

SustainLife - Erhalt lebender, digitaler Systeme für die Geisteswissenschaften

Barzen, Johanna

johanna.barzen@iaas.uni-stuttgart.de
Universität Stuttgart, Deutschland

Blumtritt, Jonathan

jonathan.blumtritt@uni-koeln.de
Universität zu Köln, Deutschland

Breitenbücher, Uwe

uwe.breitenbuecher@iaas.uni-stuttgart.de
Universität Stuttgart, Deutschland

Kronenwett, Simone

simonekronenwett@gmail.com
Universität zu Köln, Deutschland

Leymann, Frank

frank.leymann@iaas.uni-stuttgart.de
Universität Stuttgart, Deutschland

Mathiak, Brigitte

bmathiak@uni-koeln.de
Universität zu Köln, Deutschland

Neuefeind, Claes

c.neuefeind@uni-koeln.de
Universität zu Köln, Deutschland

Der digitale Wandel verändert die Wissenschaft grundlegend (Kramp 2013). Das exponentielle Wachstum, die steigende Komplexität sowie der zunehmende Gebrauch von digitalen Forschungsdaten beeinflussen den Forschungsprozess signifikant. Um das Potential der fortschreitenden Digitalisierung optimal nutzen zu können, müssen entsprechende Infrastrukturen geschaffen werden, die das Management von Forschungsdaten, die Möglichkeit ihrer Vernetzung, ihre dauerhafte Verfügbarkeit und einen freien Zugang gewährleisten.

Die Vielzahl von wissenschaftspolitischen Empfehlungen, Bestandsaufnahmen, Umfragen und institutionellen Richtlinien rund um Forschungsdaten, die in den vergangenen Jahren veröffentlicht wurden, sind ein Zeichen zunehmenden Problembewusstseins, politischen Handlungswillens, aber eben auch anhaltenden

Handlungsbedarfs (Pampel / Bertelmann 2011; Wissenschaftsrat 2012; DV-ISA 2016; Rat für Informationsinfrastrukturen 2016). Durch die großen europäischen und nationalen Infrastrukturprojekte in den Geisteswissenschaften (CLARIN, DARIAH), die Einrichtung von fachspezifischen Datenzentren und spezifisch geisteswissenschaftlichen Datenzentren, wie dem Data Center for the Humanities (DCH, s. <http://dch.uni-koeln.de>), hat sich die Versorgungslage für Forschungsdaten stetig verbessert. Doch längst werden nicht alle produzierten Daten und digitalen Forschungsergebnisse tatsächlich für die Nachnutzung verfügbar gemacht bzw. sind für eine dauerhafte Verfügbarkeit in einer hochdynamischen digitalen Welt gerüstet (Razum / Neumann 2014; Wissenschaftsrat 2012).

Die Diskussion fokussiert bisher stark auf Forschungsinformationssysteme zur standardisierten Vorhaltung von Forschungsprimärdaten. Dabei bleibt weitgehend unberücksichtigt, dass ein Großteil der digitalen Produkte in den Geisteswissenschaften in einer Form vorliegen, die sich einer normierten Versorgung bislang entzieht. Gemeint sind Forschungsanwendungen, oder "lebende Systeme" (Sahle/Kronenwett 2013), die einen wesentlichen Bestandteil digitaler Ergebnissicherung darstellen: Präsentationssysteme, interaktive Visualisierungen, Recherche-Datenbanken, digitale Editionen und digitale Arbeitsumgebungen – um nur einige Formen zu nennen – sind Arbeitswerkzeuge, Plattformen der Dissemination und Aggregation von Forschungsergebnissen und sind aus dem Arbeitsalltag in Geisteswissenschaften nicht mehr wegzudenken. Ihre dauerhafte Erhaltung, Betreuung und Bereitstellung über den Zeitraum der Projektförderung hinaus stellt eine große organisatorische und letztlich finanzielle Herausforderung dar.

Ein grundlegendes Problem ist die Sorge um softwaregestützte Präsentationen und Funktionalitäten, die in Forschungsprojekten entstehen (Sahle/Kronenwett 2013). Nicht selten sind diese die eigentlichen Träger von „Informationsgehalt beziehungsweise wissenschaftliche[m] Mehrwert“ der im Projekt erbrachten Forschungsleistung (Wuttke et al. 2016). Eine Nachhaltigkeitsstrategie, die auf eine Abtrennung und Archivierung allein der Datenbasis zurückfällt, führt unweigerlich zum Verlust von Information und reduziert im Extremfall den wissenschaftlichen Nutzen auf null (Blumtritt/Mathiak 2016).

Forschungsanwendungen sind keineswegs statische Objekte, sondern unterliegen einem kontinuierlichen Veränderungszyklus. Viele Anwendungen fungieren als Plattformen, die User-Input entgegennehmen und damit ihre Datenbasis laufend verändern. Browser-Updates und Veränderungen der Nutzungsgewohnheiten können bestimmte Komponenten unbrauchbar oder obsolet machen und damit eine Überarbeitung des Codes anstoßen. Notwendige Sicherheitsupdates erfordern regelmäßiges Eingreifen und können Kaskaden von weiteren Updates und Softwareanpassungen nach sich ziehen. Der Verzicht

auf kontinuierliche Wartung spart kurzfristig Kosten, verschärft mittelfristig aber das Problem. Erfahrungen aus dem "LAZARUS-Projekt" zeigen, dass Anwendungen, die über längere Zeit brach liegen, nur unter großem Ressourceneinsatz wieder revitalisiert werden können (Bingert et al. 2016). Ein erster Ansatz zur Archivierung von Forschungssoftware auf Basis des TOSCA-Standards wurde vorgestellt, löst jedoch nur das Problem der initialen automatisierten Bereitstellung, nicht der Revitalisierung terminierter Systeme (Breitenbücher et al. 2017).

Darüber hinaus sehen sich Forschungseinrichtungen mit dem Dilemma konfrontiert, dass befristete Projektlaufzeiten einen dauerhaften Betrieb organisatorisch deutlich erschweren. Dies ist insbesondere in den Geisteswissenschaften fatal, da hier oftmals ein anderer Maßstab bezüglich der Nachhaltigkeit angelegt wird, als dies in den vergleichsweise schnelllebigen Naturwissenschaften üblich ist. Die institutionelle Vorhaltung von Forschungsdaten für einige Jahrzehnte erfüllt den Zweck der Überprüfbarkeit und Reproduzierbarkeit im Sinne der guten wissenschaftlichen Praxis, für die Vorhaltung von Gegenständen des kulturellen Erbes ist eine zeitliche Beschränkung nicht sinnvoll. Selbst dann, wenn sich eine Abschaltung aus Sicherheits- oder Kostengründen nicht vermeiden lässt, so ist zumindest zu gewährleisten, dass Anwendung und Datenbasis auf eine Art und Weise archiviert werden, dass sie sich jederzeit verlustfrei und benutzbar wiederherstellen lassen.

Das Projekt „SustainLife“, das in Zusammenarbeit zwischen dem Data Center for the Humanities (DCH) aus Köln und dem Institut für Architektur von Anwendungssystemen (IAAS) der Universität Stuttgart durchgeführt wird, adressiert diese Probleme. In diesem Poster stellen wir das Projekt vor und präsentieren den ersten Teil einer Anforderungsanalyse an lebende Systeme der Digital Humanities. Als zweiten Teil soll vor Ort eine Umfrage unter den Teilnehmern der Konferenz durchgeführt werden, um die Anforderungen an lebende Systeme aus unterschiedlichsten Domänen möglichst zielgenau zu analysieren. Diese Anforderungsanalyse stellt die Basis für im Projekt SustainLife geplante Erweiterungen der open-source Software „OpenTOSCA“ (Binz et al. 2013) dar, welche auf dem OASIS-Standard TOSCA (OASIS 2013) basieren. Dieser Standard ermöglicht die Automatisierung des Deployments und Managements von Anwendungen und stellt damit eine vielversprechende Grundlage zur kostengünstigen Sicherung der Nachhaltigkeit lebender Systeme dar.

Bibliographie

Bingert, Sven / Blumtritt, Jonathan / Buddenbohm, Stefan / Engelhardt, Claudia / Kronenwett, Simone / Kurzawe, Daniel (2016): "Anwendungskonservierung und die Nachhaltigkeit

von Forschungsanwendungen", in: Forschungsdaten in den Geisteswissenschaften (FORGE 2016) 13-16 <https://www.fdm.uni-hamburg.de/ueber-uns/anachrichten/aktivitaeten/forge16/presentationen/programmheft.pdf> [letzter Zugriff 20. September 2017].

Binz, Tobias / Breitenbücher, Uwe / Haupt, Florian / Kopp, Oliver / Leymann, Frank / Nowak, Alexander / Wagner, Sebastian (2013): "OpenTOSCA - A Runtime for TOSCA-based Cloud Applications", in: Proceedings of the 11th International Conference on Service-Oriented Computing (ICSOC 2013). Springer.

Blumtritt, Jonathan / Mathiak, Brigitte (2016): "Consulting Workflow for Humanities Research Data", in: Forschungsdaten in den Geisteswissenschaften (FORGE 2016) 21-22 <https://www.fdm.uni-hamburg.de/ueber-uns/anachrichten/aktivitaeten/forge16/presentationen/programmheft.pdf> [letzter Zugriff 20. September 2017].

Breitenbücher, Uwe / Barzen, Johanna / Falkenthal, Michael / Leymann, Frank (2017): "Digitale Nachhaltigkeit in den Geisteswissenschaften durch TOSCA: Nutzung eines standardbasierten Open-Source Ökosystems", in: DHd 2017: Digitale Nachhaltigkeit 235-238 http://www.dhd2017.ch/wp-content/uploads/2017/01/Abstractband_def_24.1.17-1.pdf [letzter Zugriff 20. September 2017].

DV-ISA (2016): "Umgang mit digitalen Daten in der Wissenschaft: Forschungsdatenmanagement in NRW. Eine erste Bestandsaufnahme", Version 0.7, <https://www.dvisa-nrw.de/veroeffentlichungen/veroeffentlichungen-container-oeffentlich/dv-isa-vorstudie-bestandsaufnahme-forschungsdatenmanagement> [letzter Zugriff 20. September 2017].

Kramp, Leif (2013): "Wie die Digitalisierung die Wissenschaft verändert", in: sueddeutsche.de, 20.11.2013, <http://www.sueddeutsche.de/wissen/digitales-morgen-debatte-zur-digitalisierung-wie-die-digitalisierung-die-wissenschaft-veraendert-1.1823133> [letzter Zugriff 20. September 2017].

OASIS (2013): Topology and Orchestration Specification for Cloud Applications Version 1.0.

Pampel, Heinz / Bertelmann, Roland (2011): "Data Policies im Spannungsfeld zwischen Empfehlung und Verpflichtung", in: Handbuch Forschungsdatenmanagement 49–61, Bad Honnef: Bock + Herchen.

Rat für Informationsinfrastrukturen (2016): "Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland", <http://www.rfii.de/download/rfii-empfehlungen-2016/> [letzter Zugriff 20. September 2017].

Razum, Matthias / Neumann, Janna (2014): "Das RADAR Projekt: Datenarchivierung und -publikation als Dienstleistung - disziplinübergreifend, nachhaltig, kostendeckend", in: o|bib Das offene Bibliotheksjournal 1/1: 30–44 <https://www.o-bib.de/article/view/2014HIS30-44/117> [letzter Zugriff 20. September 2017].

Sahle, Patrick / Kronenwett, Simone (2013): "Jenseits der Daten: Überlegungen zu Datenzentren für die Geisteswissenschaften am Beispiel des Kölner 'Data Center for the Humanities'", in: LIBREAS. Library Ideas 23: 76–96.

Wissenschaftsrat (2012): "Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020", <http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf> [letzter Zugriff 20. September 2017].

Wuttke, Ulrike / Engelhardt, Claudia / Buddenbohm, Stefan (2016): "Angebotsgenese für ein geisteswissenschaftliches Forschungsdatenzentrum", in: Zeitschrift für digitale Geisteswissenschaften 2017, text/html Format, doi:10.17175/2016_003.