

Forschungssoftware rezensieren – Konzeption, Durchführung und Umsetzung

Homburg, Timo

timo.homburg@hs-mainz.de
Hochschule Mainz, Mainz, Deutschland

Klammt, Anne

aklammt@hotmail.com
Deutsches Forum für Kunstgeschichte Paris,
Frankreich

Offert, Fabian

offert@ucsb.edu
University of California, Santa Barbara, USA

Thiery, Florian

florian.thiery@rgzm.de
Römisch-Germanisches Zentralmuseum, Mainz,
Deutschland

Der Workshop adressiert eine signifikante Lücke in den Praktiken und Formaten der digitalen Geisteswissenschaften als *Open Humanities*: die bisher noch untergeordnete Rolle von Softwarerezensionen von *Forschungssoftware*.

Wissenschaftliche Rezensionen von Forschungssoftware schätzen deren Beitrag zur Lösung einer Aufgabe im Forschungsprozess, ihre Anwendbarkeit und Zielgruppe, sowie ihre handwerkliche Qualität und Nachhaltigkeit ein. Softwarerezensionen wären daher ein wesentlicher Bestandteil einer Wissenschaftspraxis, die ihre Methoden offenlegt und kritisch reflektiert. Bislang aber erscheinen noch wenige Rezensionen und das, obwohl es mittlerweile in den Geisteswissenschaften im deutschsprachigen Raum erste Zeitschriften zu ihrer Veröffentlichung gibt. Es finden sich jedoch noch kaum Autor*innen.¹ Der Workshop soll Interessierten einen Einstieg bieten und somit zur Verbreitung und Anerkennung dieses wichtigen wissenschaftlichen Formats beitragen.

Gemeinschaftlich mit den Teilnehmer*innen möchten wir am Beispiel kleiner, überschaubarer Tools alle Arbeitsschritte einer Rezension von Forschungssoftware am praktischen Beispiel durchführen. Im Mittelpunkt steht das Kennenlernen und Anwenden von Aspekten, mit denen eine Software besprochen werden kann. Die Beurteilung von Forschungssoftware muss ihre Aufgabe in der Forschung, ihre Nutzbarkeit aus Anwender*innensicht und ihre Nachhaltigkeit aus technischer Sicht umfassen. Damit fragen Rezensionen von Software ganz unterschiedliche Kompetenzen ab und sind daher be-

sonders gut und effizient als Team zu bearbeiten. Im Idealfall entsteht während des Workshops genug Material, um unterstützt von den Workshopanbieter*innen ohne umfangreiche Nacharbeiten eine Rezension beim Journal CKIT oder den Archäologischen Informationen einzureichen.

Forschungssoftware ist ein wesentlicher Bestandteil (digitaler) geisteswissenschaftlicher Forschung. Sie ermöglicht und steuert oft den gesamten Forschungsprozess (Katerbow u. Feulner 2018; Schmidt u. Marwick 2020). Dieser Einfluss reicht von der Auswahl und Strukturierung der Daten, über die angewandten rechnerischen Verfahren bis hin zu den Ausgabeformaten. Entsprechend können fehlerhaft implementierte Algorithmen, irreführende Nutzeroberflächen und unvollständige Dokumentationen die Forschung erschweren oder sogar zum Scheitern von Forschungsprozessen führen. Auch bei einem positiven Verlauf schreiben sich die Tools so tief in die Ergebnisse ein, dass eine digitale Quellenkritik immer wieder auch die Betrachtung der vorhergehenden Prozessierung und somit der Tools einbezieht. Dokumentierte, quelloffene Software ermöglicht schließlich, sie gezielt für spezifische Bedarfe weiterzuentwickeln oder ihre grundlegende Idee und Konzeption in Softwareprojekten mit aktuellen Bibliotheken erneut umzusetzen.

Aus unterschiedlichen Gesichtspunkten ist somit die kritische Betrachtung und Würdigung von Software, die in spezifischer Weise in der Forschung eingesetzt wird, sehr wünschenswert. Bislang wird der Bedarf, verstreut und über verschiedene Nutzergruppen verteilt, oft durch Erfahrungsberichte, Tutorials und zitierfähige Publikationen der Software etwa über Github oder Zenodo bedient. Insbesondere Erfahrungsberichte enthalten vielfach bereits Informationen, die Auswahlkriterien für die Nutzer*innen sind. Dies betrifft nicht alleine Anmerkungen zur Form und Verständlichkeit der Interaktion mit dem Programm oder die Import- und Exportfunktionen, sondern die Berichte zeigen zudem, wie die Software für die Bearbeitung einer geisteswissenschaftlichen Forschungsfrage eingesetzt wird. Allerdings finden Merkmale zur Beurteilung der Stabilität der Programme, zu ihrer Erweiterbarkeit oder auch ihrem Einsatz in anderen technischen Setups kaum Berücksichtigung. Tutorials, oft von den Entwickler*innen selbst verfasst, geben meist ebenfalls umfassend Einblick in die Funktionen, bewerten diese aber nicht. *Benchmark papers* und Softwarepublikationen richten sich schließlich nur an Entwickler*innen.

Nach der Bestimmung von Zielen und Blickwinkeln für Softwarerezensionen in der geisteswissenschaftlichen Forschung stellt sich die Frage nach den Kriterien, der Vorgehensweise und spezifischen Anforderungen. Die Zeitschrift Archäologische Informationen hat 2021 einen Vorschlag veröffentlicht (Homburg et al. 2021), der gemeinsam von Fachwissenschaftlerinnen und Informatiker*innen unterschiedlicher Schwerpunktsetzungen verfasst, und inzwischen kommentiert wurde (Carloni 2021; High-Steskal 2021). Der Text war eine wichtige Grundlage für die Ausformulierung der von den Herausgeber*innen der Zeitschrift CKIT formulierten Leitfragen, an denen sich die erwünschten Beiträge ausrichten sollen (CKIT 2021). Sie sind im Expert*innenforum der Task Area 3 "Research tools and data services" von NFDI4Culture 2022

diskutiert und angenommen worden, so dass sie heute in einer inhaltlichen und funktionalen Verbindung mit der entstehenden Software Registry stehen. Sie sind damit ein Anfang, um Gewohnheiten und *best practices* in der Forschungsgemeinschaft zu entwickeln. Entsprechend werden sie auch von den Workshopanbieter*innen als Orientierung verwendet.

Gemeinsame Ziele

- Erwerb der Kenntnis von wesentlichen Parametern zur Beurteilung einer Software im Forschungskontext
- Sammeln von Erfahrungen in der Zusammenstellung von spezifischen Nutzeranforderungen an eine Software
- Erwerb der Kenntnis von Vorgehensweisen, um Informationen zur handwerklichen Qualität von Software zu sammeln
- Sammeln von Erfahrungen im Einschätzen eigener Kompetenzen zur Beurteilung von Software
- Sammeln von Erfahrungen im (gemeinsamen) Verfassen einer Softwarerezension im Forschungskontext

Ergebnisformat

Ziel des Workshops ist es, auf Grundlage der genannten Handreichung (CKIT 2021) gemeinsam die Funktionalität und handwerklichen Qualität einer Software zu beschreiben sowie ihre Nutzbarkeit im Zuge der Bearbeitung einer geisteswissenschaftlichen Frage zu beurteilen.

Im Idealfall soll daraus eine gemeinsam verfasste Rezension entstehen, die bei der Zeitschrift CKIT oder den Archäologischen Informationen zur Veröffentlichung eingereicht wird.

Beispiele zur Rezension vorgeschlagener Forschungssoftware

1. SPARQLing Unicorn QGIS Plugin

Beschreibung: QGIS-Plugin (noch kein *stable release*, experimental), das eine einfache Integration von Geodaten aus Wikidata und anderen Linked Open Data SPARQL Endpoints ermöglicht. (Plugin: <https://plugins.qgis.org/plugins/sparqlunicorn/>; Github: <https://github.com/sparqlunicorn/sparqlunicornGoesGIS>)

2. PixPlot

Beschreibung: WebApp / eigenständige Software zum Clustern und anschließender Visualisierung von Bildern auf Grundlage von *neural network features*. (Projekt: <https://dhlab.yale.edu/projects/pixplot/> ; Github: <https://github.com/YaleDHLab/pix-plot>)

3. Wax

Beschreibung: Eine Software-Lösung, um einfach digitale Ausstellungen mit IIIF-Technologien zu realisieren. Die Software verfolgt einen *minimal computing-An-*

satz. (Projekt: <https://minicomp.github.io/wax/> ; Github: <https://github.com/minicomp/wax>)

Workshoporganisation

Der Workshop wird in fünf Schritten durchgeführt. Dabei übernehmen die Workshopanbieter*innen die Funktion als Lotsen und geben zu jedem Schritt einen inhaltlichen Input. Anschließend wird gemeinsam von den Teilnehmer*innen die Untersuchung der Software auf die zuvor festgelegten Kriterien vorgenommen und die Ergebnisse von ihnen dokumentiert. Als Arbeitsumgebung wird ein GitHub-Repositorium genutzt. Der Fokus liegt auf der Evaluierung der Software in der Gruppe und dem Austausch über die Anwendbarkeit und Bedeutung der verschiedenen Kriterien zur Beschreibung der ausgewählten Software. Der Workshop schließt mit einer gemeinsamen Reflexion zum Verlauf des Workshops ab.

Ablauf

Vorfeld

Schritt 1 (im Vorfeld der DHD)

- Im Vorfeld des Workshops installieren die Teilnehmer*innen die Software und benutzen sie in einem vorgegebenen *use case* mit einem Testdatensatz. Sie notieren dabei, unterstützt durch einige Leitfragen, für sich ihre Erfahrungen.
- In einer gemeinsamen Tabelle geben sie anonym Kennwerte zu ihrem technischen Setup an. Arbeitsumfang: etwa 120 min.

Workshop (0,5 Tage)

- Kennenlernen
- Erste Eindrücke zur Software sammeln

Schritt 2 - Forschungskontext der Software und grundlegende Funktionen

- Impuls 1 (5-10 min.): *Vorstellung des Anwendungsbereichs der Software in der geisteswissenschaftlichen Forschung*
- Gemeinsame Auswahl und Anwendung von Kriterien, um die generelle Funktion der Software zu beschreiben und zu beurteilen.
- Beurteilung der Software nach diesen Kriterien (Gruppenarbeit)

Schritt 3 - Perspektive Anwender*innen

- Impuls 2 (5-10 min.): *Den Wald vor lauter Bäumen nicht sehen - Schwerpunkte in der Beschreibung aus Anwender*innensicht setzen*
- Gemeinsame Auswahl von Kriterien
- Einbeziehung subjektiver Aspekte (Kompetenzen, Forschungsinteresse, Gewohnheiten)
- Beurteilung der Software nach diesen Kriterien (Gruppenarbeit)

Workshop (0,5 Tage)

Schritt 4 - Perspektive Entwickler*innen

- Impuls 3 (5-10 min.): *Finden, was für Entwickler*innen interessant ist*
- Gemeinsame Auswahl und Anwendung von Kriterien, um die Software mit Blick auf ihre handwerkliche Qualität, Nachhaltigkeit und Entwicklungsfähigkeit zu beschreiben
- Zwischenbilanz : Einschätzung der Software

Schritt 5 - Bilanz und Ausblick

- Zusammenführung der ausgewerteten Kriterien und der Zwischenbilanzen
- Gemeinsame Verständigung, ob eine Fortführung als Publikationsprojekt möglich, sinnvoll und machbar ist
- Feedback zum Workshop

Vorkenntnisse und Kompetenzen

- Es sind keine spezifischen Vorkenntnisse erforderlich. Wesentlich ist ein allgemeines Interesse an Software und die Motivation zum offenen Arbeiten im Team.

Technisches Setup

- Die Teilnehmer*innen bringen ihren eigenen Rechner mit. Vor Ort wird ein Bildschirm benötigt.

Teilnehmer*innenzahl

Max. 15

Anbieter*innen Workshop

- Timo Homburg ist Wissenschaftlicher Mitarbeiter am i3mainz Institut für Raumbezogene Informations- und Messtechnik der Hochschule Mainz. Er erforscht Anwendungen im Bereich (Geospatial) Semantic Web, Computerlinguistik und den Digital Humanities. Seit 2019 ist er auch aktiver Wegbereiter für neue Geodatenstandards in der Standardisierungsgruppe zu GeoSPARQL, der OGC.
- Anne Klammt ist zum Zeitpunkt der Einreichung als Forschungsleiterin am Deutschen Forum für Kunstgeschichte Paris verantwortlich für die Digital Humanities. Seit 2020 publiziert sie zur Frage, wie Forschungssoftware und Datendienste wissenschaftlich rezensiert werden können.
- Fabian Offert ist Assistant Professor for the History and Theory of the Digital Humanities an der University of California in Santa Barbara. Er ist Mitherausgeber der Zeitschrift CKIT und forscht zur Nutzung von neuesten Praktiken des maschinellen Lernens in den digitalen Geisteswissenschaften, mit einem Schwerpunkt im Bereich digitale Kunstgeschichte.
- Florian Thiery ist Research Software Engineer im "Arbeitsbereich Wissenschaftliche IT, digitale Plattformen und Tools" des Römisch-Germanischen

Zentralmuseums und forscht dort und entwickelt Forschungssoftware im Sonderforschungsgebiet "Explorative Forschung, Theorien- und Methodenentwicklung" im Handlungsfeld "Semantic Modelling and Knowledge Graphs". Seit 2020 ist er Mitglied des Vorstands der Gesellschaft für Forschungssoftware (de-RSE e.V.), Mitentwickler des SPARQLing Unicorn QGIS Plugin und forscht im Bereich semantischer Modellierung und Linked Open Data und Wikidata zu Irischen Ogham Steinen, was im Rahmen des Wikimedia Deutschland Fellow-Programms Freies Wissen gefördert wurde.

Fußnoten

1. Nach eigenen Erfahrungen der Workshopanbieter*innen und Rückmeldungen der Redaktion der Zeitschrift Archäologischen Informationen, die seit 2019 eine eigene Rubrik für Softwarerezensionen bereitstellen.

Bibliographie

«Construction KIT: a review journal for research tools and data services in the humanities». Zugegriffen 26. Juli 2022. <https://journals.ub.uni-heidelberg.de/index.php/ckit/index>.

High-Steskal, Nicole. 2021. «Contribution to the discussion: Handreichung zur Rezension von Forschungssoftware in der Archäologie und den Altertumswissenschaften: (Homburg, Klammt, Mara et al., 2020)». *Archäologische Informationen* 44: 257–59. <https://doi.org/10.11588/ai.2021.1.89200>.

Homburg, Timo, Anne Klammt, Hubert Mara, Clemens Schmid, Sophie C. Schmidt, Florian Thiery, und Martina Trognitz. 2021. «Diskussionsbeitrag: Handreichung zur Rezension von Forschungssoftware in der Archäologie und den Altertumswissenschaften.». *Archäologische Informationen* 43: 357–72.

Katerbow, Mathias und Feulner, Georg. 2018. «Handreichung zum Umgang mit Forschungssoftware». Zenodo, 27.2.2018. <https://doi.org/10.5281/zenodo.1172970>.

Carloni, Massimiliano. 2021. «Einige Anmerkungen zur ‚Handreichung zur Rezension von Forschungssoftware‘ : (Homburg, Klammt, Mara et al., 2020)». *Archäologische Informationen* 44: 253–56. <https://doi.org/10.11588/ai.2021.1.89199>.

Schmidt, Sophie. C. und Ben Marwick. 2020. «Tool-Driven Revolutions in Archaeological Science». *Journal of Computer Applications in Archaeology* 3: 18–32. <https://doi.org/10.5334/jcaa.29>.