

# QUADRIGA-Fallstudien zur Datenkompetenz in den Humanities: Jupyter Books als “scalable Open Educational Resources”

## Sluyter-Gäthje, Henny

henny.sluyter-gaethje@uni-potsdam.de  
Universität Potsdam, Deutschland  
ORCID: 0000-0003-2969-3237

## Skorinkin, Daniil

daniil.skorinkin@uni-potsdam.de  
Universität Potsdam, Deutschland  
ORCID: 0000-0002-1845-9974

## Trilcke, Peer

peer.trilcke@uni-potsdam.de  
Universität Potsdam, Deutschland  
ORCID: 0000-0002-1421-4320

## Chlastak, Maria

maria.chlastak@gi.de  
Gesellschaft für Informatik, Deutschland  
ORCID: 0000-0003-1046-3634

## Samoilova, Evgenia

evgenia.samoilova@uni-potsdam.de  
Universität Potsdam, Deutschland  
ORCID: 0000-0003-3858-901X

## Seltmann, Melanie

melanie.seltmann.1@hu-berlin.de  
Humboldt-Universität zu Berlin, Deutschland  
ORCID: 0000-0002-7588-4395

## Einleitung

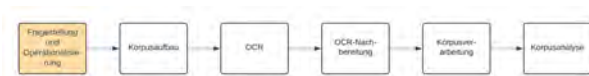
Datenkompetenz, die sowohl den kritischen Umgang mit Daten (von Sammeln über Verwalten bis hin zu Bewerten) als auch die Analyse und Interpretation von Daten (vgl.

Ridsdale et al., 2015) umfasst, wird auch für traditionell ausgebildete Geisteswissenschaftler:innen zunehmend wichtiger. Im Berlin-Brandenburgischen Datenkompetenzzentrum QUADRIGA<sup>1</sup> (Buchholz et al., 2024) werden deshalb fallstudienbasierte Open Educational Resources (OER) entwickelt, die in Form von Jupyter Books (Exe-

cutable Books Community, 2020) zur Verfügung stehen und sowohl die Möglichkeit der individuellen Lerndauer als auch der Interaktivität und des Assessments bieten (vgl. z.B. Chen & Asta, 2022 oder Betlem et al., 2024). Eine detaillierte Darstellung des Vorgehens zur Entwicklung der OER unternehmen Samoilova et al. (2024). Das Poster legt seinen Fokus auf die Potentiale von Jupyter Books als OER.

## Fallstudienbasierte Open Educational Resource

Um den verschiedenen Dimensionen der Kompetenzvermittlung gerecht zu werden, die sich von basalen technischen Fähigkeiten über methodenspezifisches Wissen bis hin zur Daten- und Methodenkritik erstrecken, orientieren sich die von QUADRIGA entwickelten OER an Fallstudien (Foran, 2001) und setzen so problembasiertes Lernen (Kay et al., 2000) ein. Das heißt, QUADRIGA geht von einer konkreten Fragestellung aus<sup>2</sup>. Die sich aus der Fragestellung ergebende Forschungsstudie wird basierend auf workflow-basiertem Lernen (Mischke et al., 2021) und dem Vier-Komponenten-Instructional-Design-Modell (4C/ID) (van Merriënboer, 1997) in kleinere Aufgaben zerlegt (Abb. 01). Sie dienen als Basis für die Entwicklung von Lernzielen.



Aufgaben innerhalb der Fallstudie "Spanische Grippe".

Jede dieser Aufgaben ist in strukturell identische Komponenten unterteilt:

1. Präsentation der Aufgabe
2. Unterstützende Informationen
3. Interaktive Übungen der Lernaufgaben/Probleme
4. Reflexion
5. Prozedurale Informationen zu Routineaspekten der Lernaufgaben
6. Assessment

## Jupyter Books als Open Educational Resource

### Struktur

Jupyter-Book<sup>3</sup> ist eine Python-Bibliothek, die es erlaubt, aus einer Sammlung von Markdown-Dateien und Jupyter Notebooks eine Website zu erstellen (Jupyter Book genannt). Im Fall von QUADRIGA liegen die Dateien des Books auf GitHub, die Website der Fallstudie wird über GitHub-Pages gehostet und mittels einer GitHub-Action de-

ployed. Dies ist sowohl in der Erstellung als auch in der Anwendung von Vorteil, da die Inhalte schnell geupdated und Fehler sowie Feedback zur Fallstudie über GitHub-Issues gemeldet werden können.

Die Struktur des Books wird über ein Inhaltsverzeichnis hierarchisch festgelegt. Mit Hilfe dieser Struktur lassen sich die Aufgaben und ihre Komponenten gut abbilden. Wie auf der Landing-Page (Abb. 02) zu sehen, besteht die Website aus drei vertikalen Abschnitten.

Landing-Page der Fallstudie "Spanische Grippe".

Dadurch kann die Fallstudie zum einen chronologisch durchlaufen werden, zum anderen können einzelne Aufgaben und Komponenten direkt angesteuert werden. So kann selbst ein Lernpfad im Sinne des aktiven Lernens (Markant, 2016) gewählt werden. Aufgaben, Komponenten und ihre Bestandteile können über URLs referenziert werden, was die Modularität des Books weiter erhöht.

## Elemente

Die interaktiven Komponenten werden in Jupyter Notebooks erstellt, in denen sowohl Markdown als auch Python-Code geschrieben werden kann. So können kleine UI-Elemente erzeugt werden wie die Worteingabe in Abb. 03. Zusätzlich kann das gesamte Jupyter Notebook über Google Colab<sup>4</sup> im Browser ausgeführt werden.

Interaktive Elemente im Jupyter Notebook.

Markdown ermöglicht eine multimodale Gestaltung (durch Tabellen, Abbildungen, GIFs, Querverweise etc.), bibliographische Angaben werden über BibTeX erzeugt. Mittels der MyST (Markedly Structured Text) Markdown-Erweiterung lassen sich weitere UI-Elemente hinzufügen, z.B. Informationsboxen (Abb. 04). Diese Strukturierungsmöglichkeit macht Informationen leicht auffindbar (scannable) und Inhalte schnell überfliegbar (skimmable).

## Fragestellung

Dass Pandemien ihre je eigenen Verlaufsprofile entwickeln, konnte Zeitgenoss:innen vor wenigen Jahren am Verlauf der COVID19-Pandemie miterleben. Schnell wurde die Rede von den "Wellen" der Corona-Pandemie sprichwörtlich. Auch während der Spanischen Grippe wurde bereits von "Wellen" bzw. von "Waves" gesprochen und diese spezifische Verlaufsform graphisch dokumentiert (siehe Fig. 1).

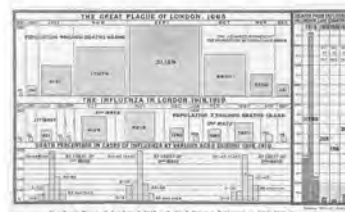


Fig. 3 Vergleichende Grafik zum Verlauf der Pest (1665) und der Spanischen Grippe (1918/19) in London. Quelle: [Staveley-Wadham, n.d.]

Aus Perspektive einer medienwissenschaftlich informierten historischen Epidemiologie überlagern sich dabei unterschiedliche Wellen: Während 'Fallzahlen' (z.B. die Anzahl der Infektionen oder die Anzahl der Todesfälle), wie sie in Fig. 1 dargestellt werden, ein in erster Linie medizinisch zu erhebendes Maß sind, sind die unterschiedliche Intensität und Extensität der öffentlichen Wahrnehmung des pandemischen Geschehens ein Untersuchungsgegenstand der (historischen) Medienwissenschaft. Auch hier ist davon auszugehen, dass sich charakteristische "Medienzeiten" oder eben: "Medienwellen" zeigen: Mal dominiert die Pandemie den öffentlichen Diskurs, mal tritt sie in der öffentlichen Wahrnehmung zurück.

Diese Fallstudie führt durch eine Digital Humanities-Forschungsprojekt, das aus medienhistorischer Perspektive den Verlauf der "Medienwellen" während der Spanischen Grippe 1918/19 – bekannt auch als "the Mother of All Pandemics" [Taubenberger and Morens, 2006] – in Berlin und Brandenburg untersucht hat. Das Forschungsprojekt wurde von der folgenden Forschungsfrage geleitet:

Multimodale Elemente in der Markdown-Datei zur Fragestellung.

## Potentiale von Jupyter Book-basierten OER: Modularisierung, Skalierbarkeit

Die Realisation als Jupyter Book ermöglicht es, eine Modularisierung der Fallstudie vorzunehmen und so atomisierte Einheiten zu spezifizierten Lernzielen aus dem Feld des Datenkompetenzerwerbs zu erstellen und referenzierbar zu machen. Darüber hinaus birgt das technische Konzept von Jupyter Books die Möglichkeit, etwas zu entwickeln, was wir "scalable OER" nennen. Die dargestellte Fallstudie kann nämlich auf verschiedenen "levels of expertise" durchgeführt werden:

- als traditionelles Buch (ohne Interaktion)
- mit interaktiven Elementen auf der Website
- durch die Öffnung eines Jupyter Notebooks via Binder oder Google Colab im Browser (ermöglicht individuelle Anpassungen und Erweiterung der Studie)
- mit dem Herunterladen des Books oder der Notebooks (ermöglicht die lokale Ausführung der Notebooks und die Anwendung auf eigenen Daten)

Die Modularisierbarkeit sowie die flexible Skalierbarkeit von Jupyter Books bieten das Potential, eine Vielzahl von Lernenden mit unterschiedlichem Vorwissen zu erreichen.

Fördervermerk

QUADRIGA wird im Rahmen der Richtlinie Förderung von Projekten zum Aufbau von Datenkompetenzzentren in der Wissenschaft des Bundesministeriums für Bildung und Forschung unter Kennzeichen 16DKZ2034A, 16DKZ2034G und 16DKZ2034H gefördert.

## Fußnoten

1. URL: <https://www.quadriga-dk.de/> (letzter Zugriff am 14.11.2024)
2. Vgl. dazu die textwissenschaftliche Fallstudie: <https://dh-network.github.io/quadriga/markdown/intro.html>, Stand: 28.11.2024
3. URL: <https://pypi.org/project/jupyter-book/> (letzter Zugriff am 15.07.2024)
4. Da sich andere Ausführungsoptionen wie mybinder.org oder Live Coding über thebe (<https://thebe.readthedocs.io/>) momentan als unzuverlässig zeigen, steht nur Google Colab zur Verfügung.

## Bibliographie

**Betlem, Peter, Nil Rodes, Sara Mollie Cohen und Marie Vander Kloet.** 2024. "Jupyter Book as an open online teaching environment in the geosciences: Lessons learned from Geo-SfM and Geo-UAV". Preprint. <https://doi.org/10.5194/gc-2024-6>.

**Buchholz, Bettina, Ulrike Lucke, Sonja Schimmler, Jan-Hendrik Bakels, Jan Bernoth, Frank Fischer, Lena Gieseke, Matthias Grotkopp, Torsten Hiltmann, Daniel Krupka, Skadi Loist, Julia Meisner, Dennis Mischke, Robin Möser, Heike Neuroth, Vivien Petras, Juliane Schmeling, Sibylle Söring und Peer Trilcke.** 2024. "Umsetzungskonzept QUADRIGA: Berlin-Brandenburgisches Datenkompetenzzentrum für Digital Humanities, Verwaltungswissenschaft, Informatik und Informationswissenschaft". Zenodo. <https://doi.org/10.5281/zenodo.10805016>.

**Chen, Enze und Mark Asta.** 2022. "Using Jupyter Tools to Design an Interactive Textbook to Guide Undergraduate Research in Materials Informatics". *Journal of Chemical Education* 99(10): 3601–3606. <https://doi.org/10.1021/acs.jchemed.2c00640>.

**Executable Books Community.** 2020. "Jupyter Book (v0.10)". Zenodo. <https://doi.org/10.5281/zenodo.4539666>.

**Foran, John.** 2001. "The Case Method and the Interactive Classroom". *Thought & Action* 17(1): 41–50.

**Kay, Judy, Michael Barg, Alan Fekete, Tony Greening, Owen Hollands, Jeffrey H. Kingston, und Kate Crawford.** 2000. "Problem-Based Learning for Foundation Computer Science Courses". *Computer Science Education* 10(2): 109–128. [https://doi.org/10.1076/0899-3408\(200008\)10:2;1-C;FT109](https://doi.org/10.1076/0899-3408(200008)10:2;1-C;FT109).

**Markant, Douglas B., Azzurra Ruggeri, Todd M. Gureckis und Fei Xu.** 2016. "Enhanced Memory as a Common Effect of Active Learning". *Mind, Brain, and Education* 10 (3): 142–152. <https://doi.org/10.1111/mbe.12117>.

**Mischke, Dennis, Peer Trilcke und Henny Sluyter-Gäthje.** 2021. "Workflow-basiertes Lernen in den Geisteswissenschaften: digitale Kompetenzen forschungsnah vermitteln". In *Bildung in der digitalen Transformation*, hg. von Heinz-Werner Wollersheim, Marios Karapanos und Norbert Pengel, 190–195. Münster, New York: Waxmann. <https://www.waxmann.com/index.php?eID=download&buchnr=4456>, letzter Zugriff: 28.11.2024.

**Samoilova, Evgenia, Henny Sluyter-Gäthje, Daniil Skorinkin, Hannes Schnaitter, Peer Trilcke und Ulrike Lucke.** 2024. "Forschungsgeleitete Vermittlung von Datenkompetenz. Mediendidaktische Aufbereitung von Fallstudien zu Bildungsangeboten." *INFORMATIK 2024*: 2153–2166. [https://doi.org/10.18420/inf2024\\_187](https://doi.org/10.18420/inf2024_187).

**van Merriënboer, Jeroen J. G.** 1997. *Training complex cognitive skills. A four-component instructional design model for technical training*. Educational Technology Publ, Englewood Cliffs.

**Ridsdale, Chantel, James Rothwell, Mike Smit, Hossam Ali-Hassan, Michael Bliemel, Dean Irvine, Daniel Kelley, Stan Matwin und Brad Wuetherick.** 2015. "Strategies and Best Practices for Data Literacy Education Knowledge Synthesis Report". Dalhousie University. <https://doi.org/10.13140/RG.2.1.1922.5044>.