

Data Literacy Framework

als Grundlage für die Entwicklung
didaktischer Vermittlungskonzepte von Data Literacy

i-science Tag, Potsdam

Katharina Schüller

15. Juli 2019

Visualisierungen: Markus Armbruster | pictomind GmbH

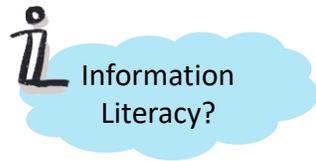


Data Literacy als 21st century skill

Hintergrund und Zielsetzung der Studie

- Kompetenzrahmen für Digitalkompetenzen am Beispiel von Data Literacy entwickeln und umsetzbares Wissen für Hochschulen verfügbar machen
- Messung von Wirkung und Qualität von Lehre und Studium im digitalen Zeitalter wie auch die Testentwicklung für Digitalkompetenzen am Beispiel von Data Literacy vorbereiten

“Wir ertrinken in Informationen und dürsten nach Wissen.” (Naisbitt, 1982)



„Wenn wir mündige Bürger in einer modernen technologischen Gesellschaft möchten, dann müssen wir ihnen drei Dinge beibringen: Lesen, Schreiben und statistisches Denken, das heißt den vernünftigen Umgang mit Risiken und Unsicherheiten.“ (Wells, 1903)

... an einen Kompetenzrahmen

- Bildet alle Stufen des Wissens- bzw. Wertschöpfungsprozesses aus Daten ab
- Erfasst alle Kompetenzdimensionen: (a) Wissen, (b) Fertigkeiten, (c) Fähigkeiten, (d) Motivation und (Wert-)Haltung
- Erlaubt es, die erfassten Kompetenzen in konkrete und testbare Lern- oder Kompetenzziele zu überführen
- Reflektiert die Interdisziplinarität der Aufgabe:
 - Was will ich? (Fachexperte)
 - Was kann ich? (Datenexperte)
 - Was darf ich? (Datenschützer)
 - Was soll ich? (Datenethiker)

... an Mess- und Testinstrumente

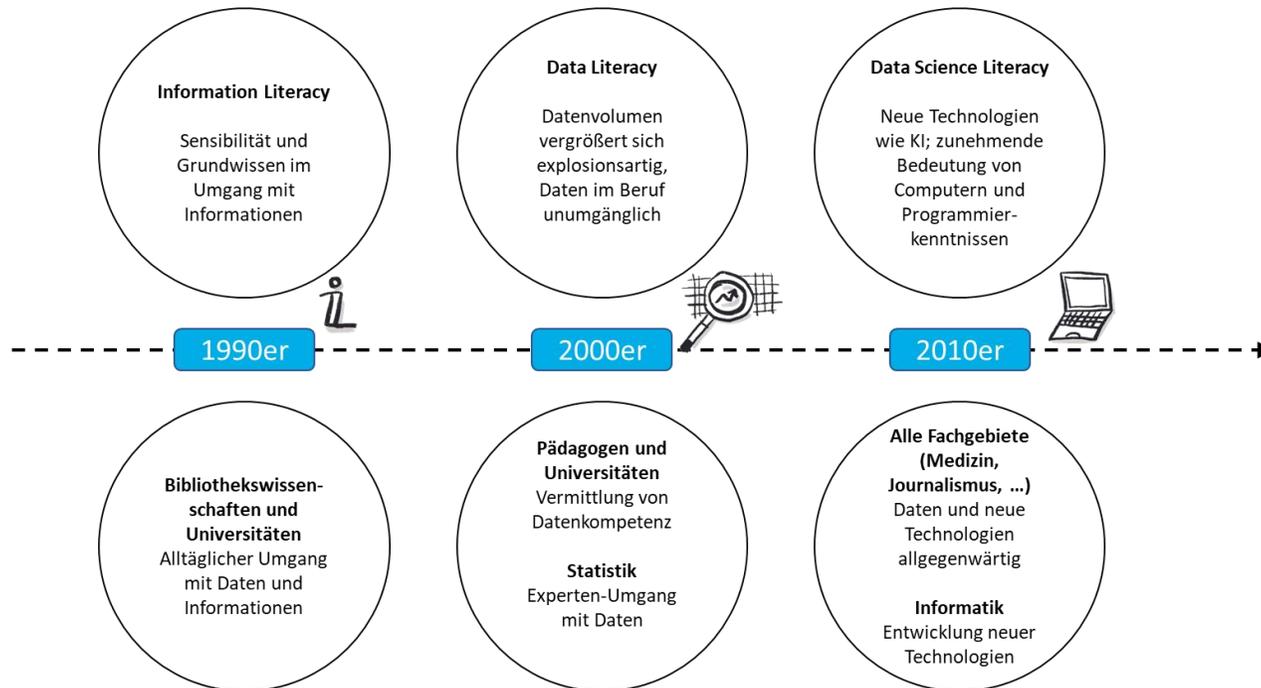
- Erfassen kognitive und affektive Lernbereiche
- Erfassen möglichst viele Lernstufen: (a) Reaktion, (b) Lernerfolg, (c) Verhalten, (d) Ergebnis
- Sind transparent bezüglich der Möglichkeiten und Grenzen einer Schlussfolgerung von beobachtbarem Verhalten auf dahinter liegender Kompetenz
- Genügen den Testgütekriterien der Validität, Reliabilität und Objektivität
- Sind mit vertretbarem Aufwand (Geld, Zeit, benötigte Fähigkeiten der Prüfer) durchzuführen

Der Kompetenzrahmen

Vom Aufgabenverständnis zu den Kompetenzen

Was verstehen wir unter Data Literacy – und was nicht?

GEPRÄGTE BEGRIFFE



PRÄGENDE BEREICHE

▪ Data Science Literacy

- nicht 21st century skill, sondern eher **Kompetenzprofil** eines neuen Berufes

▪ Data Literacy

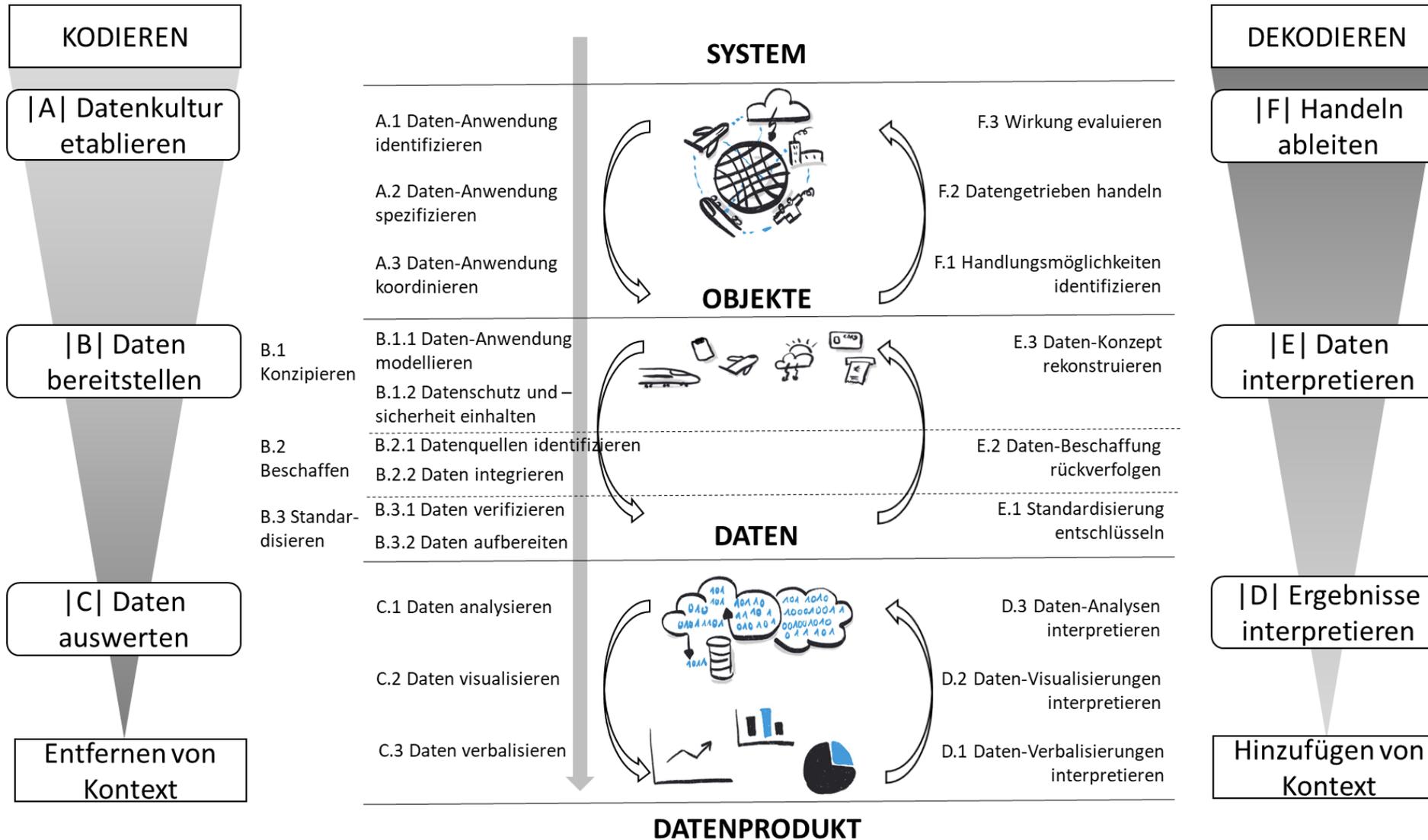
- Cluster aller effizienten Verhaltensweisen und Einstellungen für die effektive Erfüllung sämtlicher **Prozessschritte zur Wertschöpfung bzw. Entscheidungsfindung aus Daten.**

▪ Haltung („Ethical Literacy“):

- Fähigkeit, die **Bedeutung von Daten zur Entscheidungsfindung** vollständig zu erfassen, indem mögliche Interpretationen dieser Daten **in unterschiedlicher Kontextualisierung reflektiert** und kritisch bewertet werden.

Vom kompetenzdefinierenden Prozess zum Kompetenzrahmen

Data Literacy als übergreifende Kompetenz



Beispiel C.1 Daten analysieren (Kodieren = produktive Seite)

Kompetenz		Kompetenzdimensionen			Kompetenzniveaus
Bezeichnung	Beschreibung	Beispiele für Wissen („Knowledge“)	Beispiele für Fähigkeiten („Skills“)	Beispiele für Haltung („Attitude“)	Beispiele für aufsteigende Niveaus
C.1: Daten analysieren	Setzt Analyseverfahren aus verschiedenen Gebieten (Statistik, Analytics, Machine Learning), unter Zuhilfenahme der geeigneten Werkzeuge sach- und zweckorientiert ein	<ul style="list-style-type: none"> • Kenntnis von Verfahren für unterschiedliche Aufgaben (Deskription, Explikation, Prognose) sowie über deren Voraussetzungen, Stärken und Schwächen • Wissen über Verfahren für gerichtete und ungerichtete Fragestellungen • Wissen über Stärken und Grenzen von Algorithmen • Wissen über Möglichkeiten der Modelldiagnostik und der Modelloptimierung • Wissen über Möglichkeiten der Robustheit und allgemeine Gültigkeit der Inferenzsicherzustellen (z.B. Kreuzvalidierung) • Wissen über mögliche Ursachen von Artefakten • Wissen über Best Practices und Standards der Modellentwicklung und Datenanalyse, z.B. Validierung • Wissen, dass bei der Analyse von Daten Information verloren geht und dass ein Auswahl- und Bewertungsprozess nötig ist, auf welche Information verzichtet werden kann 	<ul style="list-style-type: none"> • Fähigkeit, messbare Beziehungen in Modellen abzubilden • Fähigkeit, anhand der jeweiligen Fragestellung und der verfügbaren Daten geeignete Analyseverfahren zu identifizieren und auszuwählen • Fähigkeit, das Modell geeignet zu spezifizieren (z.B. durch Festlegung von Optimierungskriterien) • Fähigkeit, das Modell auf Schwächen und Artefakte (z.B. Overfitting, Multikollinearitätsprobleme) zu untersuchen und diesen entgegenzuwirken • Fähigkeit, die Unsicherheit der Modellergebnisse zu beurteilen und die benötigte Genauigkeit zu bestimmen • Fähigkeit, zukünftige Nutzungen der Analyseergebnisse zu antizipieren • Fähigkeit, relevante und irrelevante Informationen im Analyseprozess zu trennen (z.B. Variablen zu selektieren) 	<ul style="list-style-type: none"> • Bereitschaft, in einem iterativen und oft zeitaufwändigen Prozess Modelle zu implementieren und anzupassen • Skeptische Grundhaltung bei der Datenanalyse • Bereitschaft, Informationsverluste im Prozess der Analyse abzuwägen und zu akzeptieren • Bereitschaft, „good analytics standards“ einzuhalten, selbst wenn diese nicht explizit definiert sind • Bereitschaft, ressourcenschonend zu arbeiten, z.B. das Modell nicht zu „over-engineeren“ • Bereitschaft, die benötigte Präzision gegen Widerstände durchzusetzen, um Trugschlüsse zu verhindern • „Analytical Fairness“ als Grundhaltung, d.h. Bereitschaft, Analysen nicht durchzuführen, wenn das Risiko eines Missbrauchs hoch ist • Objektivität als Grundhaltung, insbesondere in Situationen, in denen Datenlage und Fragestellung Spielraum für die Analyse lassen 	<ul style="list-style-type: none"> • Kann mit grundlegenden statistischen Methoden wie etwa Mittelwert und Standardabweichung umgehen • Beherrscht und verwendet komplexere Analyseverfahren und erkennt die Grenzen der Analyseergebnisse, kann genau einschätzen, welche Methoden für welche Fragestellungen und Daten die besten Ergebnisse liefern • Beherrscht und verwendet hochgradig komplexe Modelle und erkennt, wenn die Analyse keine sinnvollen Ergebnisse liefern kann bzw. wann die Informationen der Analyse keine Relevanz für die Fragestellung haben und der Analyseprozess beendet werden sollte

- Wissen über Schätzverfahren und Algorithmen
- Wissen über mögliche Ursachen von Artefakten

Fähigkeit, messbare Beziehungen in Modellen abzubilden
 Fähigkeit, zukünftige Nutzungen der Analyseergebnisse zu antizipieren

„Analytical Fairness“ als Grundhaltung, d.h. Bereitschaft, Analysen nicht durchzuführen, wenn das Risiko eines Missbrauchs hoch ist

Beispiel D.3 Datenanalysen interpretieren (Dekodieren = rezeptive Seite)

Kompetenz		Kompetenzdimensionen			Kompetenzniveaus
Bezeichnung	Beschreibung	Beispiele für Wissen („Knowledge“)	Beispiele für Fähigkeiten („Skills“)	Beispiele für Haltung („Attitude“)	Beispiele für aufsteigende Niveaus
D.1: Daten-Analysen interpretieren	Interpretiert Datenprodukte (Statistiken, Modellergebnisse) in verbalisierter Form bzw. prüft kritisch die explizit oder implizit gelieferte Interpretation	<ul style="list-style-type: none"> Wissen über statistische Kennzahlen wie Mittelwerte, Prozente, Prozentpunkte, deren Aussagekraft und Zusammenhänge Wissen über Zusammenhänge zwischen statistischen Kennzahlen und den zugrundeliegenden Daten Wissen über den Bezug zwischen Kennzahlen (z.B. Geometrische Häufigkeiten) Wissen, dass die Wahl einer bestimmten Kennzahl das Ergebnis eines bewussten Entscheidungsprozesses sein kann Kenntnis statistischer Fachbegriffe Kenntnis statistischer Fehlschlüsse (z.B. Korrelation vs. Kausalität) Wissen, dass statistische Aussagen i.d.R. keine Einzelfallaussagen darstellen Wissen, dass die Interpretation das Hinzufügen von Kontextinformation erfordert 	<ul style="list-style-type: none"> Kann Schlüsse ziehen, über welche Charakteristika der Daten eine Kennzahl Aussagen trifft Versteht, welche Kennzahlen (auch verbalisiert) äquivalent verwendet werden Kann hinterfragen, ob spezielle Datensituationen das Ergebnis beeinflussen Kann hinterfragen, ob die Darstellungsform die Aussage eines Datenprodukts beeinflusst Kann hinterfragen, wie die Interpretation der Ergebnisse vom eigenen Kontextwissen abhängt Kann hinterfragen, inwiefern die Darstellung der Ergebnisse durch die Wahl der Form (z.B. Prozente) oder die Verbalisierung die Interpretation manipuliert beeinflusst 	<ul style="list-style-type: none"> Bereitschaft, explizit kommunizierte, vorgegebene Interpretationen in Daten-Verbalisierungen zu hinterfragen Bereitschaft, nach implizit kommunizierten Interpretationen zu suchen und diese zu hinterfragen Bereitschaft, das eigene Kontextwissen in Bezug auf dessen Einfluss auf die Interpretation zu hinterfragen Offenheit gegenüber neuen Erkenntnissen, auch wenn diese den bisherigen Überzeugungen widersprechen Bereitschaft, die Bedeutung der Ergebnisse für den Sachverhalt, in dem sie gestellt werden, zu hinterfragen 	<ul style="list-style-type: none"> Kann einfache statistische Terminologie verstehen und deren Bezug zu den Daten interpretieren, kennt grundlegende Formen der Manipulation durch Statistiken und Berichte und die zu beachtenden Kriterien Besitzt ein fortgeschrittenes Verständnis der Terminologie und kann zwischen verschiedenen Begriffen sauber differenzieren, kennt Bausteine expliziter Kommunikation und kann prüfen, ob explizite Interpretationen aus den Ergebnissen ableitbar sind Besitzt ein vertieftes Verständnis der Terminologie, kann implizite Aussagen und Interpretationen erkennen, und zueinander sowie in Bezug auf den Sachverhalt abwägen und beurteilen

- Kenntnis statistischer Fachbegriffe
- Kenntnis statistischer Fehlschlüsse (z.B. Korrelation vs. Kausalität)

- Kann Schlüsse ziehen, über welche Charakteristika der Daten eine Kennzahl Aussagen trifft
- Kann hinterfragen, inwiefern die Interpretation der Ergebnisse vom eigenen Kontextwissen abhängt

- Offenheit gegenüber neuen Erkenntnissen, auch wenn diese den bisherigen Überzeugungen widersprechen

Vorschläge für Instrumentarien

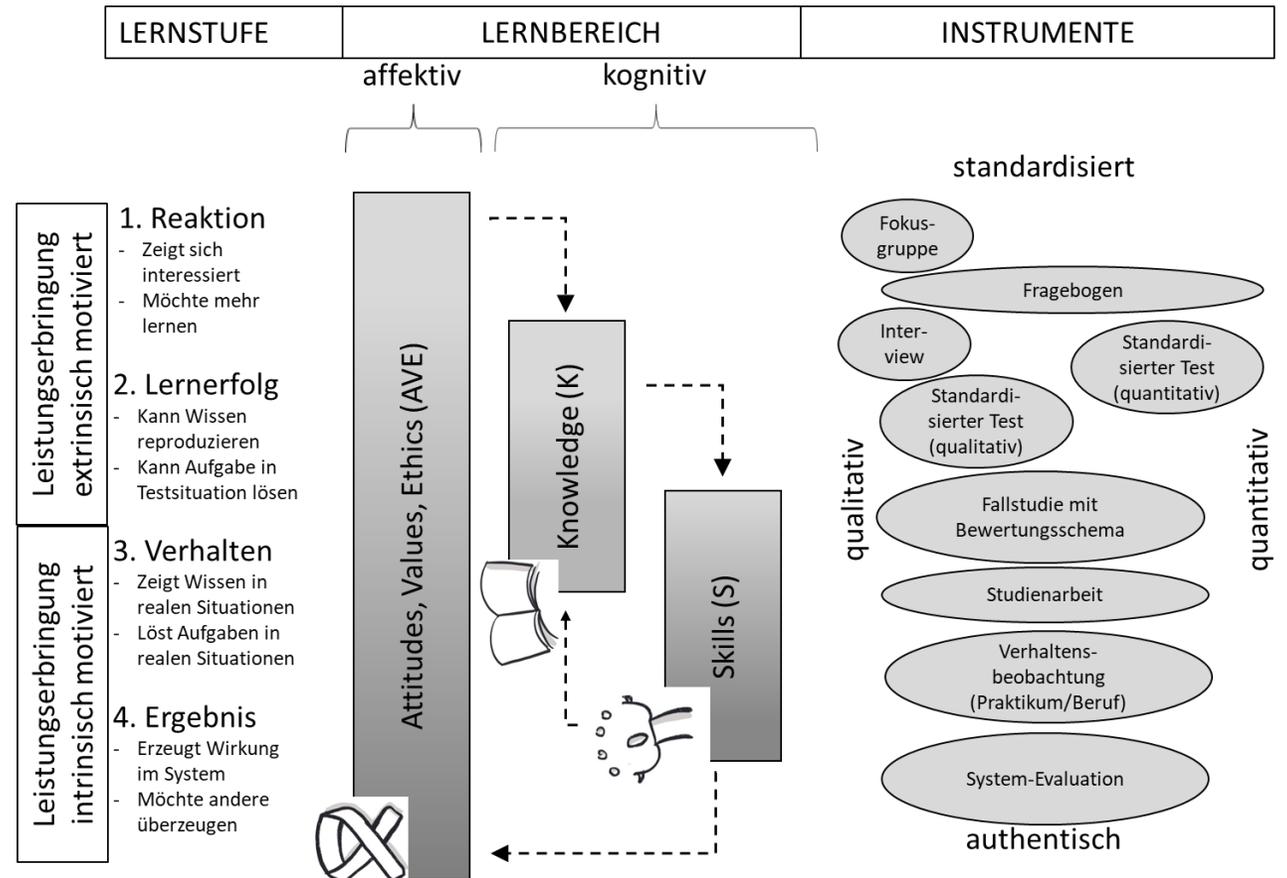
Einordnung und Impulse für die Testentwicklung

▪ Lernstufen-Modell nach Kirkpatrick

▪ **Problem:** Wie lassen sich Lernstufen messen, die gekennzeichnet sind durch eine **intrinsische Motivation** der Leistungserbringung, in einem System, das durch Noten einen **extrinsischen Motivator** einführt?

▪ Standardisierte Instrumente: **Validierung und Benchmarking** möglich

▪ Empfehlung: **Multimethodischer Ansatz.**
Kombination von (mehrstufigem) Multiple Choice und Fallstudien mit Bewertungsschema



Beispielaufgabe zu den Kompetenzen D.1 „Daten-Verbalisierungen interpretieren“ und E.1 „Standardisierung entschlüsseln“

Unter Bezugnahme auf die Allbright-Studie „Die Macht der Monokultur“ titelte das Manager-Magazin:

„Der Frauenanteil in den Vorständen steigt (um 0,7 Prozentpunkte)“. Sie erhalten folgende zusätzlichen Informationen:

1. Der Frauenanteil ist um 9,75% gestiegen.
2. Der Frauenanteil betrug zuvor 7,3%.

Frage 1: Wie groß ist der Frauenanteil in den Vorständen heute?

Frage 2: Wie viele Frauen sind heute in den Vorständen?

- A) Information 1 reicht aus, um die Frage zu beantworten.
- B) Information 2 reicht aus, um die Frage zu beantworten.
- C) 1 und 2 werden zusammen benötigt, um die Frage zu beantworten.
- D) 1 oder 2 alleine reichen jeweils aus, um die Frage zu beantworten.
- E) 1 und 2 zusammen reichen nicht aus, um die Frage zu beantworten.

Richtig sind D (Frage 1) und E (Frage 2).

Die Original-Schlagzeile lautete dabei fälschlicherweise: „Der Frauenanteil in den Vorständen steigt (um 0,7 Prozent).“

Beispielaufgabe zu den Kompetenzen C.3 „Daten verbalisieren“

Sie erhalten die Originalschlagzeile zusammen mit den Rohdaten der Studie:

- A) Ist die Aussage der Schlagzeile korrekt? Begründen Sie Ihre Antwort.
- B) Handelt es sich Ihrer Meinung nach um eine starke oder eine schwache Veränderung? Begründen Sie Ihre Antwort.
- C) (Wie) würde sich ihre Antwort zu B) verändern, wenn die Schlagzeile gelautet hätte: „Der Frauenanteil in den Vorständen steigt (um 9,75 Prozent).“
- D) Wie würden Sie die Schlagzeile formulieren, um zu signalisieren, dass Sie die Frauenquote befürworten?
- E) Wie würden Sie sie formulieren, um zu signalisieren, dass Sie die Frauenquote ablehnen?

Fachgebiet	Medizin: Diagnostik	Ingenieurwissenschaften: Produktionsüberwachung	Betriebswirtschaftslehre/Sociologie: Gender Studies	Verkehrswissenschaften: Mobilitätsplanung	Psychologie: Ethik in der Forschung
Kurzbeschreibung der Fallstudie	<p>Die Fallstudie behandelt die Frage, wie die Validität des BEHAVIOUR Scores für den Einsatz als diagnostisches Instrument geprüft werden kann.</p>	<p>Die Fallstudie beschäftigt sich mit der Problematik, Sensor-Fehlfunktionen aus Messdaten zu identifizieren. Dabei wird insbesondere die Frage thematisiert, was die Daten eigentlich messen.</p>	<p>Die Fallstudie beschäftigt sich anhand einer publizierten Studie mit der Frage, wie einfach oder wie schwierig es ist, Argumente für oder gegen eine Frauenquote empirisch zu untermauern. Sie thematisiert besonders die Notwendigkeit der Unterscheidung zwischen Daten und der Interpretation dieser Daten.</p>	<p>Die Fallstudie beschäftigt sich mit der Frage, wie die Potenziale von Daten aus neuartigen Quellen (insbesondere App-Daten) hinsichtlich ihrer Bedeutung für die Mobilitätsplanung beurteilt werden können und ob sich daraus die Entscheidung für oder gegen die Integration von Ortungsdaten – die womöglich von den Nutzern negativ aufgefasst wird – unterstützen lässt.</p>	<p>Die Fallstudie beschäftigt sich mit einer Situation, in der die Daten einer Dissertation mit exzellenten Ergebnissen „verloren“ gegangen waren und stellt zur Diskussion, welche Kompetenzen ein Betreuer braucht, um solche Fälle zu vermeiden bzw. aufzudecken.</p>
Einordnung in das Data Literacy Framework					

Fragestellung

Ein*e Psychologie-Professor*in einer Universität wendet sich an eine*n Statistik-Expert*in mit der Bitte, die Ergebnisse einer Dissertation zu überprüfen. Es handle sich um herausragende Befunde, die eine wichtige Theorie bestätigen.

„Nur dummerweise scheinen sämtliche Daten verloren. Das scheint mir (wie vielen anderen) nun nicht wirklich allzu wahrscheinlich zu sein, und die genaueren Umstände lassen es denkbar erscheinen, dass die Daten erfunden worden sein könnten.“

Nun habe ich einmal gehört, dass es ausgewiesenen Statistik-Expert*innen offenbar gelingen kann, aufgrund publizierter Korrelationstabellen, Streuungsmaße, Reliabilitäten etc. die Plausibilität dessen abzuschätzen, ob die berichteten Befunde auf echten Datentabellen beruhen oder ob sie komplett frei erfunden sein könnten.“

Vorgehen des oder der Anwender*in

Schritt	Beschreibung	Kompetenz	Prozess	Zuordnung
[1]	Er oder sie bewertet die statistischen Ergebnisse im Hinblick auf deren Bedeutung für die psychologische Forschung	Erkennen, dass die Ergebnisse einen hohen Wertbeitrag für die Forschung leisten	Dekodieren	F.1
[2]	Und entscheidet, dass die Daten so werthaltig sind, dass sie vertieft ausgewertet werden sollen.	Nächste Schritte für die Forschungstätigkeit ableiten	Decodieren	F.2
[3]	Er oder sie bewertet die Umstände (z.B. Umgang mit erhobenen Daten) anhand der Regeln wissenschaftlich sauberen Arbeitens und fragt sich, ob unsauberes Arbeiten mit derartigen Ergebnissen vereinbar ist,	Erkennen, dass der behauptete Totalverlust von Daten im Gesamtkontext fragwürdig erscheint	Dekodieren	E.2
[4]	hinterfragt deshalb die Ergebnisse weiter hinsichtlich Plausibilität solcher „Superbefunde“ und vergleicht sie mit üblichen Ergebnissen bei derartigen Experimenten.	Wissen über die Möglichkeiten, was man bei Kraft-Tests erwarten kann, und Bewerten der gefundenen Effektstärken	Dekodieren	E.3
[5]	Er oder sie ist sich im Klaren, dass ein Fall von Datenverlust seiner oder ihrer Fakultät nicht tragbar wäre.	Erkennen der möglichen Folgen eines Datenbetrugs	Decodieren	F.3
[6]	Er oder sie überlegt sich, dass ein*e Expert*in wissen könnte, wie man mutmaßliche Datenfälschungen anhand der Ergebnisse aufdecken kann.	Handlungsmöglichkeit identifizieren, um die möglichen Fälschung aufzudecken	Decodieren	F.1
[7]	Er oder sie identifiziert die nötigen Informationen, die der oder die Expert*in braucht, um die Plausibilität statistischer Ergebnisse zu prüfen,	Wissen über statistische Kennzahlen und deren Aussagekraft bzw. Schwachstellen	Dekodieren	E.1
[8]	Er oder sie entscheidet sich, eine Begutachtung durch eine*n Expert*in durchführen zu lassen,	Grenzen der eigenen Kompetenz erkennen und Expertenrat einholen	Codieren	A.3
[9]	die er oder sie inoffiziell beauftragt, um den potenziellen Schaden in Grenzen zu halten.	Bewerten, welche Folgen ein möglicherweise negatives Gutachten für die Fakultät hätte.	Decodieren	F.3

... wie passt sauberes wissenschaftliches Arbeiten dazu, dass die Daten verloren gehen? ... kann diese Art von Experimenten zu so guten Ergebnissen führen?

Vorgehen des oder der Anwender*in

Schritt	Beschreibung	Kompetenz	Prozess	Zuordnung
[10]	Der oder die Statistikexpert*in liest die Dissertation und plant sein Vorgehen.	Überlegen, welche Vorgehensweise das Problem lösen kann	Codieren	A.3
[11]	Er oder sie identifiziert mögliche Ansatzpunkte, indem er oder sie die Aussagekraft der berichteten Ergebnisse bewertet	Bewerten, welche Rückschlüsse möglich sind, weil zwischen einzelnen berichteten Kennwerten bestimmte Beziehungen bestehen	Decodieren	D.1
[12]	und extrahiert Informationen zur Methodik aus dem Fließtext	Verbale Beschreibung von Analysen entschlüsseln	Decodieren	D.3
[13]	sowie aus den Grafiken und der tabellarischen Darstellung	Visuelle Darstellung auf mehreren Ebenen entschlüsseln (inhaltlich, formal: Erkennen der Software).	Decodieren	D.2
[13]	Er oder sie überlegt, welche Einschränkungen die Darstellungsform durch einen Informationsverlust mit sich bringt	Bewerten, wie mit gerundeten Ergebnissen umzugehen ist	Decodieren	E.1
[14]	und wendet sein oder ihr Wissen über die Beziehungen zwischen berichteten Kennwerten an,	Statistische Analysen aufgrund gegebener Informationen entschlüsseln	Decodieren	D.1
[15]	indem er oder sie alle prüfbaren Beziehungen nachrechnet	Statistische Analysen mit Tertiärdaten (d.h. Sekundärdaten) durchführen	Codieren	C.1
[16]	sowie Unplausibilitäten hinsichtlich möglicher Ursachen beurteilt.	Daten-Transformationen bewerten, z.B. "Fehlern abgrenzen"	Decodieren	E.1
[17]	Er oder sie bewertet abschließend, ob eine Datenfälschung anhand des Gesamtbildes anzunehmen ist.	Überprüfen, ob die Fehler wahrscheinlicher auf unsauberes Arbeiten oder auf wissenschaftliches Fehlverhalten zurückzuführen sind.	Decodieren	E.2/E.3
[18]	Er oder sie dokumentiert sein Vorgehen und seine Ergebnisse in einem detaillierten Bericht für den oder die Anwender*in.	Vorgehen, Methodik und Ergebnisse in einem für den Anwender verständlichen Text zusammenfassen	Codieren	C.3

... in welchem Bezug stehen die berichteten Kennzahlen (deskriptive Statistiken, Teststatistiken, p-Werte)

zueinander?

... was lässt sich durch Rundungsfehler erklären, was ist

sicher falsch?

Beispiel 1: Eher einfach (Basisniveau)

Deskriptive Angaben: $n=140$, Mean = 24.8, SD = 0.34, age range 20 – 36 Jahre

Passt das zusammen?

- Faustregel bei annähernd symmetrisch verteilten Beobachtungen: Range in etwa 6x SD
- Hier: Spannweite = 16, 6x SD = 2.04
- SD minimal, wenn alle Werte außer min und max das Mean(min,max) annehmen (hier: 28)
- → SD = 0.9596162 >> 0.34 und Mean = 28 >> 24.8.

Beispiel 2: Eher anspruchsvoll (Fortgeschrittenen-Niveau)

- Unterscheidet sich das Delta der Messungen in t_1 und t_2 zwischen zwei Gruppen EG/CG?
- Angabe: $mean(t_1, t_2)$ je Gruppe, $SD(t_1, t_2)$ je Gruppe
- Daraus lassen sich nachträglich berechnen:
 - $mean(t_2 - t_1)$, $SD(t_2 - t_1)$, $cov(t_1, t_2)$, $corr(t_1, t_2)$, $beta$ der Regression t_2 auf t_1 je Gruppe

	Mean (t1)	SD (t1)	Mean (t2)	SD (t2)	Mean (t2-t1)	SD (t2-t1)	Cov(t1.t2)	Corr(t1.t2)	Sig. (approx.)	Beta (approx.)
EG (n=60)	86.36	29.94	82.66	29.52	-3.70	5.0	871.42	.99	<.001	0.97
CG (n=60)	84.38	25.42	83.46	25.60	-0.94	1.0	650.27	>.99	<.001	1.01
Gesamt (n=120)	85.37	27.77	83.06	27.63	-2.32	3.61	764.79	> .99	<.001	0.99

→ Nahezu perfekte Vorhersage von t_2 aus t_1 – plausibel für Daten aus psychologischen Experimenten?

Vielen Dank für Ihre
Aufmerksamkeit!

