

Die Datenfallstudie – eine Lehr-Lern-Methode für den Informatikunterricht

Viktoriya Olari ¹ und Ralf Romeike ¹

Abstract: Themen aus dem Bereich Künstliche Intelligenz (KI) werden in die Informatiklehrpläne aufgenommen. Dabei stellt sich die Frage, welche Methoden geeignet sind, um diese Themen – insbesondere die bislang wenig beachteten Datenkonzepte und -praktiken – im Unterricht zu behandeln. Die zentrale Lehr-Lern-Methode in der akademischen Bildung in KI und Datenwissenschaft ist die Datenfallstudie. Trotz ihrer Verbreitung sind keine Berichte über ihre didaktische Weiterentwicklung für den Einsatz in der Schule bekannt. In diesem Beitrag wird über ein Forschungsprojekt berichtet, in dem theoriebasiert und empirisch die Herausforderungen bei der Umsetzung der Datenfallstudie in der Schule identifiziert wurden. Darauf aufbauend wurde die Methode schulspezifisch weiterentwickelt. Das Ergebnis ist eine Spezifizierung der Datenfallstudie als handlungsorientierte und lernendenzentrierte Lehr-Lern-Methode, die an die Ziele der schulischen KI-Bildung angepasst ist und zur Förderung datenbasierter Urteils- und Problemlösekompetenz eingesetzt werden kann.

Keywords: Datenfallstudie, Data Case Study, Lab, Sekundarstufe, KI-Bildung, Datenbildung

1 Einleitung

Die datenwissenschaftliche Fallstudie (Datenfallstudie) ist eine in der akademischen Bildung im Kontext von Künstlicher Intelligenz (KI) und Datenwissenschaft verbreitete handlungsorientierte und lernendenzentrierte Lehr-Lern-Methode [HI18; NS01; Wr24].² Studierende an Hochschulen bearbeiten Datenfälle um in Vorlesungen theoretisch vermittelte Datenkonzepte und -praktiken in realen sowie authentischen Kontexten anzuwenden und hierbei eine datenbasierte Urteils- sowie Problemlösekompetenz zu entwickeln [La21]. Im Prozess verinnerlichen die Studierenden die fundamentale Bedeutung von Daten für KI-Systeme, insbesondere bei solchen, die auf maschinellen Lernverfahren basieren. Auch in der schulischen Bildung zum Thema „KI“ wird die Förderung eines fundamentalen Datenverständnisses und einer datenbasierten Urteils- und Problemlösekompetenz als zentral angesehen. Gemäß den Bildungsstandards Informatik sollen Schülerinnen und Schüler (SuS) beispielsweise den Einfluss der verwendeten Daten bei Anwendung eines Verfahrens des maschinellen Lernens beschreiben können [He24]. Es stellt sich deshalb die Frage, ob mit der Datenfallstudie auch die schulische Bildung zum Thema „KI“ methodisch bereichert werden kann.

Die Fallstudie ist als Lehr-Lern-Methode zur Förderung der eigenständigen Problemlösefähigkeit im Informatikunterricht bereits bekannt [Ze18]. Die Datenfallstudie stellt jedoch

¹ Freie Universität Berlin, Didaktik der Informatik, Königin-Luise-Str. 24-26, 14195 Berlin, Deutschland, viktoriya.olari@fu-berlin.de, <https://orcid.org/0000-0002-5113-6624>; ralf.romeike@fu-berlin.de, <https://orcid.org/0000-0002-2941-4288>

² Anschauliche Beispiele (z. B. „Predicting Annual Air Pollution“) bieten Wright et al. [Wr24].

eine komplexere Form des Lernens dar, bei deren Einsatz in der Schule Herausforderungen anzunehmen sind. Beispielsweise verfügen viele SuS nicht über die für die Durchführung der Datenvorbereitung erforderlichen Programmierkenntnisse. Auch die Zeitressourcen der Lehrkräfte bei der Unterrichtsplanung sind beschränkt. Die Vorbereitung der akademischen Datenfälle setzt jedoch umfassende Programmierkenntnisse voraus [NS01] und wird als zeitintensiv angesehen [La21]. Deshalb ist die Frage zu klären, wie die Datenfallstudie als Lehr-Lern-Methode in der Schule fruchtbar eingesetzt werden kann.

Um ihr Potenzial als Lehr-Lern-Methode mit den Herausforderungen bei der Umsetzung in der Schule in Einklang zu bringen, wurde die Datenfallstudie für den Einsatz im Informatikunterricht adaptiert und dem Ansatz der fachdidaktischen Entwicklungsforschung folgend ausgebaut. Dabei wurden folgende Forschungsfragen bearbeitet:

F1: Welche Herausforderungen sind zu berücksichtigen, um die Datenfallstudie als Lehr-Lern-Methode im Informatikunterricht zum Thema „KI“ fruchtbar einzusetzen?

F2: Wie muss die Datenfallstudie als Lehr-Lern-Methode adaptiert werden, um die festgestellten Herausforderungen zu bewältigen und den Zielen des Informatikunterrichts zum Thema „KI“ zu entsprechen?

Zunächst wird der Forschungsstand skizziert. Anschließend wird das methodische Vorgehen erläutert, und die Ergebnisse des Forschungsprozesses werden präsentiert. Mit dieser Arbeit wird ein zweifacher Beitrag geleistet: Neben der theoriebasierten und empirischen Adaption der Datenfallstudie als Lehr-Lern-Methode werden Beispiele erprobter Datenfälle für den Informatikunterricht geliefert.³

2 Forschungsstand

Das Thema „Daten“ ist zentral in der Schulinformatik [He24] und erhält mit der Integration der Inhalte zum Bereich „KI“ in die Informatiklehrpläne eine zusätzliche Relevanz. Daten sind das Fundament von KI-Systemen, insbesondere solchen, die auf maschinellen Lernverfahren basieren [Zh25]. Um solche Systeme zu verstehen, zu bewerten, verantwortungsvoll zu nutzen und zu gestalten, ist ein erweitertes Datenverständnis notwendig. Dazu gehören Kenntnisse zu Datenpraktiken wie Datenerfassung, -bereinigung, -konstruktion und -aufteilung [OR24b] sowie zu grundsätzlichen Datenkonzepten wie Datenrauschen oder -verzerrung [OR24a]. In Studien zur Vermittlung der Datenkonzepte und -praktiken wird über Experimente mit Unplugged-Ansätzen [Wa20] sowie über Projektarbeiten in textuellen [PHS22] und grafischen Umgebungen berichtet [GR19]. Jedoch wurde auch gezeigt, dass in derzeitigen Unterrichtsansätzen das Thema „Daten“ nur oberflächlich behandelt wird [OR21]. Gleichzeitig wird betont, dass die Auseinandersetzung mit Daten SuS substanzielle Lerngelegenheiten eröffnen kann, um KI-Technologien zu verstehen [TDT21].

³ Die Datenfälle werden im Rahmen der Heftreihe *Data Cases. Datenfallstudien für den Informatikunterricht zum Thema Daten und Künstliche Intelligenz* veröffentlicht: <https://computingeducation.de/proj-datacases/>.

Die Datenfallstudie als Lehr-Lern-Methode (auch bekannt als „Case Study“ oder „Lab“) ist in der akademischen Bildung in KI und Datenwissenschaft weit verbreitet [Ch21; La21; NS01; Wr24]. Im Zentrum der Methode steht die Arbeit mit einem konkreten Datenfall. Ein akademischer Datenfall besteht aus einer Beschreibung einer problematischen Situation, einem Datensatz sowie einem exemplarischen Weg zur Bearbeitung des Datenfalls [Ch21; NS01; Wr24]. Methodisch kann die Arbeit am Datenfall beispielsweise nach einer Schritt-für-Schritt-Anleitung [Wr24] oder in offener Form mit Aufgaben und Übungen durchgeführt werden [Ho20; La21]. Aus lerntheoretischer Sicht steht die Datenfallstudie in der Tradition des konstruktivistischen [Gi11; Ho20], aktiven [He11] und situierten [Ho20] Lernens. In Forschungsarbeiten aus dem akademischen Kontext ist gezeigt worden, dass sich das Lernen mit der Datenfallstudie positiv auf das Interesse, die Motivation und das Engagement der Studierenden sowie auf den Wissens- und Kompetenzzuwachs auswirkt [He11; Ho20; La21; To23]. In der schulischen Bildung ist die Lehr-Lern-Methode „Fallstudie“ im Informatikunterricht bekannt [Ze18], beispielsweise im Kontext von Datenbanken [Pe94].

3 Methodik

Um Potenziale und schulspezifische Herausforderungen in Einklang zu bringen, wurde die Datenfallstudie als Lehr-Lern-Methode adaptiert und weiterentwickelt. Hierzu wurde ein Expertenteam gebildet, das aus zwei Informatiklehrkräften, einem Domainexperten und zwei Forschenden der Informatikdidaktik bestand, die sich während eines Zeitraums von zwei Jahren regelmäßig trafen. Als methodischer Rahmen wurde der Ansatz der fachdidaktischen Entwicklungsforschung gewählt [Pr12]. Der Forschungsprozess bestand aus vier eng miteinander vernetzten Phasen, die in drei Zyklen durchgeführt wurden: (1) Lerngegenstand spezifizieren und strukturieren, (2) Design (weiter)entwickeln, (3) Design-Experimente durchführen und auswerten sowie (4) lokale Theorien (weiter)entwickeln.

(1) Was soll gelernt werden? Die erste Phase umfasste die Spezifizierung des Lerngegenstands. Um Datenkonzepte und -praktiken zu identifizieren, die im Unterricht zum Thema „KI“ relevant sind, wurde eine Analyse der Fachliteratur und der Informatiklehrpläne durchgeführt [OR24a; OR24b]. Daraus wurden Konzepte und -praktiken für den ersten Zyklus ausgewählt, die auch in akademischen Datenfällen bearbeitet werden (z. B. Daten aufbereiten) (vgl. etwa [Wr24]). Im zweiten Zyklus wurde die inhaltliche Auswahl weiter eingeschränkt.

(2) Wie soll gelernt werden? In der zweiten Phase wurde die akademische Datenfallstudie für den schulischen Kontext adaptiert sowie inhaltlich umgesetzt. Wenngleich Herausforderungen und Ansprüche an das Design sich teilweise aus der Forschung ableiten ließen – etwa die Bedeutung kognitiver Aktivierung der SuS bei der Gestaltung der lernwirksamen Aktivitäten [KI20] –, waren bei der Umsetzung unbekannte Schwierigkeiten zu erwarten. Um diese frühzeitig zu identifizieren und zu berücksichtigen, wurde jeder Datenfall in vier Schritten entwickelt: 1) Austausch über potenzielle Kontexte und Herausforderungen im Expertenteam, 2) Ermittlung geeigneter Kontexte und Datensätze sowie Erarbeitung eines

Vorschlags für den Datenfall durch Didaktikforschende, 3) fachliche Sicherung des Vorschlags durch Domainexperten, 4) didaktische Sicherung des Vorschlags durch Lehrkräfte. So entstanden in zwei Jahren 18 Datenfälle.

(3) Wie gelingt die Umsetzung? Die dritte Phase umfasste die Erprobung der Datenfälle mit SuS und die Identifizierung weiterer Herausforderungen. Die Erprobung fand wöchentlich während der regulären Unterrichtszeit im Rahmen von drei dreimonatigen Enrichmentkursen in der Fächerkombination Informatik/Mathematik statt. An den Kursen nahmen 44 SuS der 9. und 10. Klassen eines Gymnasiums teil (15 SuS in Kurs 1 ($w = 5, m = 8; o. A. = 2; \emptyset\text{-Alter} = 14,46 \text{ Jahre}$), 13 SuS in Kurs 2 ($w = 4, m = 7; o. A. = 2; \emptyset\text{-Alter} = 14,7 \text{ Jahre}$), 16 SuS in Kurs 3 ($w = 4, m = 10, d = 1, o. A. = 1; \emptyset\text{-Alter} = 14,72 \text{ Jahre}$)). Jeder Kurs beinhaltete eine zweitägige Exkursion zum GEOMAR Helmholtz-Zentrum für Ozeanforschung, der Arbeitsstelle des Domainexperten. Zur Untersuchung von Herausforderungen und Gelingensbedingungen wurde eine breite Palette direkter und indirekter Indikatoren analysiert [Va99], über die in einem anderen Artikel bereits berichtet wurde [OR25]. So wurden beispielsweise wöchentlich Wissen und Motivation in Pre-Post-Tests evaluiert und Arbeitsergebnisse eingesammelt. Didaktikforschende beobachteten und videographierten den Unterricht; zudem wurden mit Lehrkräften sowie SuS Interviews durchgeführt und deduktiv sowie induktiv im Hinblick auf Herausforderungen, Wirksamkeit, Motivation und Gelingensbedingungen ausgewertet.

(4) Welche Faktoren sind für das Lernen wesentlich? Die vierte Phase diente der Reflexion bezüglich der Herausforderungen und der Wirksamkeit der Ausgestaltung der Datenfälle für den Lernprozess unter realen Unterrichtsbedingungen. Das Ergebnis ist eine Spezifizierung der Datenfallstudie als Lehr-Lern-Methode, die Herausforderungen und Zielen des Informatikunterrichts zum Thema „KI“ entspricht.

4 Ergebnisse

Im Folgenden werden die exemplarischen Herausforderungen und Bewältigungsstrategien im Entwicklungsforschungsprozess (F1) erläutert, gefolgt von einer Spezifizierung der Datenfallstudie als Lehr-Lern-Methode (F2).

4.1 Alte Herausforderungen, neue Herausforderungen

Die Identifizierung und Bewältigung von Herausforderungen war ein iterativer, auf Theorie und Praxis gestützter Prozess. Sobald bekannte Herausforderungen bewältigt waren, traten neue, detailliertere auf. Im Folgenden werden zentrale und spezifische Herausforderungen sowie Bewältigungsstrategien bei der Adaptierung der Datenfallstudie für die Schule erörtert.

1. Fachliche Herausforderungen. Als handlungsorientierte Methode setzt die Datenfallstudie in der akademischen Bildung umfangreiche Programmierkenntnisse voraus, insbesondere

in Datenmanipulation, Visualisierung und Modellierung. In der Sekundarstufe I, in der die Modellierung und Implementierung von Datenflüssen behandelt wird [He24], ist mit heterogenen Vorkenntnissen zu rechnen – was sich in den Vorbefragungen bestätigte. Die SuS, die die Datenfälle erprobten, hatten heterogene Programmierkenntnisse und wenig Erfahrung im Umgang mit Daten, etwa in Tabellenkalkulationsprogrammen. Zur Bewältigung dieser initialen Herausforderung wurde im ersten Zyklus ein Satz stark angeleiteter Bottom-up-Datenfälle für die intuitive, flussbasierte Umgebung „Orange3“ entwickelt. Jeder Datenfall startete mit der leeren Arbeitsfläche in Orange3 (Abb. 1 links). Die SuS wurden Schritt für Schritt durch Datenpraktiken in einem Kontext geleitet. Die Evaluation zeigte, dass die Herausforderung der heterogenen Vorkenntnisse durch die Bottom-up-Architektur erfolgreich adressiert wurde. Alle SuS waren in der Lage, Datenflüsse nach Anleitung zu erstellen. Nachteilig wirkte jedoch im ersten Zyklus die notwendige Überbetonung der Bedienung von Orange3, wodurch wichtigere Kompetenzen vernachlässigt wurden, beispielsweise die Reflexion bezüglich des Prozesses und die Hinführung auf die abstrakte Ebene.

Aus diesem Grund wurde im zweiten Zyklus in starker Orientierung an die methodische Vorgehensweise im akademischen Kontext ein Satz Datenfälle mit Top-down-Architektur entwickelt (Abb. 1 mittig). Hierbei begann jeder Datenfall mit einem vorbereiteten Datenfluss, also dem vollständigen Weg der Daten von der ersten bis zur letzten Recheneinheit (in Orange3 „Widget“ genannt), und beinhaltete eine Reihe Interpretations- sowie Reflexionsaufgaben. Damit wurde die kleinschrittige Anleitung durch eine kriteriengeleitete Reflexion des Prozesses ersetzt. Die wöchentlichen Arbeitsergebnisse und Abschlussprojekte zeigten, dass alle SuS vorbereitete Datenflüsse nachvollziehen, interpretieren und modifizieren sowie eigene erfolgreich erstellen konnten. Dennoch traten zwei neue Herausforderungen auf: (1) Einige SuS hatten Schwierigkeiten, das Datenflusskonzept zu verstehen. So stellte der Lehrer nach dem vierten Datenfall im zweiten Kurs fest: *„Die Hilfekarten zu den Widgets haben sie auch aktiv benutzt. [...] Aber zum Beispiel eine Gruppe hatte gar nicht verstanden, wie das mit der Verbindung [...] der Widgets funktioniert, dass da die Daten von einem Widget zum nächsten fließen.“* (SIT2W5, Pos. 63) (2) Zudem fehlten SuS Debuggingstrategien: *„Die haben zum Beispiel auch Probleme dabei gehabt, dieses Datenentdecken zu machen [...] Die Debuggingstrategien fehlen.“* (SIT2W5, Pos. 63) .

Beide Herausforderungen ließen sich durch die explizite Vermittlung des Datenflusskonzepts und von Debuggingstrategien im dritten Kurs mit puzzleartigen Datenfällen erfolgreich adressieren. Bei dieser Architektur starten die SuS mit einem Satz von Spielsteinen (bestehend aus Widgets, Datenübersichten, Vorlagen zur Konfiguration der Widgets) zu einem Kontext und müssen rückwirkend den zugrundeliegenden Datenfluss ermitteln und Widgetkonfigurationen rekonstruieren (Abb. 1 rechts). Im Rekonstruktionsprozess entdeckten SuS intuitiv Elemente, die für das Debugging der Datenflüsse wesentlich sind (etwa die Anzahl der Daten, die in ein Widget ein- und aus ihm ausgehen).

2. Fachdidaktische Herausforderungen. Neben unterschiedlichen Voraussetzungen muss die Datenfallstudie verschiedene pädagogische Ziele adressieren. Zu Beginn des ersten

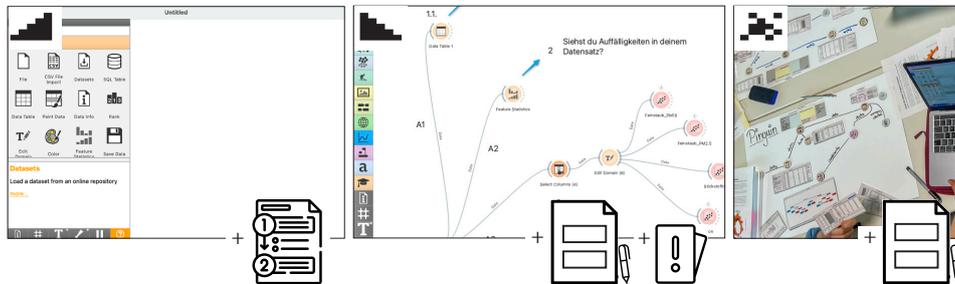


Abb. 1: Architektur eines Datenfalls: bottom-up (, top-down (, puzzelartig ().

Zyklus durchgeführte Lehrkräfteinterviews zeigen, dass ein zentrales Bildungsziel darin besteht, SuS zu sinnstiftender Kommunikation anzuleiten. Dafür muss die Datenfallstudie als Lehr-Lern-Methode entsprechende Gelegenheiten bieten. In der akademischen Bildung sind Datenfälle jedoch oft für Einzelarbeit konzipiert und ermöglichen wenig Austausch [Wr24]. Um diese Herausforderung zu adressieren, wurden im ersten und zweiten Zyklus verschiedene Formen der Partner- und Gruppenarbeit in die Arbeit an Datenfällen integriert (z. B. Lerntempoduett, Gruppenpuzzle). Diese Maßnahmen waren teilweise erfolgreich. In der Partnerarbeit wurde oft inhaltlicher Austausch zwischen SuS beobachtet; Gruppenarbeiten hingegen erwiesen sich als weniger effektiv. Lehrkräfte kritisierten, dass Datenfälle nur wenige Gelegenheiten für Gespräche bieten: „*Wo ich wirklich einem anderen mal was sagen muss, wo ich wirklich (...) ins echte Gespräch komme.*“ (SIT1W3, Pos. 97). Zudem führten Gruppenarbeiten zu Problemen, da SuS unterschiedlich schnell waren und oft auf ihre Teammitglieder warten mussten. Auch vertiefende Aufgaben als Lösung erwiesen sich als ungeeignet, da nicht die Leistungsstärkeren schneller waren, sondern oft oberflächlich arbeitende leistungsschwächere SuS (SIT1W3, Pos. 16). Durch die Entwicklung der puzzelartigen Architektur, bei der SuS gemeinsam die Aufgabe lösen mussten, konnte sinnstiftende Kommunikation im dritten Zyklus erfolgreich eingesetzt werden.

3. Domänenspezifische Herausforderungen. Zur Beurteilung der Plausibilität eines Datensatzes (z. B. im Kontext von Datenverzerrungen) oder zur Interpretation der Ergebnisse im Kontext ist Domänenwissen erforderlich. In der akademischen Bildung erhalten Studierende Informationsmaterial zur Einarbeitung in die Domäne [NS01]. Dafür fehlt Lehrkräften und SuS jedoch die Zeit. Eine Lösung besteht darin, Kontexte zu wählen, in denen Lehrkräfte und SuS Vorwissen haben. Es wurde das Thema „Umwelt“ gewählt, das in früheren Klassenstufen behandelt wird und persönlich sowie gesellschaftlich relevant ist. Im Unterricht konnte jedoch beobachtet werden, dass das Vorwissen der SuS häufig nicht ausreichte, um Ergebnisse im Kontext zu bewerten. Obwohl diese Herausforderung vermutet war und durch Informationsmaterial adressiert wurde, erwies sich dessen Einsatz im ersten Zyklus als nur mäßig erfolgreich: „*Diese Idee, dass ich das Wissen aus vorherigen Aufgaben in die neue Aufgabe einpflanzen muss, ist für mich als Schüler schon schwer.*“ (SIT1W3, Pos. 42). Im zweiten Zyklus wurden die Interpretationsaufgaben mit expliziten

Verweisen auf das Informationsmaterial ergänzt (z. B.: „Vergleiche das Ergebnis mit deiner Intuition. Wenn du keine Intuition hast, lies das Feinstaub-Infoblatt. Nimm Stellung dazu, ob die Vorhersage plausibel ist.“). Zusätzlich wurden Interpretationskarten mit Kriterien zur Ergebnisbeschreibung entwickelt. Diese Maßnahmen erwiesen sich als hilfreich. Eine Lehrkraft fasste in einem Interview zusammen: „Sie [die SuS] waren in der Lage, sich das Vorwissen quasi zu beschaffen, über die Informationskarten, und es dann anzuwenden.“ (SIT2W5, Pos. 103).

4. Konzeptionelle Herausforderungen. Manche Unterrichtsmethoden können Lehrkräfte selbst aufbereiten, andere erfordern spezialisierte Expertise. Ein Beispiel sind Leitprogramme – Arbeitshefte mit 50 bis 100 Seiten, deren Erstellung Fachkompetenz, didaktische Phantasie, Sprachtalent und Zeit verlangt [Ki97]. Datenfälle sind in dieser Hinsicht vergleichbar. Das zentrale Element eines Datenfalls ist ein aufgearbeitetes Szenario mit zugehörigem Datensatz. Aus der akademischen Bildung in Datenwissenschaften ist bekannt, dass die Suche nach einem Datensatz und dessen Aufbereitung aufwändig ist [HI18]. Das bestätigte sich für die Sekundarstufe I. Zwar wurde diese Herausforderung adressiert, indem der Kontext auf ein Fachgebiet eingeschränkt wurde, zu dem viele Datensätze frei verfügbar sind und für das der Domainexperte die Plausibilität der Szenarien und die Eignung der Datensätze einschätzen konnte. Dennoch betrug der Entwicklungsaufwand pro Datenfall 40 bis 120 Stunden. Die Suche nach geeigneten Datensätzen – also solchen, die gut dokumentiert, für SuS der Sekundarstufe I nachvollziehbar und gleichzeitig für das zu vermittelnde Datenkonzept oder die Datenpraktik sinnvoll sowie fachlich tragfähig sind – erwies sich als mit Abstand anspruchsvollste Aufgabe. Eine enge Zusammenarbeit zwischen Fachdidaktik, Domänenexperten und Lehrkräften ist für die Entwicklung neuer Datenfälle deshalb wesentlich.

4.2 Spezifizierung der Datenfallstudie als Lehr-Lern-Methode

Neben den dargestellten Schwierigkeiten wurden im Prozess weitere schulische Herausforderungen erfolgreich adressiert, darunter hardwareseitige Einschränkungen für komplexe Datenanalysen, Frustrationserlebnisse als integraler Bestandteil der Arbeit mit Daten und begrenzte Hausaufgabenzeit. Im Folgenden wird die aus dem Forschungsprozess resultierende schulspezifische Charakterisierung der Datenfallstudie als Lehr-Lern-Methode erläutert.

Die Datenfallstudie ist eine handlungsorientierte, lernendenzentrierte Lehr-Lern-Methode. Im Zentrum der Datenfallstudie steht die Arbeit der SuS an einem Datenfall. Zentrale Bestandteile eines schulspezifischen Datenfalls sind: (1) die Beschreibung einer problematischen Situation aus einem für SuS relevanten Kontext, (2) ein dokumentierter, alltagsrelevanter, auf die Situation ausgerichteter, gegebenenfalls didaktisch angereicherter Datensatz, an dem ausgewählte Datenkonzepte und -praktiken sinnstiftend sowie fachgerecht angewendet werden können, (3) ein vollständiger oder teilweise ausgearbeiteter Datenfluss mit Aufgaben zur Beschreibung und begründeten Bewertung der Datenverarbeitungsergebnisse sowie

Aufgaben zur Modifikation und Erweiterung des Datenflusses, einschließlich Ideen für weiterführende Datenanalysen, (4) eine in Aufgabenstellungen explizit eingebettete Sammlung von Domäneninformationen, (5) Informationsmaterial zum Datenanalysewerkzeug und zur Interpretation der Ergebnisse, (6) optional: mehrstufiges Hilfematerial, Lernkontrollen und Übungen. Für den Einsatz im Unterricht wird ein Datenfall außerdem (7) von einem Erwartungshorizont mit Hinweisen zu Problemen im Datensatz, dem vorausgesetzten Wissen der SuS bezüglich des Datenanalysewerkzeugs, Datenkonzepten und -praktiken, Zeitangaben sowie Empfehlungen zu Sozialformen begleitet.

Für seinen schulischen Einsatz wird der Datenfall methodisch in drei in Kapitel 4 vorgestellten Architekturen realisiert: top-down, bottom-up und puzzleartig. Die Bottom-up- und die Top-down-Architektur ähneln der methodischen Umsetzung der Datenfallstudie im akademischen Kontext (z. B. [Wr24]). Mit einer Kombination aus Top-down- und puzzleartiger Architektur werden viele schulspezifische Herausforderungen adressiert.

Bei der Bottom-up-Architektur bauen SuS den Datenfluss nach Anleitung auf und bearbeiten nach jedem Teilschritt Aufgaben zum Kontext. Lernprodukte sind der Datenfluss und die schriftliche Interpretation der Erkenntnisse daraus. Besonders leistungsstarke und interessierte SuS sowie Lehrkräfte profitieren von dieser Variante, denn diese Architektur bietet eine geleitete Einführung in ein bestimmtes Thema (z. B. „Exploration der Daten“) und ist für Einzelarbeit gut geeignet.

Bei der Top-down-Architektur erhalten SuS einen fertigen Datenfluss und werden dazu angeleitet, ausgewählte Datenpraktiken und -konzepte in vorbereiteten Szenarien nachzuvollziehen sowie kriteriengeleitet zu bewerten. Dafür arbeiten SuS im Dreischritt: (1) Ergebnisse des Datenflusses beschreiben, interpretieren und mit der Intuition vergleichen (im Kontext von Orange3 ist dies die Arbeit mit Blattwidgets), (2) Zwischenschritte untersuchen und ihre Bedeutung im Datenfluss reflektieren (in Orange3 ist dies die Untersuchung der Vorbereitungswidgets, die zu Blattwidgets führen), (3) Phasen im Datenfluss identifizieren. Lernprodukte sind ein schriftliches Protokoll mit der Interpretation der Erkenntnisse aus dem Datenfluss sowie die Beschreibung von dessen Phasen. Die Top-down-Architektur eignet sich besonders, wenn der Fokus auf der kritischen Reflexion über den Datenfluss und der Übersicht über die Konzepte und Phasen im Datenfluss liegen soll.

Im puzzleartigen Datenfall rekonstruieren die SuS kollaborativ und ohne Computer den Datenfluss aus gegebenen Elementen. Diese Architektur eignet sich für Lerngruppen mit heterogenen Vorkenntnissen und auch besonders, wenn es darum geht, Debuggingkompetenzen, die für die selbständige Entwicklung der Datenflüsse erforderlich sind, zu vermitteln oder sinnstiftende Kommunikation und kooperatives Lernen zu üben. Das Lernprodukt ist der auf einem Poster aus Teilen zusammengesetzte Datenfluss. Denkbar ist eine Erweiterung dieser Architektur, indem Computer eingesetzt werden, um die Ergebnisse der Rekonstruktion zu prüfen.

Lehren und Lernen mit Datenfällen erfordert ausreichend Zeit im Unterricht. Je nach

Architektur und Komplexität eines Datenfalls sollte die eingeplante Dauer eine bis vier Unterrichtsstunden betragen. Der Unterricht, in dem mit der Datenfallstudie gelernt wird, gliedert sich meistens in folgende Phasen: (1) Vorbereitung der Arbeitsumgebung; (2) Kennenlernen des Datenfalls und der Zielsetzung; (3) sinnstiftende Einzel /Partner-/Gruppenarbeit am Datenfluss, unterstützt durch Hilfematerial; (4) fachliche Sicherung der Erkenntnisse (idealerweise durch selbständigen Abgleich eigener Antworten mit Lösungskarten, da in diesem Fall SuS in ihrem Lerntempo bleiben); (5) optional: Reflexion der Datenverarbeitungsschritte im Datenfluss.

5 Fazit

In diesem Beitrag wurde der Entwicklungsforschungsprozess zur Adaptierung der Datenfallstudie als Lehr-Lern-Methode für den Informatikunterricht erläutert. Hierbei waren Herausforderungen wie heterogene Programmier- und Datenanalysevorkenntnisse, fehlendes Domainwissen bei SuS sowie Lehrkräften, hoher Entwicklungsaufwand und häufige Frustrationserlebnisse als Teil der Datenarbeit zu berücksichtigen. Beim Einsatz in der Schulpraxis kamen weitere Schwierigkeiten hinzu, darunter fehlende Debuggingkompetenzen, unterschiedliche Lerngeschwindigkeiten und begrenzte Gelegenheiten zur sinnstiftenden Kommunikation. Es zeigt sich, dass die Datenfallstudie als Lehr-Lern-Methode für den Einsatz in der Schule dennoch geeignet ist. In ihren drei Architekturen (top-down, bottom-up und puzzleartig) befähigt sie SuS, KI-relevante Datenkonzepte und -praktiken in realen sowie authentischen Kontexten zu verstehen, zu bewerten und anzuwenden. Somit kann durch den Einsatz dieser Methode im Informatikunterricht zum Thema „KI“ ein Beitrag zur Entwicklung datenbasierter Urteils- und Problemlösekompetenzen geleistet werden.

Literaturverzeichnis

- [Ch21] Chkoniya, V.: Success Factors for Using Case Method in Teaching Applied Data Science Education. *European Journal of Education* 4 (1), S. 77–86, 2021, Stand: 23. 08. 2024.
- [Gi11] Gill, T. G.: *Informing with the Case Method: A Guide to Case Method Research, Writing, & Facilitation*. Informing Science Press, Santa Rosa, 2011.
- [GR19] Grillenberger, A.; Romeike, R.: About Classes and Trees: Introducing Secondary School Students to Aspects of Data Mining. In (Pozdniakov, S. N.; Dagiene, V., Hrsg.): *Informatics in Schools. New Ideas in School Informatics*. Springer, Cham, S. 147–158, 2019.
- [He11] Herreid, C. F.: Case Study Teaching. *New Directions for Teaching and Learning* 2011 (128), S. 31–40, 2011, Stand: 14. 01. 2025.
- [He24] Hellmig, L. et al.: 66 - Empfehlungen der Gesellschaft für Informatik e.V.: *Bildungsstandards Informatik für die Sekundarstufe I*. 2024, Stand: 16. 02. 2025.
- [HI18] Hicks, S. C.; Irizarry, R. A.: A Guide to Teaching Data Science. *The American Statistician* 72 (4), S. 382–391, 2018, Stand: 04. 11. 2024.
- [Ho20] Hoffer, E. R.: Case-Based Teaching: Using Stories for Engagement and Inclusion. *IJonSES* 2 (2), S. 75–80, 2020.

- [Ki97] Kirchgraber, U. et al.: Und Dann Und Wann Ein Leitprogramm! Zentralblatt für Didaktik der Mathematik 29 (6), S. 199–206, 1997.
- [Kl20] Klieme, Eckhard: Unterrichtsqualität. In: Handbuch Schulpädagogik. Utb-Titel Ohne Reihe, Waxmann, S. 394–409, 2020, Stand: 28. 02. 2024.
- [La21] Lasser, J. et al.: Introductory Data Science across Disciplines, Using Python, Case Studies, and Industry Consulting Projects. Teaching Statistics 43 (S1), 2021, Stand: 17. 04. 2022.
- [NS01] Nolan, D. A.; Speed, T.: Stat Labs: Mathematical Statistics through Applications. Springer, New York Berlin Heidelberg, 2001.
- [OR21] Olari, V.; Romeike, R.: Addressing AI and Data Literacy in Teacher Education: A Review of Existing Educational Frameworks. In: 16th WiPSCE Conf., Proceedings. ACM, 2021.
- [OR24a] Olari, V.; Romeike, R.: Data-Related Concepts for Artificial Intelligence Education in K-12. Computers and Education Open 7, S. 100196, 2024, Stand: 24. 07. 2024.
- [OR24b] Olari, V.; Romeike, R.: Data-Related Practices for Creating Artificial Intelligence Systems in K-12. In: 19th WiPSCE Conference, Proceedings. ACM, Munich, Germany, 2024.
- [OR25] Olari, V.; Romeike, R.: Data Case Study – A Teaching and Learning Method for Computer Science Education in Schools. In: 3rd CompEd conference, Proceedings. ACM, 2025.
- [Pe94] Penon, J., Spolwig, S.: Video-Center – Eine Fallstudie Zur Einführung in Relationale Datenbanksysteme. LOG IN, 2, 32–35. LOG IN 2, S. 32–35, 1994.
- [PHS22] Podworny, S.; Hüsing, S.; Schulte, C.: A Place for a Data Science Project in School: Between Statistics and Epistemic Programming. STATISTICS EDUCATION RESEARCH JOURNAL 21 (2), S. 6, 2022, Stand: 05. 04. 2024.
- [Pr12] Prediger, S. et al.: Lehr-Lernprozesse initiieren und erforschen. Fachdidaktische Entwicklungsforschung im Dortmunder Modell. Der mathematische und naturwissenschaftliche Unterricht 65 (8), S. 452–457, 2012.
- [TDT21] Tedre, M.; Denning, P.; Toivonen, T.: CT 2.0. In: Proceedings of the 21st Koli Calling International Conference on Computing Education Research. Koli Calling '21, Association for Computing Machinery, New York, NY, USA, S. 1–8, 2021, Stand: 31. 01. 2023.
- [To23] Toogood, C.: Supporting Students to Engage with Case Studies: A Model of Engagement Principles. Educational Review, S. 1–15, 2023, Stand: 01. 12. 2024.
- [Va99] Van Den Akker, J.: Principles and Methods of Development Research. In (Van Den Akker, J. et al., Hrsg.): Design Approaches and Tools in Education and Training. Springer Netherlands, Dordrecht, S. 1–14, 1999, Stand: 08. 02. 2025.
- [Wa20] Wan, X. et al.: SmileyCluster: Supporting Accessible Machine Learning in K-12 Scientific Discovery. In: Proceedings of the Interaction Design and Children Conference. ACM, London United Kingdom, S. 23–35, 2020, Stand: 06. 01. 2024.
- [Wr24] Wright, C. et al.: Open Case Studies: Statistics and Data Science Education through Real-World Applications. Journal of Statistics and Data Science Education, S. 1–30, 2024.
- [Ze18] Zandler, A., Hrsg.: Unterrichtsmethoden für den Informatikunterricht. Springer Fachmedien Wiesbaden, Wiesbaden, 2018.
- [Zh25] Zha, D. et al.: Data-centric Artificial Intelligence: A Survey. 57 (5), 2025, <https://doi.org/10.1145/3711118>.