

# Die digitale Schulbuch-Bibliothek GEI-Digital im neuen Gewand: Ein modernes Präsentationssystem öffnet digitalisierte Schulbücher für die Open Humanities

## Klaes, Jan Sebastian

klaes@leibniz-gei.de  
Leibniz-Institut für Bildungsmedien | Georg Eckert-Institut, Deutschland

## Krüger, Katharina

katharina.krueger@leibniz-gei.de  
Leibniz-Institut für Bildungsmedien | Georg Eckert-Institut, Deutschland

## Leonhardt, Susann

leonhardt@leibniz-gei.de  
Leibniz-Institut für Bildungsmedien | Georg Eckert-Institut, Deutschland

## Nieländer, Maret

nielaender@leibniz-gei.de  
Leibniz-Institut für Bildungsmedien | Georg Eckert-Institut, Deutschland

## Sommer, Kai

sommer@leibniz-gei.de  
Leibniz-Institut für Bildungsmedien | Georg Eckert-Institut, Deutschland

## Towara, Nadine

towara@leibniz-gei.de  
Leibniz-Institut für Bildungsmedien | Georg Eckert-Institut, Deutschland

lien- und Lesebücher („GEI-Digital“ unter <https://gei-digital.gei.de/viewer/index/>).

Der Digitalisierungsprozess umfasst dabei die Herstellung hochauflösender digitaler Images, die Erzeugung von Meta- und Strukturdaten, die Volltextfassung, die Präsentation der Digitalisate auf der GEI-Digital-Website und Maßnahmen zur Langzeitsicherung. Gegenwärtig liegt der Fokus von GEI-Digital auf deutschsprachigen Werken, die in der Zeit zwischen dem ersten Aufkommen des Schulbuches im 16. Jahrhundert bis zum Ende des Kaiserreiches im Jahre 1918 erschienen sind.

Die aktuell verfügbaren Sammlungen bzw. Teilkorpora auf GEI-Digital stellen eine Auswahl an historischen Schulbüchern dar, die als repräsentativ für das jeweilige Fach, die einzelnen Schulstufen und -formen und regional unterschiedlichen Ausgaben gelten kann. Im Hinblick auf die hohe Schulbuchproduktion im Deutschen Reich und die Bedeutung der Epoche für die Konstruktion einer nationalen Identität wird das digitale Korpus fachspezifisch und zeitlich in separaten Kollektionen auf GEI-Digital präsentiert. Internationale historische Schulbücher sind punktuell vorhanden und werden perspektivisch weiter ausgebaut (Hertling und Klaes 2018a, 21-44 und Hertling und Klaes 2018b 45-68).

Digitalisiert und unter einer Public Domain-Lizenz zur freien Nachnutzung zur Verfügung gestellt, werden nach dem geltenden Urheberrechtsgesetz (UrhG) bislang nur Werke, deren Autoren seit mehr als 70 Jahren verstorben sind. Derzeit beinhaltet GEI-Digital circa 7.400 Schulbücher mit über 1,8 Millionen Seiten. Davon konnten bereits 1,4 Millionen einer Volltexterkennung zugeführt werden. Zu den nicht im Volltext vorliegenden Beständen zählen aufgrund der überwiegend grafischen Darstellungen, Atlanten, sog. Schreiblesefibeln, und Bestände in Frakturschrift mit einem Erscheinungsjahr vor 1800. Die Implementierung einer Volltexterkennung insbesondere für ältere Frakturschriften bei mittelgroßen Sammlungen ist Gegenstand eines von der DFG geförderten Verbundprojekts „OCR-D OCR4all-libraries“ (Engl et.al. 2020 und „DFG – GEPRIS“ o.J.). Dabei wird in Zusammenarbeit mit dem Zentrum für Philologie und Digitalität (ZPD) der Universität Würzburg das Open-Source-Werkzeug OCR4all (Reul et. al. 2019) so erweitert und angepasst, dass es von Bibliotheken und Archiven bei der Digitalisierung größerer Mengen eingesetzt werden kann.

In den letzten 10 Jahren haben sich die Anforderungen an das Design und Präsentation digitaler Inhalte massiv geändert. Vor dem Hintergrund hat sich das GEI entschieden den Web-Auftritt von GEI-Digital grundlegend zu erneuern und zu optimieren. Die Präsentation der Inhalte erfolgte bisher auf Basis des von der Firma IntraData und unter der Open Source-Lizenz stehenden Softwarelösung „Goobi Workflow“ und „Goobi Viewer“, die am GEI seit 2012 unverändert eingesetzt wird (Hankiewicz 2019, 77-88).

## Ausgangslage:

GEI-Digital startete als Digitalisierungsprojekt im Jahr 2009 und bietet Zugang zu digitalisierten Schulbüchern zahlreicher Fächer, Epochen und Regionen. Das Schulbuchkorpus umfasst die Themengebiete Geschichte, Geographie und Politik sowie Fibeln, Atlanten, Rea-

## Vorgehen:

Bereits 2014 wurde eine umfangreiche Nutzer:innenbefragung durchgeführt, die wichtige Bedarfe sichtbar machte. In der Folge wurden gezielte Feedbackrunden mit Forschenden in den Digital Humanities etabliert, um die sichtbaren Bedarfe zu konkretisieren. Unter Berück-

sichtigung dieser Ausgangslage begann am GEI eine interdisziplinäre Arbeitsgruppe aus den forschenden Abteilungen und der Forschungsbibliothek einen Katalog mit Optimierungsanforderungen zusammen zu stellen. Dabei kristallisierten sich Volltextverfügbarkeit, Sammlungsverwaltung und Schnittstellen für die Datenbereitstellung als zukünftige Kernbedarfe dieser Zielgruppe heraus.

Allerdings erforderte die 2020 veraltete Softwareinfrastruktur des Goobi Viewers als Fundament für die Realisierung der Kernbedarfe ein umfassendes Upgrade. Um aber den weiterentwickelten Goobi Viewer auch für andere Anwenderinstitutionen möglichst einfach nachnutzen zu können, hat das GEI zusammen mit dem Entwicklerteam von IntraData entschieden, diese Software in Form einer unter Open Source laufenden Containervirtualisierung, auch Docker genannt, zur Verfügung zu stellen. Docker erlaubt eine vereinfachte Bereitstellung von Software-Anwendung, da alle relevanten Softwarepakete enthalten und somit schnell für eine Installation bereitstehen. Die gesamte Goobi Viewer-Software steht jetzt als Docker-Anwendung unter GitHub zur Nachnutzung bereit (Goobi-Viewer-Docker o.J.). Eng verbunden mit dem Implementierungsprozess war auch eine starke Fokussierung auf die Lokalisierung, Eingrenzung und Behebung von Problemen und Fehlern. Die interdisziplinäre Arbeitsgruppe konnte gerade im Sinne eines agilen Vorgehens schnell und effektiv Herausforderungen angehen und Lösungsansätze erarbeiten.

## Zusammenfassung und Ausblick:

Das abgestufte Vorgehen mit einer Nutzer:innenbefragung, gezielten Feedbackrunden mit Forschenden und einer interdisziplinären Arbeitsgruppe erwies sich in dem hier vorliegenden Szenario als gewinnbringend. Mit Abschluss der Phase eines ersten technischen Upgrades und einer umfassenden Erneuerung der Software-Infrastruktur steht die digitale Schulbuchbibliothek GEI-Digital in einem neuen Outfit für die Nutzer:innen in den Digital Humanities zur Verfügung. Gerade die interdisziplinäre Zusammenarbeit von Akteur:innen aus den Forschungsinfrastrukturen und den wissenschaftlichen Anwender:innen konnte einen erfolgreichen Neustart sicherstellen und die nachhaltige Grundlage für weitere Optimierungen legen.

Der kontinuierliche Betrieb einer digitalen Bibliothek und deren Optimierung sollte als eine Daueraufgabe verstanden werden. Vor dem Hintergrund ist vorgesehen, die Nutzungsmöglichkeiten von GEI-Digital für die vielfältigen Bedarfe weiter zu verbessern. Zu den nächsten Schritten zählen die Bereitstellung nachnutzbarer Datenformate nach aktuellen Standards und die Persistenz der Daten im Hinblick auf z.B. die verschiedenen Volltextversionen bei Anwendung verschiedener OCR-Engines im Kontext von Forschungsdaten.

## Bibliographie

„DFG - GEPRIS - OCR4all-libraries - Volltexterkennung historischer Sammlungen“. o. J. Zugriffen 14. Dezember 2022. <https://gepris.dfg.de/gepris/projekt/460665940?context=projekt&task=showDetail&id=460665940&>.

Engl, Elisabeth, Konstantin Baierer, Matthias Boenig, Volker Hartmann, und Clemens Neudecker. 2020. „Volltexte – die Zukunft alter Drucke: Bericht zum Abschlussworkshop des OCR-D-Projekts“. *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* 7 (2): 1–4. <https://doi.org/10.5282/o-bib/5600>.

„GEI-Digital“. 2022. Zugriffen 10. Dezember 2022. <https://gei-digital.gei.de/viewer/index/>.

Goobi-Viewer-Docker. (2022), Abgerufen 3. August 2022, von <https://github.com/intranda/goobi-viewer-docker>

Hankiewicz, Steffen. 2018. „Goobi entwickeln – Eine Open-Source Software zur Verwaltung von Workflows in Digitalisierungsprojekten“. In *Digitalisierung in Bibliotheken*, herausgegeben von Gregor Neuböck, 77–88. De Gruyter. <https://doi.org/10.1515/9783110501094-006>.

Hertling, Anke, und Sebastian Klaes. 2018a. „Historische Schulbücher als digitales Korpus für die Forschung: Auswahl und Aufbau einer digitalen Schulbuchbibliothek“. In *Digital Humanities in der internationalen Schulbuchforschung*, Volume 9:21–44. Eckert. Expertise, Volume 9. V&R unipress. <https://doi.org/10.14220/9783737009539.21>.

Hertling, Anke, und Sebastian Klaes. 2018b. „GEI-Digital als Grundlage für Digital-Humanities-Projekte: Erschließung und Datenaufbereitung“. In *Digital Humanities in der internationalen Schulbuchforschung*, Volume 9:45–68. Eckert. Expertise, Volume 9. V&R unipress. <https://doi.org/10.14220/9783737009539.45>.

Reul, Christian, Dennis Christ, Alexander Hartelt, Nico Balbach, Maximilian Wehner, Uwe Springmann, Christoph Wick, Christine Grundig, Andreas Büttner, und Frank Puppe. 2019. „OCR4all—An Open-Source Tool Providing a (Semi-)Automatic OCR Workflow for Historical Printings“. *Applied Sciences* 9 (22): 4853. <https://doi.org/10.3390/app9224853>.