

# Die Werkbänke der Digital Humanities: Zur Rolle von Tools und Software für die Forschungsarbeit

Frank Fischer (Higher School of Economics, Moskau), Manuel Burghardt (Universität Leipzig), Jan Luhmann (Universität Leipzig), Laure Barbot (DARIAH-EU), Yoann Moranville (DARIAH-EU), Alireza Zarei (GWDG Göttingen)

## Einleitung

Quinn Dombrowski hat prägnant das „Directory Paradox“ beschrieben: Obwohl es einen breiten Konsens dahingehend gibt, dass die in den DH eingesetzten Tools in sog. Tool Directories gesammelt werden sollten, hat es bisher kein entsprechendes Projekt geschafft, sich nachhaltig durchzusetzen, wenn man vom kanadischen TAPoR absieht (Barbot et al. 2020, Dombrowski 2020). Derlei Projekte wären als Referenzpunkt auch die Voraussetzung dafür, den Einsatz von Tools in der Forschung besser zu beobachten. Drei Gründe, warum dies wichtig ist:

- 1. Häufig genutzte Tools könnten bei der Finanzierung ihrer Nachhaltigkeit priorisiert werden.
- 2. Häufig genutzte Tools könnten für Trainingseinheiten bevorzugt werden.
- 3. Ein Tool-Überblick würde auch die Analyse von Lücken in der Toolchain der DH leichter identifizierbar machen.

Ziel dieses Panels ist es, diese Gedanken anhand eigener Experimente produktiv weiterzudenken. Dafür werden relevante Projekte in der DH-Forschungslandschaft vorgestellt und überlegt, wie sich diese verschränken ließen, um die Praxis der DH besser zu erforschen und auf aktuellem Stand zu diskutieren.

## Forschungsstand

„Es gibt keine Software“ polemisierte Friedrich Kittler 1993 gegen die Auffassung, dass erst vorgefertigte, „anwenderfreundliche“ Software die Hardware von Computern nutzbar mache. Trotz dieser Kritik entwickelt sich seit spätestens 2008 das Feld der Software Studies, die angesichts der wachsenden Bedeutung von Software eine neue Art von Kulturtheorie fordert (Fuller 2008). Das 2010 in Großbritannien gegründete Software Sustainability Institute hat seine Arbeit unter dem Slogan „Better Software – Better Research“ begonnen. Dem Umstand, dass Softwareentwicklung eine immer wichtigere Stellung in der Forschungsarbeit einnimmt, wurde außerdem 2018 durch die Gründung des Vereins „Research Software Engineerig“ Rechnung getragen (<https://de-rse.org/en/index.html>). Parallel dazu entwickelt sich auch ein medienarchäologischer Blick auf Software – so verschlagwortet das oben genannte kanadische Portal TAPoR einige Werkzeuge als „Historic Tool (developed before 2005)“.

Zur Nutzung von Tools in der DH-Forschung gibt es erste empirische Ansätze. Der ToolXtractor etwa ermöglicht es, in DH-Publikationen erwähnte Software und Tools zu extrahieren und zu analysieren (Barbot et al. 2019, vgl. auch <https://bbrause.github.io/tools-in-dh/> für chronologisch orientierte Visualisierungen). Eine auf Machine Learning basierende Software, die ähnliches tut, ist GROBID. Voraussetzung für eine bessere Erkennungsrate sind Goldannotationen von Forschungspapern aus dem Bereich der DH (der ToolXtractor hat eine entsprechende Funktion für Prä-Annotationen).

## Thesen zur Softwarenutzung in den Digital Humanities

In Wolff 2015 wird „Tool Science“ als eigene Unterdisziplin im Spannungsfeld von Software Engineering und Information Behavior vorgeschlagen. Darin findet sich auch eine grundlegende Taxonomie zur Kategorisierung von Tools:

- 1. generic development tools: Programmiersprachen, Libraries, Frameworks, IDEs etc.
- 2. standard commercial off-the-shelf (COTS) software: Microsoft Office etc.
- 3. software infrastructures: VREs wie CLARIN oder TextGRID; aber auch SPSS oder Atlas.TI
- 4. software specifically designed for specific research projects or research questions: CATMA? AntConc? ...

Angewandt auf die DH lässt sich konstatieren:

Ad 1: In der frühen Zeit des Humanities Computing spielten Programmiersprachen selbst eine größere Rolle (COBOL, FORTRAN, Perl), das Resultat waren viele selbstgebaute Skripte und proprietäre Software. Ein Blick in aktuelle Forschungspapiere zeigt, dass in den DH (und der Data Science) vor allem Python und R als De-Facto-Programmiersprachen gelten können, meist gepaart mit entsprechenden Bibliotheken (Barbot et al. 2019).

Ad 2: COTS-Software spielt in den DH eine untergeordnete Rolle, wenn man das Fach dahingehend definiert, dass digitale Tools so eingesetzt werden, dass man in der Forschung neue Erkenntnisse erlangt, die ohne Tools nicht möglich gewesen wären.

Ad 3: (Virtuelle) Forschungsumgebungen sind vermutlich der Bereich, in den im Rahmen von DH-Projekten das meiste Fördergeld geflossen ist. Als Motivation dahinter stehen die Nachhaltigkeit und eine breite Einsetzbarkeit.

Ad 4: Software, die für spezifische Forschungsfragen entwickelt wurde, dürfte die DH-Tool-Landschaft immer noch maßgeblich prägen. Ein häufiges Merkmal ist die Kurzlebigkeit dieses Typus, nach dem Auslauf eines Projekts wird ein Tool oft nicht mehr gepflegt. Dies gilt ironischerweise auch für Tool Directories, wenn man sich etwa die konkreten Fälle TERESAH (<http://teresah.dariah.eu/>) oder DiRT Directory anschaut (Dombrowski 2014). Gleichwohl gibt es auch positive Gegenbeispiele, etwa CATMA, das ursprünglich für ein spezifisches Projekt entwickelt wurde und sich sukzessive zum General-Purpose-Tool für Annotationen weiterentwickelt hat, mit entsprechender Förderung und viel Energie für Dissemination.

## Software und Training

Die Voraussetzung für die Adaption von Softwarelösungen auch über Projektkontexte hinaus sind Workshops und breit verfügbare Trainingsmaterialien. Vorreiter auf diesem Gebiet ist der „Programming Historian“, der mittlerweile Lerneinheiten in vier Sprachen zur Verfügung stellt und häufig genutzte DH-Tools explizit mit einbezieht (Fischer/Moranville 2020).

Eine Korrelierung zwischen Tools und entsprechenden Schulungsmaterialien ist die raison d'être des SSH Open Marketplace, der im Rahmen des EU-Projektes „Social Sciences and Humanities Open Cloud“ (SSHOC) bis 2022 entwickelt wird (Beta-Version: <https://marketplace.sshopencloud.eu/>).

## Ablauf des Panels

Das Panel wird in drei Kapiteln vorgetragen (1. Einleitung & Forschungsstand; 2. Thesen zur Softwarenutzung in den DH; 3. Ausblick).

## Bibliografie

Barbot, Laure; Fischer, Frank; Moranville, Yoann; Pozdniakov Ivan: Which DH Tools Are Actually Used in Research? In: weltliteratur.net, 6 December 2019. (URL: <https://weltliteratur.net/dh-tools-used-in-research/>)

Barbot, Laure; Dombrowski, Quinn; Fischer, Frank; Rockwell, Geoffrey; Spiro, Lisa (2020): Who Needs Tool Directories? A Forum on Sustaining Discovery Portals Large and Small. In: DH2020: „carrefours/intersections“. 22–24 July 2020. Book of Abstracts. University of Ottawa. (URL: [https://dh2020.adho.org/wp-content/uploads/2020/07/126\\_WhoneedstooldirectoriesAforumonsustainingdiscoveryportalslargeandsmall.html](https://dh2020.adho.org/wp-content/uploads/2020/07/126_WhoneedstooldirectoriesAforumonsustainingdiscoveryportalslargeandsmall.html))

Dombrowski, Quinn (2014): What Ever Happened to Project Bamboo? In: Literary and Linguistic Computing, Vol. 29, Issue 3, September 2014, S. 326–339, doi:10.1093/lc/fqu026.

Dombrowski, Quinn (2020): The Directory Paradox. In: Anne McGrail et al. (Hg.): Debates in Digital Humanities: Institutions, Infrastructures at the Interstices. University of Minnesota Press (forthcoming).

Fischer, Frank; Moranville, Yoann (2020): DH Tools Mentioned in „The Programming Historian“. In: weltliteratur.net, 17 Jan 2020. (URL: <https://weltliteratur.net/dh-tools-programming-historian/>)

Fuller, Matthew (Hg.) (2008): Software Studies. A Lexicon. Cambridge: MIT Press.

Wolff, Christian (2015): The Case for Teaching „Tool Science“. Taking Software Engineering and Software Engineering Education beyond the Confinements of Traditional Software Development Contexts. In: 2015 IEEE Global Engineering Education Conference (EDUCON), Tallinn, S. 932–938, doi:10.1109/EDUCON.2015.7096085.

Anm.: Die vDHD 2021 ist eine durch die Community des Verbandes Digital Humanities im deutschsprachigen Raum organisierte virtuelle Konferenz. Die Abstracts haben kein Peer-Review-Verfahren durchlaufen.