

Bildungsserver:Next – Ihre Ideen für unser Bildungsportal

Innovationswettbewerb des Deutschen Bildungsservers
vom 14. April bis 9. Juni 2026

Platz 3 | Kategorie „Content“

EvidenzLabel — quality labels for education

Nicht alle Bildungsinhalte sind gleichermaßen evidenzbasiert – doch der Deutsche Bildungsserver stellt sie derzeit so dar, als wären sie es.



© KI-generiertes Bild Stefan Fischer/Hyve

Zusammenfassung

Mit „EvidenzLabel“ wird auf dem Deutschen Bildungsserver ein fünfstufiges Kennzeichnungssystem für Evidenz eingeführt. Jede verlinkte Ressource erhält eine Kennzeichnung, die ihren epistemischen Status angibt: empirisch belegt (E1), wissenschaftlich plausibel (E2), praxisbasiert (E3), normativ/wertebasiert (E4) oder noch nicht bewertet (E?). Ein Evidenz-Index fasst diese Kennzeichnungen auf Dossier-Ebene

zu einem Punktwert von 0 bis 100 zusammen und liefert Nutzern so auf einen Blick einen Qualitätsindikator für ein gesamtes Themengebiet. Das System wird als einzelnes Metadatenfeld in die bestehende Datenbank integriert – es sind weder ein neues CMS noch eine neue Infrastruktur erforderlich. Die Umsetzung erfolgt schrittweise: Neue Inhalte werden sofort gekennzeichnet, bereits bestehende, stark frequentierte Dossiers nachträglich; zudem ist ein moderierter Mechanismus für Beiträge aus der Fachcommunity geplant. Dank seiner Expertise in Redaktion und Forschung ist das DIPF in der idealen Position, dieses System aufzubauen und zu pflegen. Bisher gibt es in Deutschland kein Äquivalent zum „Toolkit“ der britischen Education Endowment Foundation oder zum US-amerikanischen „What Works Clearinghouse“. EvidenzLabel schließt diese Lücke.

Beschreibung

Das Problem

Eine Lehrkraft, die auf einem Bildungsportal nach Methoden für den Leseunterricht sucht, stößt auf Links zu fundierten Meta-Analysen, die direkt neben ideologisch geprägten Schriften und traditionellen Praxisleitfäden stehen – und optisch gleichwertig präsentiert werden. In einer Zeit, die von KI-generierten Inhalten und Bildungs-Influencern geprägt ist, wiegt das Fehlen eines solchen erkenntnistheoretischen Kontexts schwerer denn je.

Das Problem ist strukturell tief verwurzelt. Nicht alle pädagogischen Inhalte sind auch nur im Prinzip falsifizierbar. Manche Methoden wurden in kontrollierten Studien überprüft; andere entstammen Traditionen der Unterrichtspraxis, in denen randomisierte Studien methodisch gar nicht möglich sind; wieder andere basieren auf normativen politischen Überzeugungen, bei denen die Frage „Funktioniert das?“ schlicht am Thema vorbeigeht. All diese Ansätze als gleichartige Aussagen zu behandeln, stellt für sich genommen bereits eine Verzerrung der Realität dar.

Die fünf Evidenzstufen

E1 — Empirisch belegt. Mehrere randomisierte kontrollierte Studien (RCTs) oder Metaanalysen bestätigen die Wirkung. Über verschiedene Kontexte hinweg reproduzierbar. Beispiele: Spaced-Retrieval-Praxis (verteilt wiederholen), formatives Feedback, expliziter Phonetikunterricht.

E2 — Wissenschaftlich plausibel. Fundierte theoretische Basis; gewisse empirische Unterstützung, wobei die Studienlage jedoch begrenzt oder uneinheitlich ist. Beispiele: Projektbasiertes Lernen, Flipped Classroom.

E3 — Praxisbasiert. Basiert auf der Erfahrung von Praktikern oder auf etablierter Tradition. Eine systematische Überprüfung hat nicht stattgefunden oder ist methodisch schwierig. Beispiele: Waldorfpädagogik, Montessori-Methoden.

E4 — Normativ. Die Inhalte spiegeln pädagogische Werte oder politische Zielsetzungen wider. Eine empirische Überprüfung ist ihrer Natur nach nicht anwendbar. Beispiele: Demokratische Bildung, inklusiver Unterricht als Wert.

E? — Noch nicht bewertet. Keine Einstufung vorgenommen. Redaktionelle Prüfung steht noch aus. Diese Kennzeichnung macht gezielt auf Lücken in der Evidenzlage aufmerksam und lenkt die Aufmerksamkeit der Forschung auf bisher wenig untersuchte Bereiche.

Ein entscheidendes Gestaltungsprinzip: Das System stellt keine Rangfolge nach Wertigkeit dar. E4 ist E1 nicht unterlegen – die Kategorie legt lediglich offen, welcher Art sie ist. Die Bezeichnungen dienen der Kontextualisierung, nicht der Abwertung.

Der Evidenz-Index

Auf Dossier-Ebene generiert EvidenzLabel einen numerischen Qualitätsindex (0–100), indem es den gewichteten Durchschnitt der Evidenzstufen aller verknüpften Ressourcen ermittelt. Ein Dossier mit einem Index von 80 signalisiert der Lehrkraft, dass sie dort vorwiegend empirisch evaluierte Inhalte vorfindet. Ein Index von 25 weist hingegen auf einen Themenbereich hin, der eher auf Praxisberichten oder normativen Ansätzen basiert. Zudem dient der Index als redaktionelles Instrument, um systematisch aufzuzeigen, wo Evidenzlücken bestehen.

Umsetzung

Das Kennzeichnungssystem fungiert als Metadatenebene auf der bestehenden Datenbank – mit einem neuen Feld pro Eintrag, einer neuen Anzeigekomponente (Badge und Tooltip) sowie einem neuen Suchfilter. Es ist kein neues CMS erforderlich. Die Umsetzung erfolgt in drei Phasen:

Phase 1 (Monate 1–6): Neue Inhalte werden im Rahmen des regulären redaktionellen Workflows gekennzeichnet. Dies verursacht nahezu keine Kosten.

Phase 2 (Monate 6–18): Nachträgliche Kennzeichnung bereits vorhandener Dossiers mit hoher Zugriffszahl; Forschende des DIPF steuern hierfür Einschätzungen zu ihren jeweiligen Fachgebieten bei.

Phase 3 (ab Monat 18): Einführung eines moderierten Mechanismus für Community-Beiträge, der es verifizierten Lehrkräften und Forschenden ermöglicht, Änderungen der Einstufung vorzuschlagen (vorbehaltlich der redaktionellen Freigabe).

Verantwortungsvolle Gestaltung

Keine Einstufung wird als überlegen dargestellt. Die „E4“-Kennzeichnung wird konsistent über das gesamte politische Spektrum hinweg angewandt. Eine kommerzielle Einflussnahme auf die Einstufung ist ausgeschlossen. Es werden keine Nutzerdaten erhoben. Sämtliche Definitionen der Einstufungen werden öffentlich zugänglich gemacht.

Internationale Vorbilder

Das Konzept stützt sich auf das „Teaching and Learning Toolkit“ der britischen Education Endowment Foundation, das US-amerikanische „What Works Clearinghouse“ sowie das GRADE-System aus der evidenzbasierten Medizin. In Deutschland existiert derzeit keine vergleichbare Ressource. Mit „EvidenzLabel“ würde sich der Bildungsserver als Vorreiter einer evidenzinformierten Bildungsinfrastruktur positionieren – eine logische Erweiterung des Forschungsauftrags des DIPF.