



Im Beitrag zeigt der Autor Brenzau, dass die Arbeit mit PIAAC-Daten anspruchsvoll ist und das ständige Wechseln zwischen vier Dokumenten zu Verwirrung führt. Deshalb hat er ein Werkzeug "The PIAAC Variable Finder" entwickelt, das alle relevanten Metadaten auffindbar verknüpft.

Schlagworte: Bildungsforschung; Metadaten; Datenanalyse; Weiterbildung; Weiterbildungsangebot

E-Journal Einzelbeitrag
von: Nate Brenzau

Weniger Frustration für Forschende bei der Suche nach Variablen und Metadaten

aus: Grundbildung & Grundkompetenzen (WBDIE2503W)
Erscheinungsjahr: 2025
Seite: 32
DOI: 10.3278/WBDIE2503W011

Weniger Frustration für Forschende bei der Suche nach Variablen und Metadaten Der PIAAC Variable Finder

NATE BREZNAU

Die Arbeit mit PIAAC-Daten ist anspruchsvoll: Die OECD und nationale Partner haben zwar viel investiert, um die Daten möglichst zugänglich und im Sinne der FAIR-Prinzipien¹ nutzbar zu machen, u. a. durch Bereitstellung der Basisdaten im csv-Format. Zum Import von Daten und Metadaten stehen zudem – in begrenztem Umfang – unterstützende Skripte für Stata und R zur Verfügung. Wer jedoch alle Aspekte einer gegebenen PIAAC-Variable über beide Zyklen hinweg verstehen will, benötigt mindestens vier Dokumententypen: die *Fragebögen*, die *Variablenlabels*, die *Trenddokumentation* und die *Derived Variables Codebooks*. Zusätzlich ist es hilfreich zu wissen, welche Variablen im *Public Use File* (PUF) tatsächlich enthalten sind – denn es kostet Forschende viel Zeit und Energie, Metadaten zu einer Variable zu recherchieren, nur um dann festzustellen, dass aus Datenschutzgründen gar keine Daten vorliegen.

Aus eigener Arbeit mit PIAAC weiß ich: Ständig zwischen vier Dokumenten (teils je Zyklus doppelt) hin- und herzuschalten, führt zu Verwirrung. Deshalb habe ich ein Werkzeug entwickelt, das alle relevanten Metadaten auffindbar verknüpft. Das verbesserte meinen Workflow so deutlich, dass ich die Lösung als Open-Access-Anwendung aufbereitet habe: den PIAAC Variable Finder. Mit der App lässt sich schnell prüfen, ob und wie die eigenen Fragestellungen in PIAAC abbildbar sind. Wer sich beispielsweise für Sprachkurse interessiert, könnte nach »Sprachkurs« suchen und feststellen, dass PIAAC keine Frage enthält, die eine spezifische Teilnahme an Sprachkursen ausweist. Sucht man jedoch nach »Fremdsprachen« oder »sprachlich«, findet man die Frage »Wo lag der Schwerpunkt dieser Weiterbildung? Bitte nennen Sie nur einen«, bei der »Fremdsprachen« auswählbar ist. Die entsprechende Ausprägung ist in der Variable *B2_Q10* enthalten. So lässt sich ermitteln, ob der zuletzt besuchte Weiterbildungskurs ein Sprachkurs war (Fremdsprache).

Die App ist webbasiert nutzbar oder kann lokal in RStudio mit dem Paket *shiny* ausgeführt werden.² Sie basiert auf der deutschsprachigen (deutscher Länderfragebogen) und der englischsprachigen Dokumentation (Vereinigtes Königreich). Die Suche ist per Stichwort oder Variablenname möglich; diverse Filter stehen bereit. Ausgegeben werden u. a. die Ursprungsfragen, die Konstruktionslogik (Formel) *Derived Variables*, Kodierungen und Labels, Begründungen dafür, warum eine Variable keine Trendvariable ist, die Unterscheidung zwischen *Strict* und *Soft Trend Variables* sowie Informationen, ob eine Variable im zweiten PIAAC-Zyklus neu ist oder entfiel. Außerdem zeigt die App an, ob eine Variable in den PUF-Daten enthalten ist. Durch den offenen Quellcode entspricht sie vollumfänglich den Prinzipien der *Open Science*; Weiterentwicklungen sowie Fehlermeldungen sind – etwa über das SocArXiv-Preprint (Brezna, 2025) und das GitHub-Repository – ausdrücklich erwünscht. Das Tool entstand aus der Praxis für die Praxis und steht allen Forschenden kostenfrei zur Verfügung. Eine ausführlichere Darstellung in der Zeitschrift *Analysis Scripts in Large-Scale Assessments in Education* ist für 2026 geplant.

NATE BREZNAU

ist Wissenschaftlicher Mitarbeiter beim
Deutschen Institut für Erwachsenenbildung in der
Abteilung Organisation und Programmplanung.

nathan.brezna@die-bonn.de

Brezna, N. (2025). *The PIAAC Variable Finder: An interactive Shiny app for cleaning, interpreting and analyzing Programme for the International Assessment of Adult Competencies data*. https://doi.org/10.31235/osf.io/xf4ja_v1

¹ Findable, Accessible, Interoperable and Reusable (auf Englisch)

² https://nate-brezna.shinyapps.io/PIAAC_Variable_Finder/