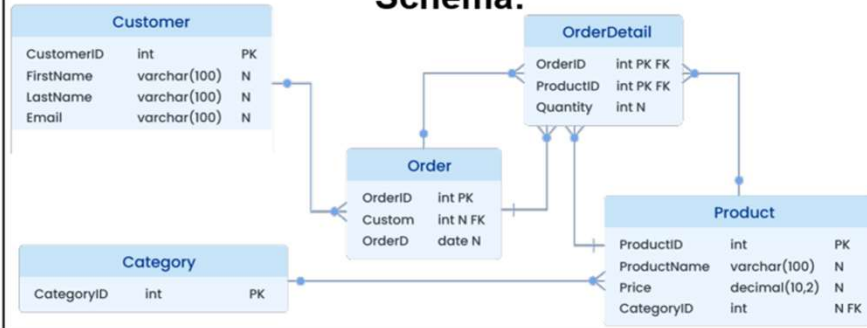
Task Description:

Write a SQL-query that extracts the described data from the database schema shown below.

Problem Statement:

Find the e-mail of all customers that ordered swimming related products between 90 and 100 dollars.

Schema:



Expected Output:

Reference-Query:

```
SELECT Email
FROM Customer AS c
JOIN ORDER o ON c.CustomerId = Custom
JOIN OrderDetail od ON o.OrderId = od.OrderId
JOIN Product p ON od.ProductId = p.ProductId
JOIN Category ca ON p.CategoryId = ca.CategoryId
WHERE ca.CategoryId = 2 AND
p.Price BETWEEN 90 AND 100;
```

Expected Result Set:

Email
Gregory.D@gmail.com
I.Brown@yahoo.com
Se.Ku@web.de
Flo.S@gmx.net

Christ, P., Munkelt, T. and Haake, J. M. (2025). Generating SQL-Query-Items Using Knowledge Graphs. In Proceedings of the 17th International Conference on Computer Supported Education - Volume 1: AIG; ISBN 978-989-758-746-7; ISSN 2184-5026, SciTePress, pages 785-796. DOI: 10.5220/0013498300003932



3. AF1: Alternative zur Generierung - Aufgabensammlungen

Arbeitskreis E-Assessment der Fachgruppe Bildungstechnologien der GI – Ansprechpartner: Prof. Dr. Michael Striewe (FH Trier)



MENÜ MITGLIED WERDEN MEINE GI

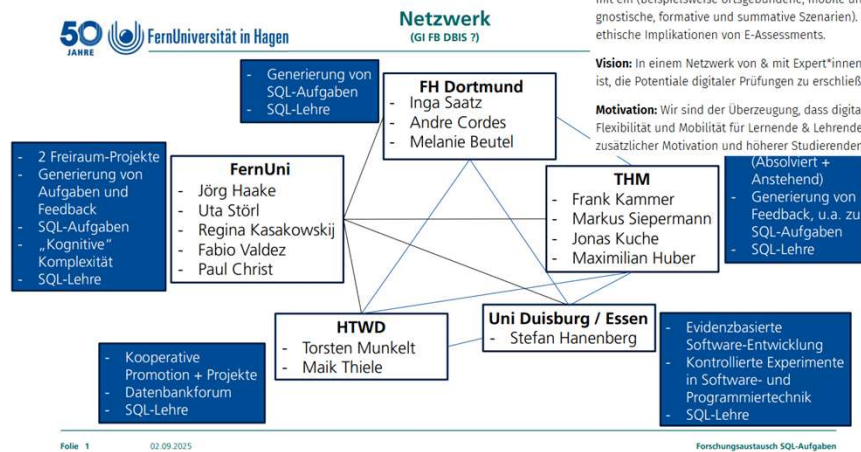
FG Bildungstechnologien > Fachgruppe > Arbeitskreise > AK E-Assessment

Arbeitskreis E-Assessment

Der Arbeitskreis "E-Assessment" befasst sich mit elektronischen Prüfungen im Hochschulkontext. Das schließt das gesamte Spektrum von Prüfungsformaten und -szenarien mit ein (beispielsweise ortsgebundene, mobile und Fernprüfungslösungen, zentrale und dezentrale Infrastrukturen, wissensbasierte und kompetenzorientierte Formate, diagnostische, formative und summative Szenarien). Thematische Schwerpunkte bilden technische Infrastrukturen, didaktische Szenarien, rechtliche Rahmenbedingungen und ethische Implikationen von E-Assessments.

Vision: In einem Netzwerk von & mit Expert*innen und Stakeholdern wollen wir für die Chancen und Herausforderungen neuer Prüfungsformen sensibilisieren. Unser Ziel ist, die Potentiale digitaler Prüfungen zu erschließen und im Sinne der Studierenden und einer bedarfsorientierten Prüfungskultur umzusetzen.

Motivation: Wir sind der Überzeugung, dass digitale Prüfungsformen neue Möglichkeiten für mehr Kompetenzorientierung, größere Lernendenzentrierung und eine höhere Flexibilität und Mobilität für Lernende & Lehrende eröffnen. Größere Freiräume für individuelle Beratung und qualitatives Feedback können zu mehr Chancengerechtigkeit, zusätzlicher Motivation und höherer Studierendenzufriedenheit beitragen.



Größe	Zuletzt geändert	Lizenz	Aktionen
	am 13.04.2023 um 23:40 Uhr		
	am 02.10.2022 um 11:46 Uhr		
	am 14.04.2023 um 00:08 Uhr		
	am 18.05.2023 um 15:52 Uhr		
	am 02.10.2022 um 11:46 Uhr		
	am 02.10.2022 um 11:46 Uhr		
	am 01.04.2023 um 11:57 Uhr		
	am 18.05.2023 um 15:26 Uhr		

Arbeitskreis DB-/SQL-Austausch – Ansprechpartner: Paul Christ (FeU)

Netzwerk Informatik – Ansprechpartner: Dr. Heiko Baum (WHZ)



3. Anwendungsfall 2: Automatische Korrektur/Bewertung/Feedback

Automatische Korrektur und Bewertung des Lösungsversuchs

Für Studis:

- Unmittelbare Rückmeldung über Korrektheit des Lösungsversuchs
 - und ggf. Feedback zu Fehlern

Für Lehrende:

- Reduzierter/Kein Korrekturaufwand
 - Kostenreduktion*

Voraussetzungen:

- Maschinenlesbare Form des Lösungsartefakts, bzw. des Lösungsversuchs
- Ortsunabhängiger Zugriff auf alle erforderlichen Übungsmaterialien für Studis

* Kosten für Korrekturkräfte an der FeU liegen jährlich derzeit noch im 7-stelligen Bereich



3. Anwendungsfall 2: Automatische Korrektur/Bewertung/Feedback

Automatische Korrektur und Bewertung des Lösungsversuchs

Für Studis:

- Unmittelbare Rückmeldung über Korrektheit des Lösungsversuchs
 - und ggf. Feedback zu Fehlern

Für Lehrende:

- Reduzierter/Kein Korrekturaufwand
 - Kostenreduktion*

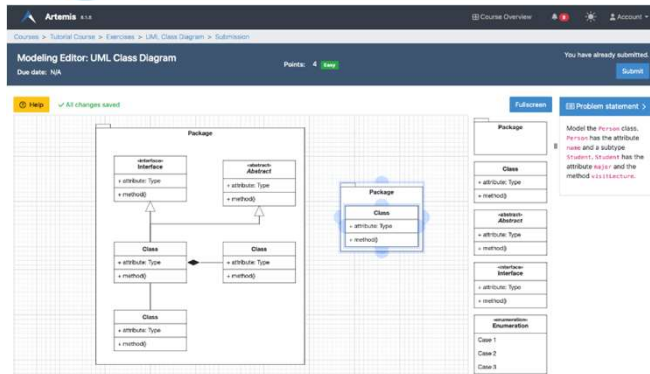
Voraussetzungen:

- Maschinenlesbare Form des Lösungsartefakts, bzw. des Lösungsversuchs
- Ortsunabhängiger Zugriff auf alle erforderlichen Übungsmaterialien für Studis

=> **Schnittmenge:**
Online-Übungs- und
Prüfungsumgebung

* Kosten für Korrekturkräfte an der FeU liegen jährlich derzeit noch im 7-stelligen Bereich

3. AF2: Bsp. für generische Umgebungen

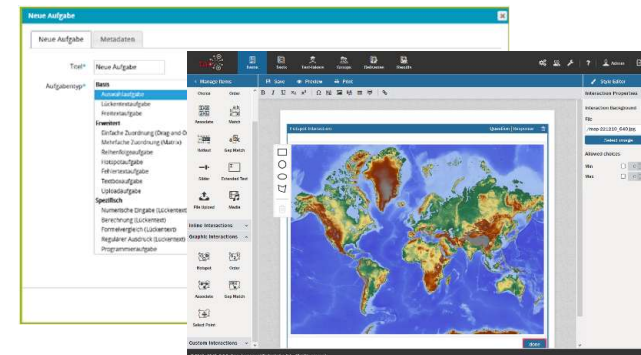


Stephan Krusche and Andreas Seitz. 2018. ArTEMiS: An Automatic Assessment Management System for Interactive Learning. In Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE '18). <https://doi.org/10.1145/3159450.3159602>

Striewe M. (2025). Automatic Item Generation Integrated into the E-Assessment-System JACK. In Proceedings of the 17th International Conference on Computer Supported Education - Volume 1: AIG; ISBN 978-989-758-746-7, SciTePress, pages 747-753. DOI: 10.5220/0013454600003932

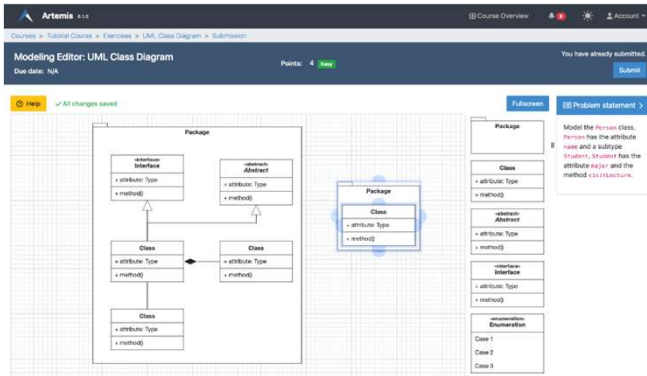
Name	Gebiete	Fat (g)	Carbs (g)	Protein (g)	Sodium (mg)	Calcium (%)
Frozen Yogurt	138	6	24	4	87	14%
Ice cream sandwich	227	9	37	4.3	129	8%
Escher	262	16	23	6	337	8%
Capicola	305	3.7	67	4.3	413	2%

Christ, Paul; Laue, Ralf; Munkelt, Torsten (2022): ALADIN – Generator für Aufgaben und Lösungshilfen aus der Informatik und angrenzenden Disziplinen. Modellierung 2022 Satellite Events. DOI: 10.18420/modellierung2022ws-028.



OPAL/ONYX, Moodle, ILIAS und co.

3. AF2: Bsp. für generische Umgebungen

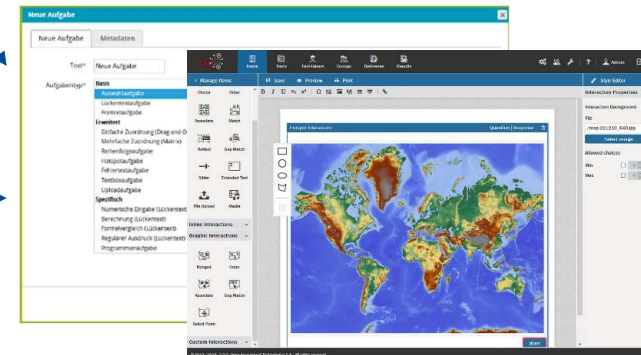


Stephan Krusche and Andreas Seitz. 2018. ArTEMiS: An Automatic Assessment Management System for Interactive Learning. In Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE '18). <https://doi.org/10.1145/3159450.3159602>

Striewe M. (2025). Automatic Item Generation Integrated into the E-Assessment-System JACK. In Proceedings of the 17th International Conference on Computer Supported Education - Volume 1: AIG; ISBN 978-989-758-746-7, SciTePress, pages 747-753. DOI: 10.5220/0013454600003932

Christ, Paul; Laue, Ralf; Munkelt, Torsten (2022): ALADIN – Generator für Aufgaben und Lösung(shif)fen aus der Informatik und angrenzenden Disziplinen. Modellierung 2022 Satellite Events. DOI: 10.18420/modellierung2022ws-028.

LTI-Schnittstelle



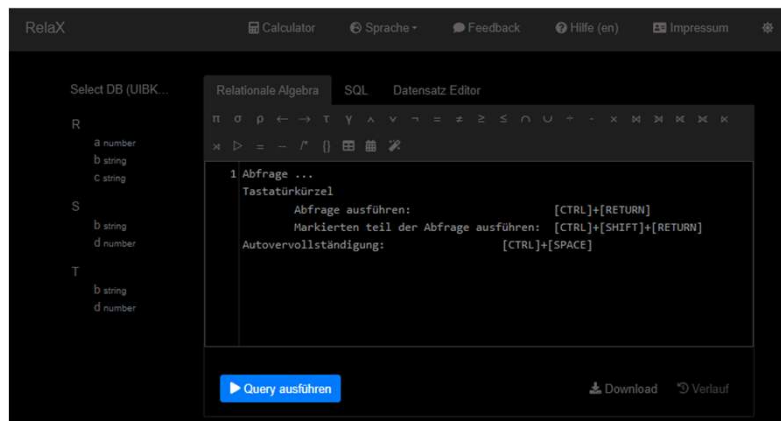
OPAL/ONYX, Moodle, ILIAS und co.



3. AF2: Bsp. für spezifische Übungsumgebungen

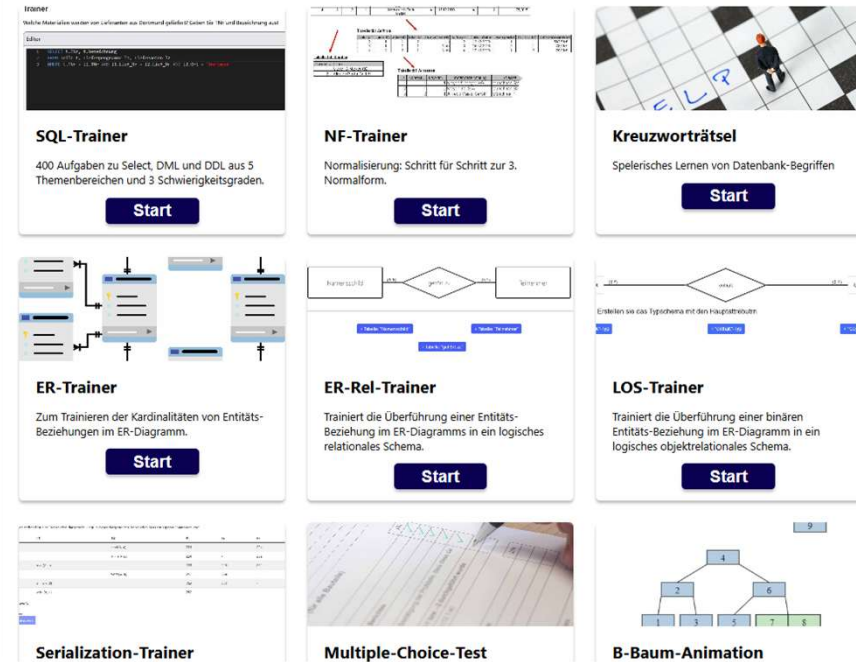


SQL-Playground im Frits-Feedbacksystem. Siehe: Reid, S. A., Kammer, F., Kunz, J., Pellekooor, T., Siepermann, M., & Wölfer, J. (2023). ItsSQL: Intelligent Tutoring System for SQL. arXiv [Cs.CY]. Retrieved from <http://arxiv.org/abs/2311.10730>



Johannes Kessler, Michael Tschuggnall and Günther Specht: RelaX: A Webbased Execution and Learning Tool for Relational Algebra, In: Proceedings of the Datenbanksysteme für Business, Technologie und Web (BTW), 2019

EILD.nrw Trainer

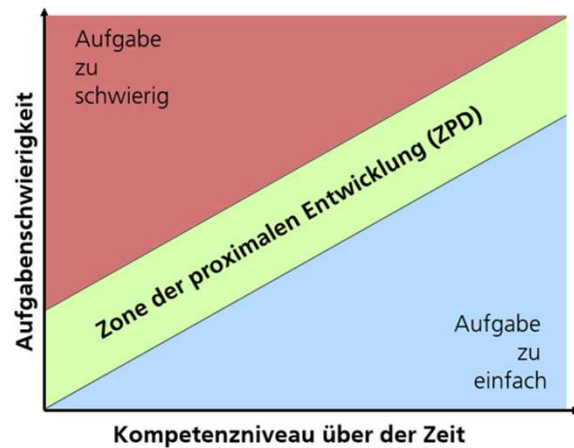


Rakow, Thomas C.; Kless, André; Hasler, Charlotte; Knolle, Harm; Faeskorn-Woyke, Heide; Saat, Inga Marina; Lambert, Jens; Focken, Mareike. Developing OERs for Teaching Database Systems – A Two-Year Effort of Four Universities of Applied Sciences. In: König-Ries, B.; Scherzinger, S.; Lehner, W.; Vossen, G. (Hrsg.), 20th Conference Database Systems for Business, Technology and Web (BTW 2023), pp. 621-631. Gesellschaft für Informatik e.V.: 2023



3. Anwendungsfall 3: Personalisierung (Aufgaben, Feedback, Lernpfade)

Kurs- und Lernmaterialien werden individuell auf die Bedürfnisse von Lernenden zugeschnitten

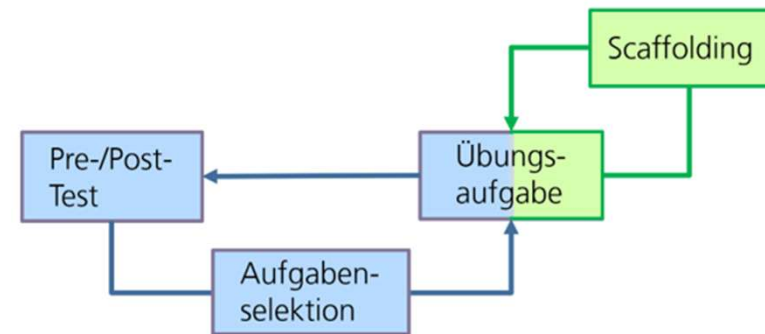


Für Studis:

- Progression im eigenen Tempo
- !Theoretisch!: Konvergenz gegen Optimum

Voraussetzungen:

- Erfüllung aller vorherigen AF und deren Voraussetzungen



Für Lehrende:

- Reduzierter Betreuungsaufwand



3. AF3: Voraussetzung – Steuerbare Komplexität/Schwierigkeit von Aufgaben

Beispiel: SQL-Aufgaben

Input:

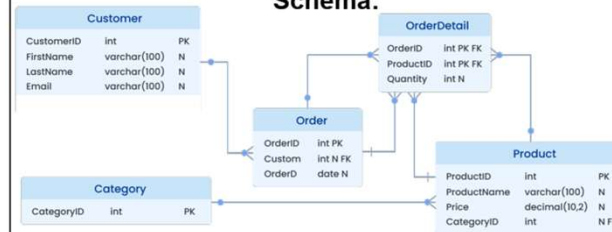
Task Description:

Write a SQL-query that extracts the described data from the database schema shown below.

Problem Statement:

Find the e-mail of all customers that ordered swimming related products between 90 and 100 dollars.

Schema:



Expected Output:

Reference-Query:

```
SELECT Email
FROM Customer AS c
JOIN ORDER o ON c.CustomerId = Custom
JOIN OrderDetail od ON o.OrderId = od.OrderId
JOIN Product p ON od.ProductId = p.ProductId
JOIN Category ca ON p.CategoryId = ca.CategoryId
WHERE ca.CategoryId = 2 AND
p.Price BETWEEN 90 AND 100;
```

Expected Result Set:

Email
Gregory.D@gmail.com
I.Brown@yahoo.com
Se.Ku@web.de
Flo.S@gmx.net



3. AF3: Voraussetzung – Steuerbare Komplexität/Schwierigkeit von Aufgaben

Beispiel: SQL-Aufgaben

Input:

Task Description:
Write a SQL-query that extracts the described data from the database schema shown below.

Problem Statement:
Find the e-mail of all customers that ordered swimming related products between 90 and 100 dollars.

Schema:

Expected Output:

Reference-Query:

```
SELECT Email
FROM Customer AS c
JOIN ORDER o ON c.CustomerId = Custom
JOIN OrderDetail od ON o.OrderId = od.OrderId
JOIN Product p ON od.ProductId = p.ProductId
JOIN Category ca ON p.CategoryId = ca.CategoryId
WHERE ca.CategoryId = 2 AND
p.Price BETWEEN 90 AND 100;
```

Expected Result Set:

Email
Gregory.D@gmail.com
I.Brown@yahoo.com
Se.Ku@web.de
Flo.S@gmx.net

Mögliche Komplexitätsdimensionen

1. Verständlichkeit der Frage
2. Benötigtes Wissen zur Projektion der Frage auf SQL
3. Komplexität des Datenbankschemas
4. Komplexität der Lösungs-SQL-Query
(bzw. dem Raum aller Lösungs-SQL-Queries)

Li, J., Hui, B., Qu, G., Yang, J., Li, B., Li, B., ... Others. (2024). Can Llm already serve as a database interface? a big bench for large-scale database grounded text-to-sqls. Advances in Neural Information Processing Systems, 36.



3. AF3: Voraussetzung – Steuerbare Komplexität/Schwierigkeit von Aufgaben

Beispiel: SQL-Aufgaben

Input:

Task Description:
Write a SQL-query that extracts the described data from the database schema shown below.

Problem Statement:
Find the e-mail of all customers that ordered swimming related products between 90 and 100 dollars.

Schema:

Expected Output:

Reference-Query:

```
SELECT Email
FROM Customer AS c
JOIN ORDER o ON c.CustomerId = Custom
JOIN OrderDetail od ON o.OrderId = od.OrderId
JOIN Product p ON od.ProductId = p.ProductId
JOIN Category ca ON p.CategoryId = ca.CategoryId
WHERE ca.CategoryId = 2 AND
p.Price BETWEEN 90 AND 100;
```

Expected Result Set:

Email
Gregory.D@gmail.com
I.Brown@yahoo.com
Se.Ku@web.de
Flo.S@gmx.net

1.
2.
3.
4.

Mögliche Komplexitätsdimensionen

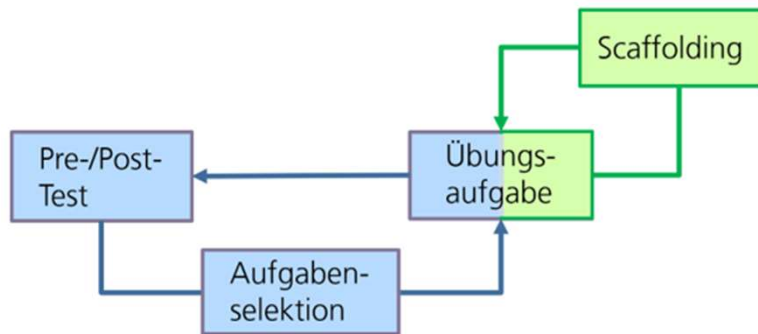
1. Verständlichkeit der Frage
2. Benötigtes Wissen zur Projektion der Frage auf SQL → T. Taipalus, "Explaining Causes Behind SQL Query Formulation Errors," 2020 IEEE Frontiers in Education Conference (FIE), Uppsala, Sweden, 2020, pp. 1-9, doi: 10.1109/FIE44824.2020.9274114.
3. Komplexität des Datenbankschemas → Taipalus, T. (2020). The effects of database complexity on SQL query formulation. Journal of Systems and Software, 165, 110576. doi:10.1016/j.jss.2020.110576
4. Komplexität der Lösungs-SQL-Query (bzw. dem Raum aller Lösungs-SQL-Queries) → Taipalus, T., Siponen, M., & Vartiainen, T. (2018). Errors and Complications in SQL Query Formulation. ACM Trans. Comput. Educ., 18(3). doi:10.1145/3231712

Li, J., Hui, B., Qu, G., Yang, J., Li, B., Li, B., ... Others. (2024). Can Llm already serve as a database interface? a big bench for large-scale database grounded text-to-sqls. Advances in Neural Information Processing Systems, 36.



3. AF3: Voraussetzung – Steuerbare Komplexität/Schwierigkeit von Aufgaben

Beispiel: SQL-Aufgaben



Mögliche Komplexitätsdimensionen

1. Verständlichkeit der Frage
2. Benötigtes Wissen zur Projektion der Frage auf SQL
3. Komplexität des Datenbankschemas
4. Komplexität der Lösungs-SQL-Query (bzw. dem Raum aller Lösungs-SQL-Queries)

Li, J., Hui, B., Qu, G., Yang, J., Li, B., Li, B., ... Others. (2024). Can Llm already serve as a database interface? a big bench for large-scale database grounded text-to-sqls. *Advances in Neural Information Processing Systems*, 36.

1.

3.

2.

4.

Input:

Task Description:
Write a SQL-query that extracts the described data from the database schema shown below.

Problem Statement:
Find the e-mail of all customers that ordered swimming related products between 90 and 100 dollars.

Schema:

Customer
CustomerID int PK
FirstName varchar(100) N
LastName varchar(100) N
Email varchar(100) N

Order
OrderID int PK
Custom int N FK
OrderD date N

OrderDetail
OrderID int PK FK
ProductID int PK FK
Quantity int N

Product
ProductID int PK
ProductName varchar(100) N
Price decimal(10,2) N
CategoryID int N FK

Category
CategoryID int PK

Expected Output:

Reference-Query:

```
SELECT Email
FROM Customer AS c
JOIN ORDER o ON c.CustomerId = Custom
JOIN OrderDetail od ON o.OrderId = od.OrderId
JOIN Product p ON od.ProductId = p.ProductId
JOIN Category ca ON p.CategoryId = ca.CategoryId
WHERE ca.CategoryId = 2 AND
p.Price BETWEEN 90 AND 100;
```

Expected Result Set:

Email
Gregory.D@gmail.com
I.Brown@yahoo.com
Se.Ku@web.de
Flo.S@gmx.net

T. Taipalus, "Explaining Causes Behind SQL Query Formulation Errors," 2020 IEEE Frontiers in Education Conference (FIE), Uppsala, Sweden, 2020, pp. 1-9, doi: 10.1109/FIE44824.2020.9274114.

Taipalus, T. (2020). The effects of database complexity on SQL query formulation. *Journal of Systems and Software*, 165, 110576. doi:10.1016/j.jss.2020.110576

Taipalus, T., Siponen, M., & Vartiainen, T. (2018). Errors and Complications in SQL Query Formulation. *ACM Trans. Comput. Educ.*, 18(3). doi:10.1145/3231712



4. (Zwischen-)Fazit

Aktueller Stand hinsichtlich der Automatisierungspotenziale in der DB-Lehre:

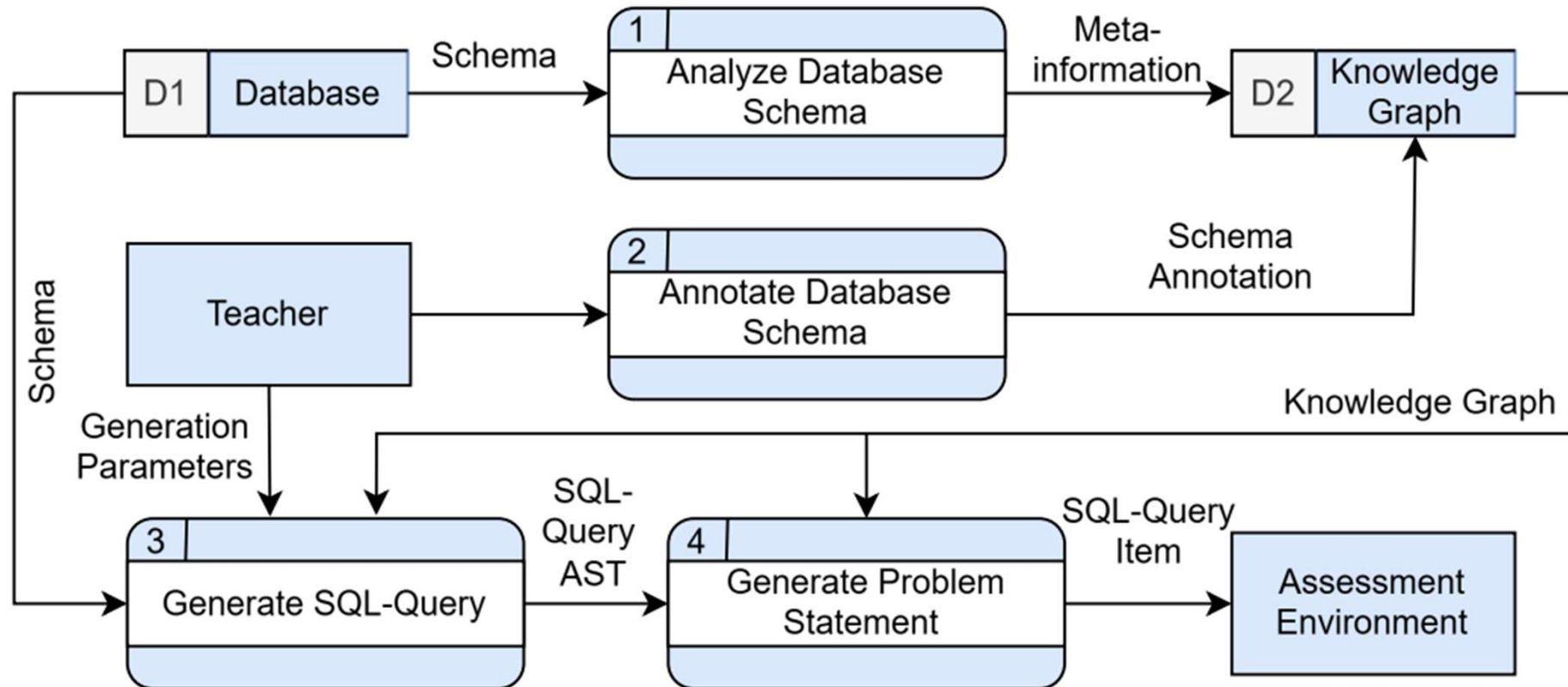
- Aktive Community mit vielen OER-Angeboten
- Existierende Tools können unmittelbar genutzt werden

Limitationen:

- Fehlende empirische Evaluation hinsichtlich der Auswirkungen auf die Lerneffekte
- Keine vollständige Abdeckung zu allen in der DB-Lehre zu vermittelnden Kompetenzen
 - Automatisierung ist initial aufwandsintensiv
=> „Amortisationszeit“ zur Erschließung fehlender Bereiche ist ggf. sehr lang



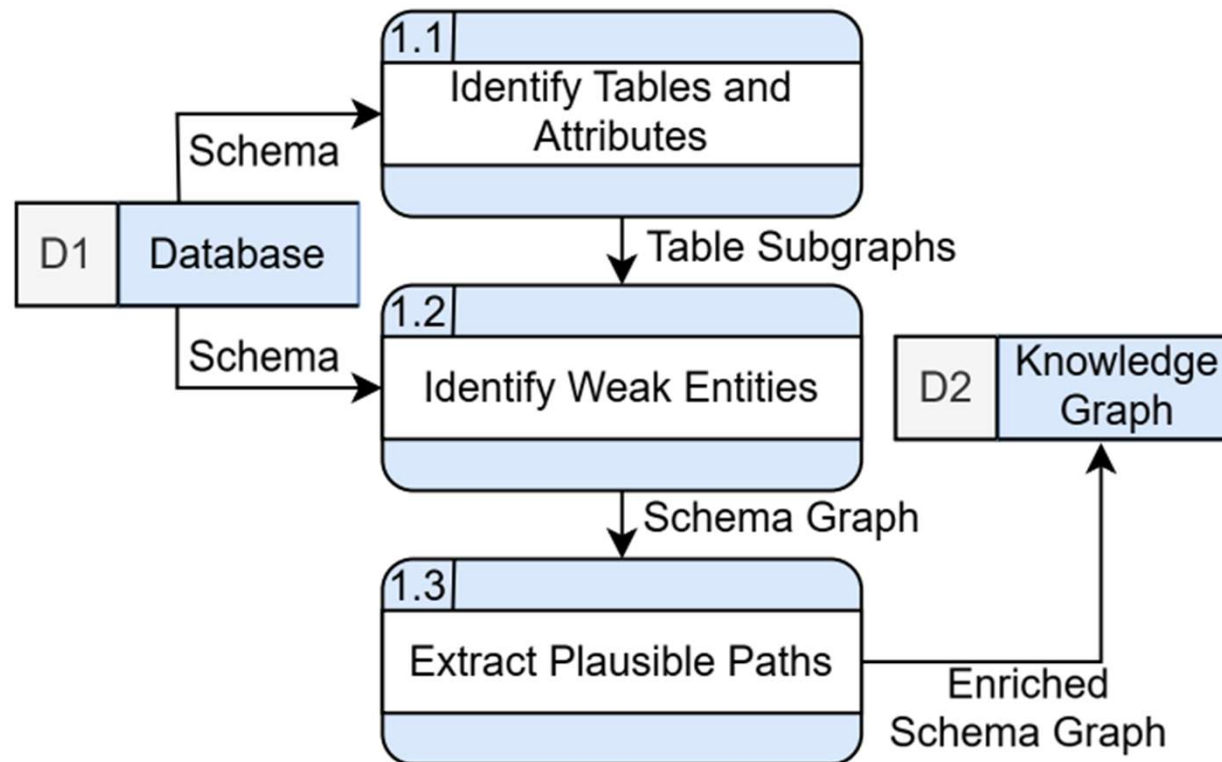
5. (Deep) Dive - 1: SQL-Aufgaben-Generierung - Überblick



Christ, P., Munkelt, T. and Haake, J. M. (2025). Generating SQL-Query-Items Using Knowledge Graphs. In Proceedings of the 17th International Conference on Computer Supported Education - Volume 1: AIG; ISBN 978-989-758-746-7; ISSN 2184-5026, SciTePress, pages 785-796. DOI: 10.5220/0013498300003932



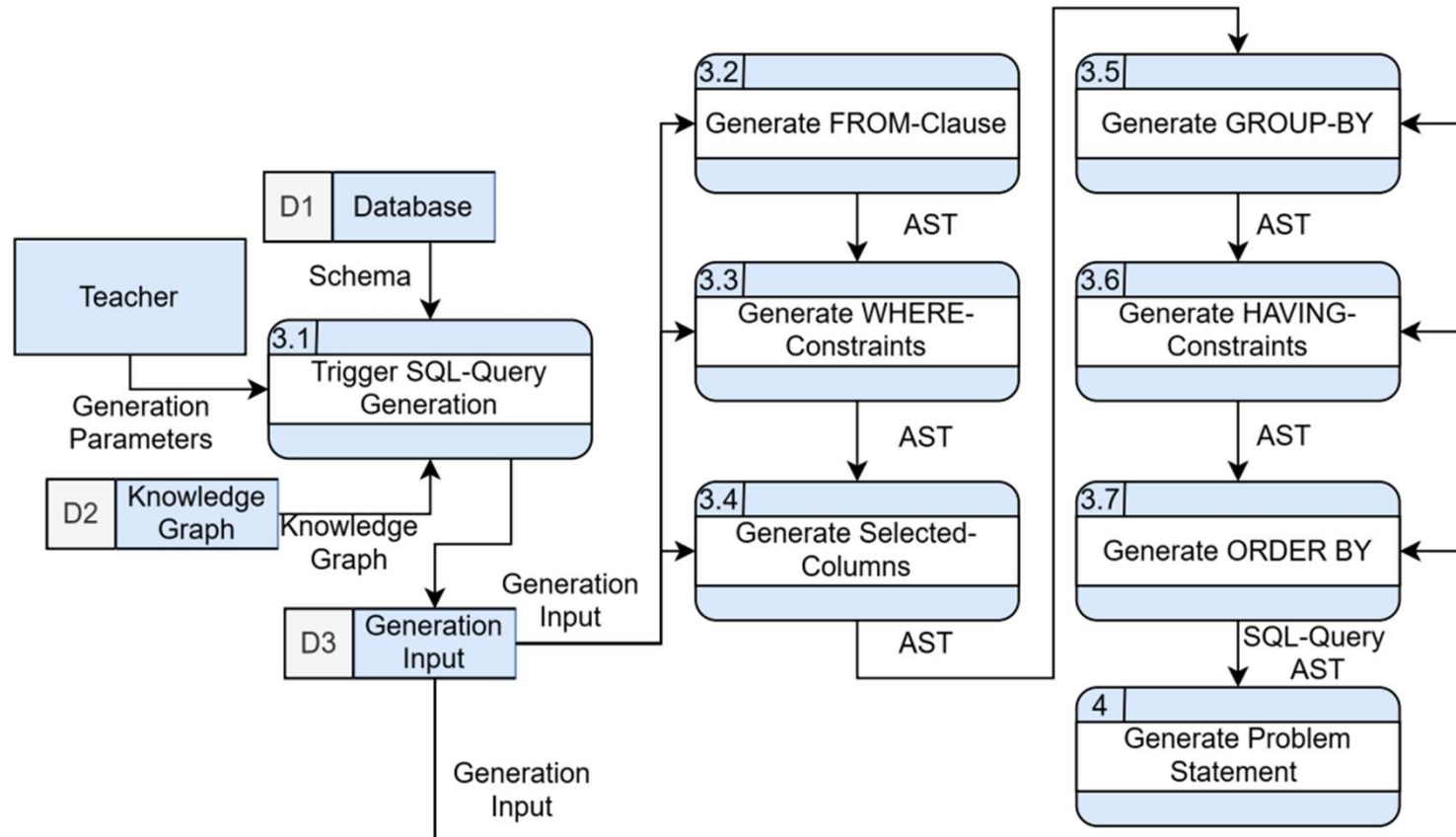
5. (Deep) Dive - 1: SQL-Aufgaben-Generierung – DB-Schema-Analyse



Christ, P., Munkelt, T. and Haake, J. M. (2025). Generating SQL-Query-Items Using Knowledge Graphs. In Proceedings of the 17th International Conference on Computer Supported Education - Volume 1: AIG; ISBN 978-989-758-746-7; ISSN 2184-5026, SciTePress, pages 785-796. DOI: 10.5220/0013498300003932



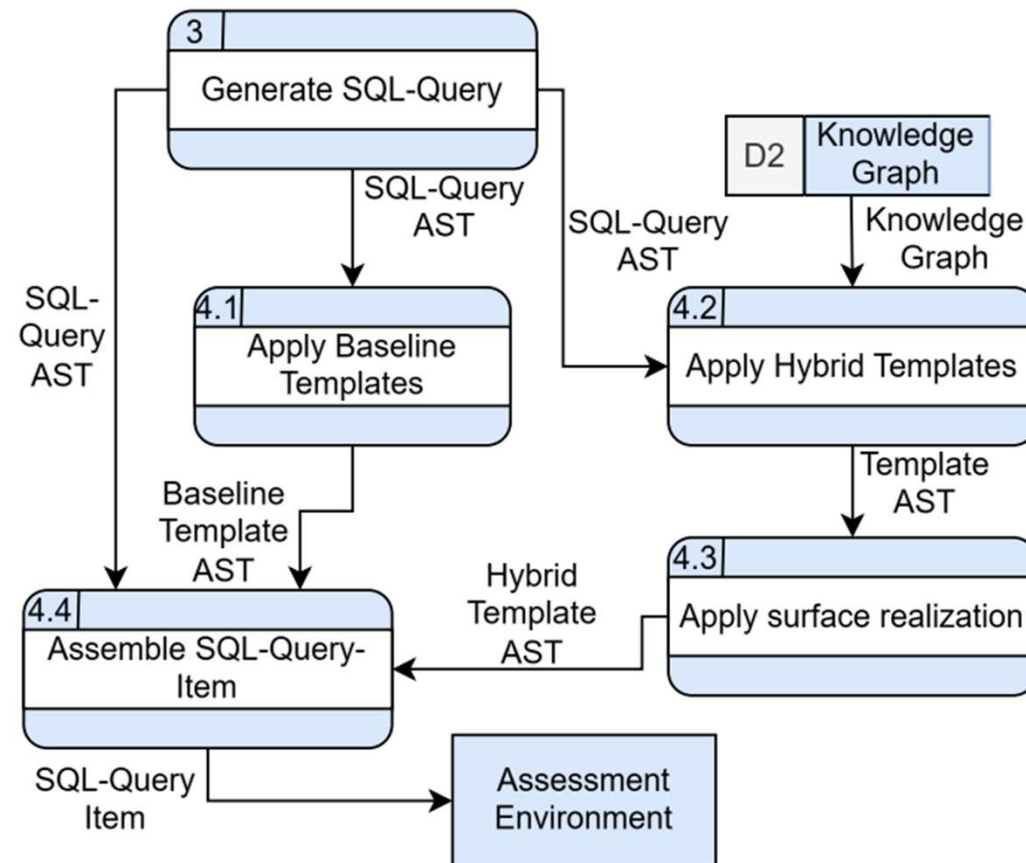
5. (Deep) Dive - 1: SQL-Aufgaben-Generierung – Query-Generierung



Christ, P., Munkelt, T. and Haake, J. M. (2025). Generating SQL-Query-Items Using Knowledge Graphs. In Proceedings of the 17th International Conference on Computer Supported Education - Volume 1: AIG; ISBN 978-989-758-746-7; ISSN 2184-5026, SciTePress, pages 785-796. DOI: 10.5220/0013498300003932



5. (Deep) Dive - 1: SQL-Aufgaben-Generierung – Generierung der Problemstellung



Christ, P., Munkelt, T. and Haake, J. M. (2025). Generating SQL-Query-Items Using Knowledge Graphs. In Proceedings of the 17th International Conference on Computer Supported Education - Volume 1: AIG; ISBN 978-989-758-746-7; ISSN 2184-5026, SciTePress, pages 785-796. DOI: 10.5220/0013498300003932



5. (Deep) Dive - 1: SQL-Aufgaben-Generierung – Evaluierung

Questions:

Q1: „How plausible is it for a human to request the information expressed by the shown SQL-query for this database?“

Q2: „How frequently would you assume that the information expressed by the shown SQL-query to be requested for this database?“

Q3: „How complex would you rate this query?“

Q4: „Which text reads more natural (1 or 2)?“

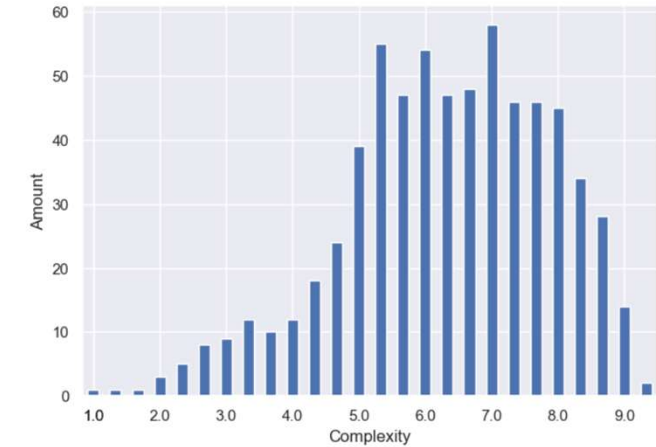
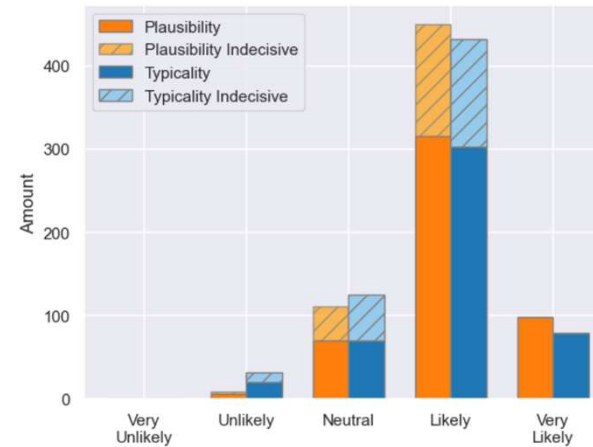


Figure 8: Distribution of plausibility and typicality ratings. Figure 7: Distribution of the average complexity rating.

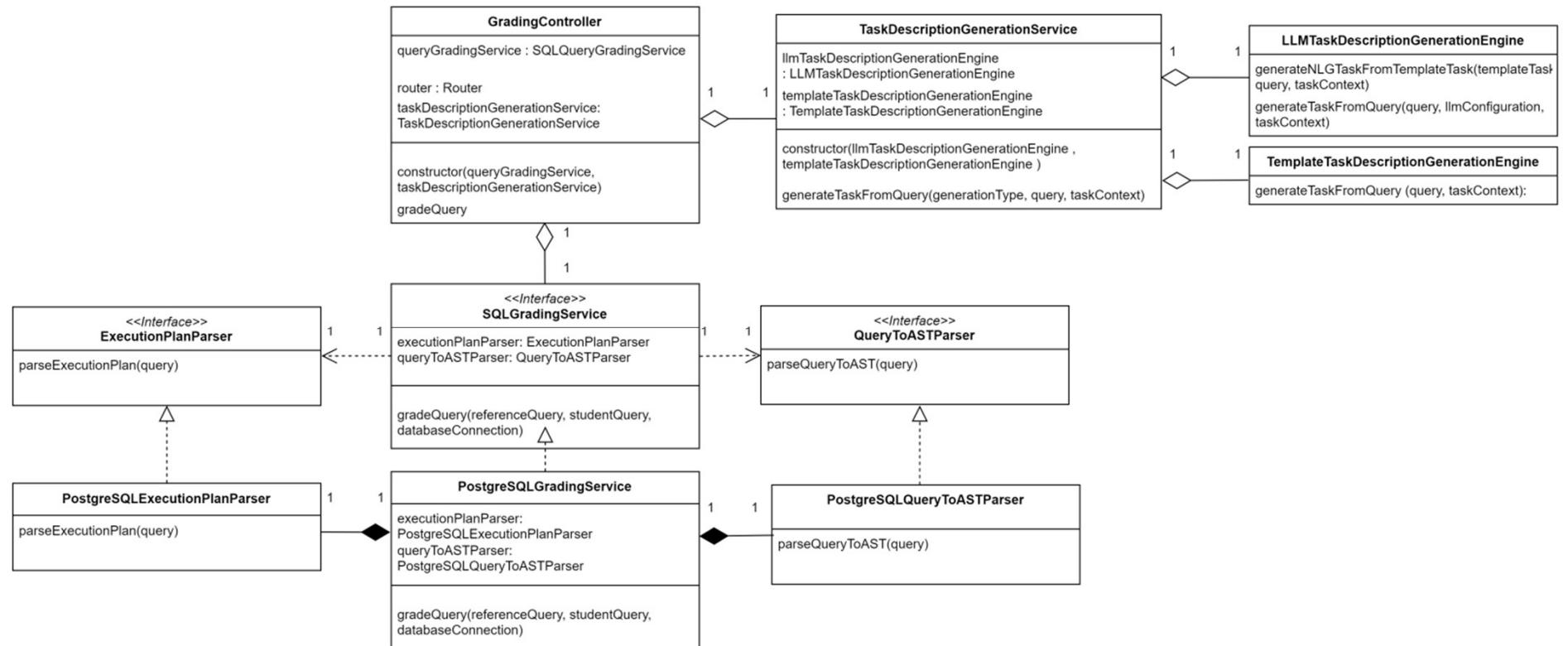
Table 4: Performance scores for the NLPS variants - numbers in bold indicate the preferable NLPS variant according to the score of the corresponding metric.

NLPS	Rater Preference	Ø Characters	Ø Term Frequency	Ø Flesch Reading Ease	Ø Cosine Similarity to Query	Passed Quality Check	Ø Dependency Distance	Ø Coherence
Baseline	289	459	31	45	63%	138	2.4	80%
Hybrid	378	225	26	65	56%	478	3.5	87%

Christ, P., Munkelt, T. and Haake, J. M. (2025). Generating SQL-Query-Items Using Knowledge Graphs. In Proceedings of the 17th International Conference on Computer Supported Education - Volume 1: AIG; ISBN 978-989-758-746-7; ISSN 2184-5026, SciTePress, pages 785-796. DOI: 10.5220/0013498300003932



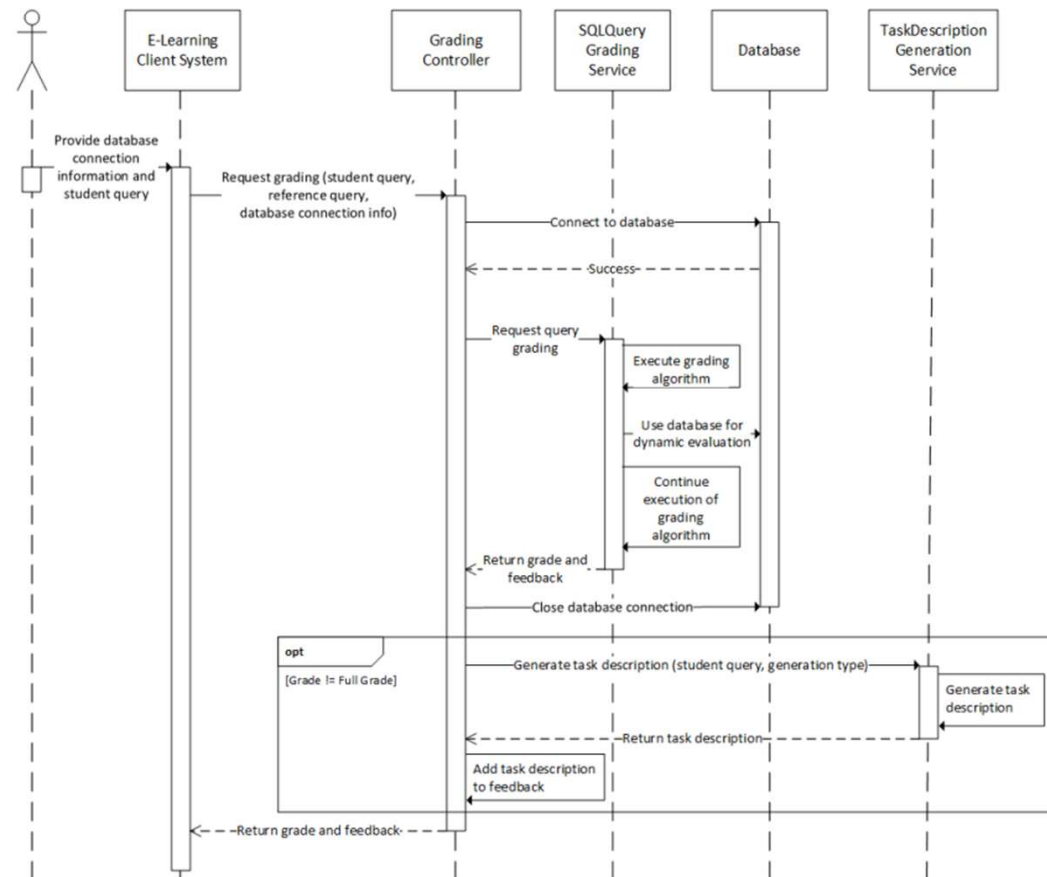
5. (Deep) Dive - 2: SQL-Feedback-Generierung – Component Overview



Jasmin Oulabi. Automatic Generation and Grading of Context-Based SQL Query Task. Master Thesis – Fernuniversität Hagen – 15.04.2025. Supervised by Paul Christ, Jörg Haake, Uta Störl



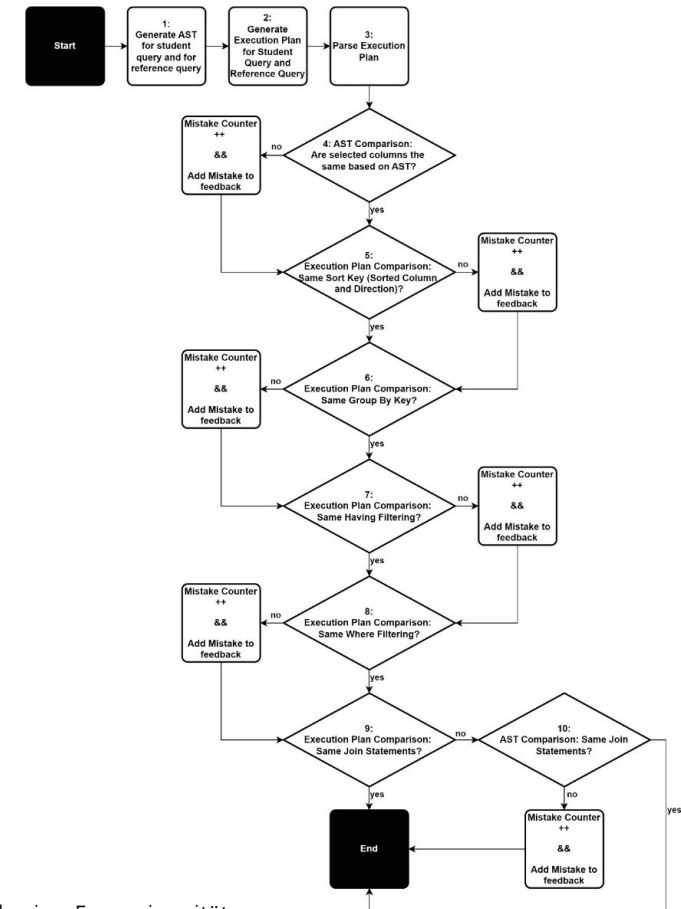
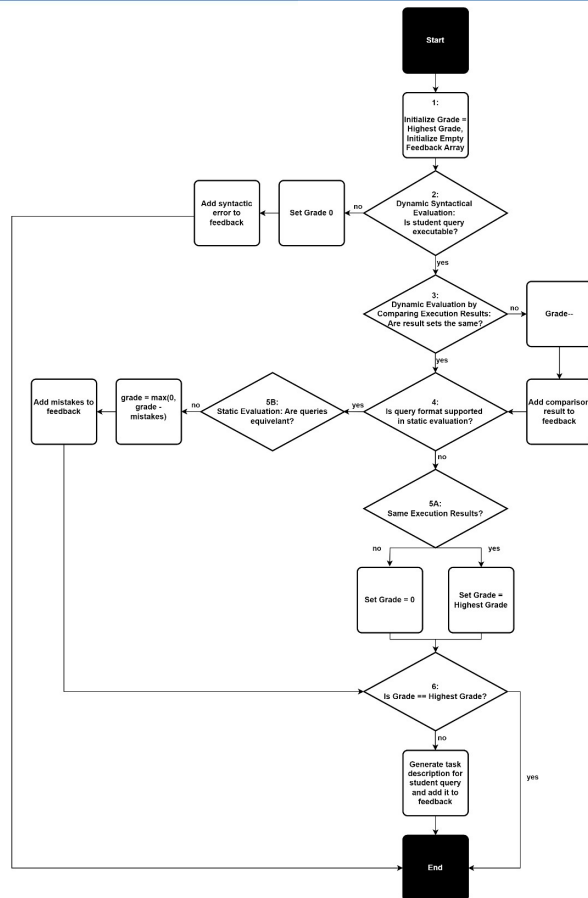
5. (Deep) Dive - 2: SQL-Feedback-Generierung - Overview



Jasmin Oulabi. Automatic Generation and Grading of Context-Based SQL Query Task. Master Thesis – Fernuniversität Hagen – 15.04.2025. Supervised by Paul Christ, Jörg Haake, Uta Störl



5. (Deep) Dive - 2: SQL-Feedback-Generierung – Grading Algorithm & Static Evaluation (AST + Ausführungsplan)



Jasmin Oulabi. Automatic Generation and Grading of Context-Based SQL Query Task. Master Thesis – Fernuniversität Hagen – 15.04.2025. Supervised by Paul Christ, Jörg Haake, Uta Störl

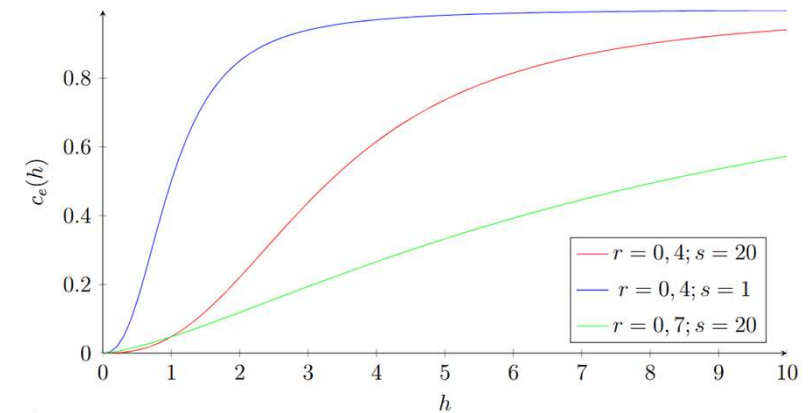


5. (Deep) Dive - 3: Personalisierung von SQL-Aufgaben – Komplexitätsmodell

Herleitung: mit Softwaremetriken, wie z. B. die Halstead-Metrik [4], Schwierigkeit schätzen, ein Programm zu schreiben oder zu verstehen

Annahmen:

- Konstrukte formaler Sprachen bestehen aus Operatoren und Operanden
- Komplexität Konstrukt $\leftarrow$ Gesamtanzahl und Anzahl distinkter Operatoren und Operanden
- Komplexität einer Aufgabe zu formalen Sprachen hinreichend durch das zu verstehende oder zu erstellende Konstrukt approximiert
- Der Komplexitätszuwachs eines Operators oder Operanden mit steigender Auftrittshäufigkeit abnehmend (siehe Diagramm)



Christ, Paul & Munkelt, Torsten & Schulz, Tim. (2025). Simulation der dynamischen Anpassung des Schwierigkeitsgrades von Übungsaufgaben zu formalen Sprachen an die Kompetenzniveaus von Lernenden. 10.13140/RG.2.2.36650.27845.

[4] Halstead, Maurice H. (1977). *Elements of Software Science*. Amsterdam: Elsevier North-Holland, Inc. ISBN 0-444-00205-7.



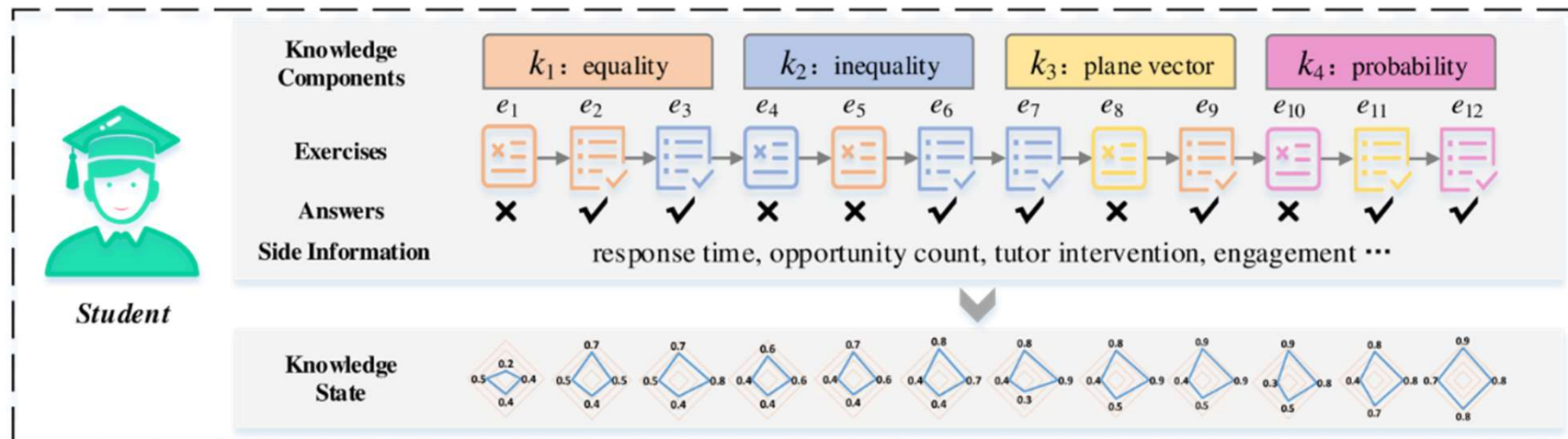
5. (Deep) Dive - 3: Personalisierung von SQL-Aufgaben – Lernermodell

- Zentrale Lernzielkompetenz: Erschaffen eines Konstrukts einer formalen Sprache (siehe z. B. Folie 4)
- in Teilkompetenzen zur Erschaffung von Teilkonstrukten aufteilbar
- „Start-Kompetenzniveau“ aller Lernenden ~normalverteilt [5]
- Ein Lerneffekt nur dann, wenn eine „passende“ Aufgabe bearbeitet worden ist („im ZPD-Korridor“)
- Lerneffekt ist unabhängig davon, ob
 - Aufgabe korrekt oder inkorrekt gelöst wurde
 - (und in welchem Maße) Hilfestellung geleistet worden ist
- Lerneffekt = zur Lösung einer *passenden* Aufgabe benötigtes Kompetenz-niveau - Kompetenzniveau der/s Lernenden vor Bearbeitung der Aufgabe

Christ, Paul & Munkelt, Torsten & Schulz, Tim. (2025). Simulation der dynamischen Anpassung des Schwierigkeitsgrades von Übungsaufgaben zu formalen Sprachen an die Kompetenzniveaus von Lernenden. 10.13140/RG.2.2.36650.27845.

[5] Myers, D. G. (2014). Intelligenz. In D. G. Myers (Hrsg.), *Psychologie* (S. 399–436). Springer. https://doi.org/10.1007/978-3-642-40782-6_11

5. (Deep) Dive - 3: Personalisierung von SQL-Aufgaben – Simulation



Shen, S., Liu, Q., Huang, Z., Zheng, Y., Yin, M., Wang, M., & Chen, E. (2024). A Survey of Knowledge Tracing: Models, Variants, and Applications. *IEEE Trans. Learn. Technol.*, 17, 1898–1919. <https://doi.org/10.1109/TLT.2024.3383325>

Trainingsdaten (70%):

- 15.000 Lernende, maximal 50 Aufgaben
- 6 Kompetenzen/Knowledge Components (JOIN, GROUP BY, WHERE, Subquery, ...)

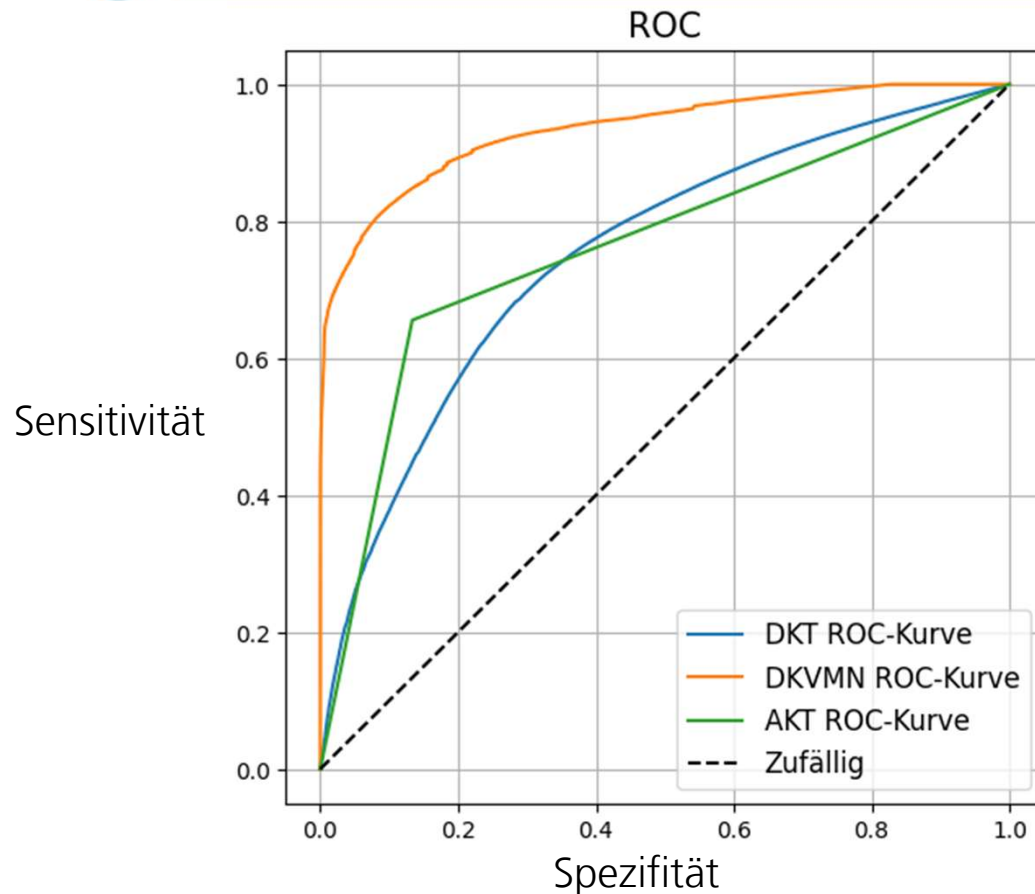
Evaluationsergebnisse (30%):

	DKT	DKVMN	AKT
Accuracy	0,74	0,86	0,79
ROC-AUC	0,75	0,93	0,76

Christ, Paul & Munkelt, Torsten & Schulz, Tim. (2025). Simulation der dynamischen Anpassung des Schwierigkeitsgrades von Übungsaufgaben zu formalen Sprachen an die Kompetenzniveaus von Lernenden. 10.13140/RG.2.2.36650.27845.



5. (Deep) Dive - 3: Personalisierung von SQL-Aufgaben – Konnten die KT-Modelle einen Lerneffekt simulieren?

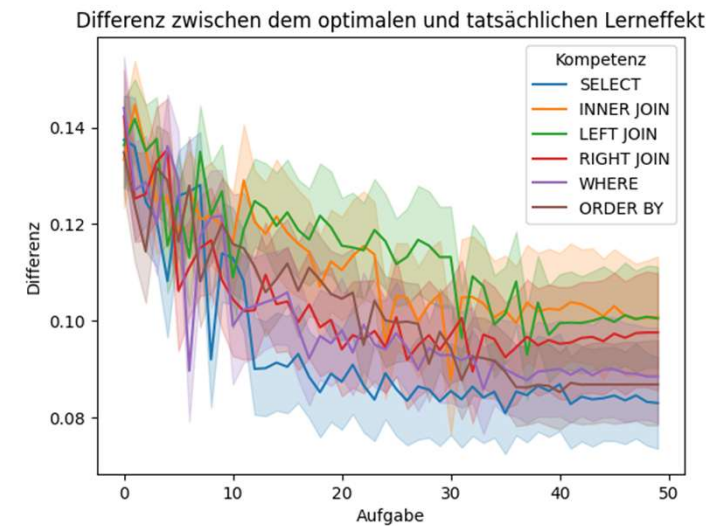
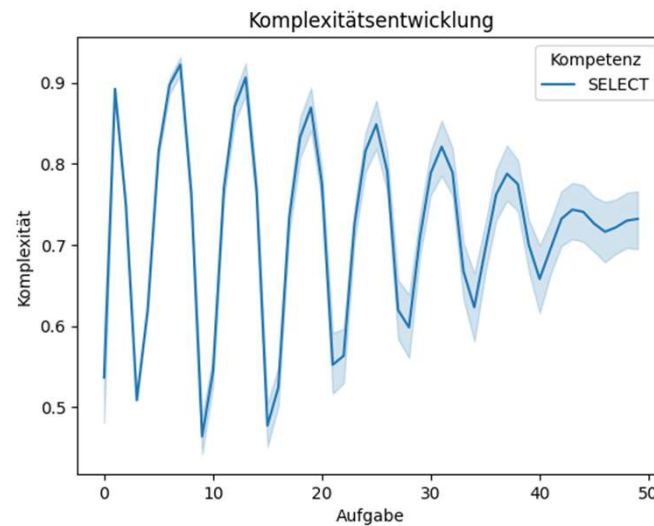
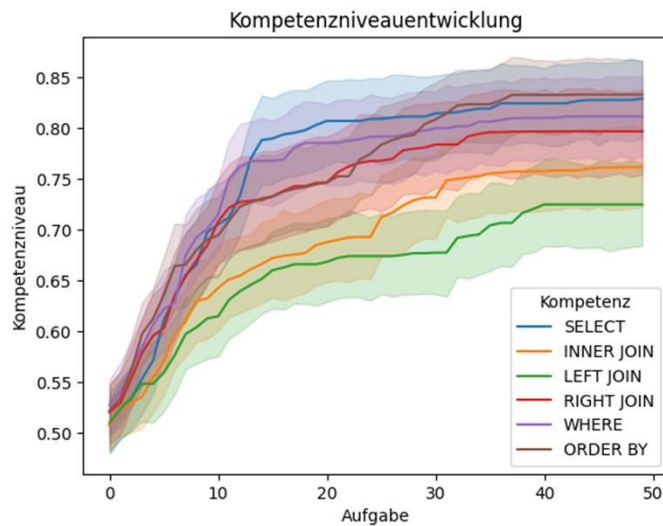


Area under the curve (AUC)	Interpretation
$0.9 \leq \text{AUC}$	Excellent
$0.8 \leq \text{AUC} < 0.9$	Good
$0.7 \leq \text{AUC} < 0.8$	Fair
$0.6 \leq \text{AUC} < 0.7$	Poor
$0.5 \leq \text{AUC} < 0.6$	Fail

Christ, Paul & Munkelt, Torsten & Schulz, Tim. (2025). Simulation der dynamischen Anpassung des Schwierigkeitsgrades von Übungsaufgaben zu formalen Sprachen an die Kompetenzniveaus von Lernenden. 10.13140/RG.2.2.36650.27845.



5. (Deep) Dive - 3: Personalisierung von SQL-Aufgaben – Anwendung eines KT-Modells innerhalb der Simulation



Christ, Paul & Munkelt, Torsten & Schulz, Tim. (2025). Simulation der dynamischen Anpassung des Schwierigkeitsgrades von Übungsaufgaben zu formalen Sprachen an die Kompetenzniveaus von Lernenden. 10.13140/RG.2.2.36650.27845.



FernUniversität
in Hagen



Vielen Dank für die Aufmerksamkeit!

Diskussion & Fragen