

Perspektiven der Benutzeraktionsanalyse im Kontext der Evaluation von Forschungspraktiken in den Digital Humanities

Walkowski, Niels-Oliver

walkowski@bbaw.de

Berlin-Brandenburgische Akademie der Wissenschaften, Deutschland

Ansätze zur Evaluation von Forschungsaktivität in den Digital Humanities

Die Dokumentation digitaler Forschungsprozesse ist seit langem ein viel diskutiertes Thema und wird im Allgemeinen mit dem Begriff Provenienz verbunden. Ziel der Erhebung von Provenienz Daten in digitaler Forschung ist es meist den Herstellungsprozess von Ergebnissen aufzuzeichnen um diese dadurch wissenschaftlich nachvollziehbar und reproduzierbar zu machen.

Seit ca. 2-3 Jahren entsteht in den Digital Humanities eine besondere Variante dieses Themas. Konkret geht es dabei um das Modellieren und Dokumentieren von Forschungsprozessen zur Identifikation digitaler geisteswissenschaftlicher Forschungspraktiken, bzw. -methoden. Zu diesem Zweck wurden mit dem *Scholarly Domain Model* (SDM) (Gradmann et al. 2015) und der *NeDiMAH Method Ontology* (NeMO) (Constantopoulos, Dallas, und Bernadou 2016) zwei Modelle geschaffen, die sowohl Konzepte für die Beschreibung von Forschungsprozessen als auch für deren Auswertung in Hinblick auf methodische Fragestellungen bieten.

Der Hintergrund für diese Aktivitäten ist zumeist das Bedürfnis von Infrastrukturprojekten, im Falle der genannten Modelle Europeana und Dariah, Anforderungen von NutzernInnen zu identifizieren und den qualitativen Gebrauch der bereitgestellten Dienste zu evaluieren. Darüber hinaus können sie aber auch als ein digitales Angebot zur Bearbeitung von Fragen der Science Studies verstanden werden. Im Kontext der Digital Humanities geht diese Perspektive mit dem Wunsch nach der Herausbildung eines methodischen Selbstbewusstseins einher (Constantopoulos, Dallas, und Bernadou 2016).

Versucht man die beiden genannten Ansätze danach zu unterscheiden wie sie sich dem Forschungsprozess nähern, so fällt zunächst ein Unterschied auf. Während in den beispielhaften Anwendungen von NeMO ("DCU OnTo - NeDiMAH Ontology Navigation" 2016) die Beschreibung

rückblickend erfolgt, wird sie bei SDM vorausschauend vollzogen. SDM nimmt hier Bezug auf das Konzept des 'modeling for' von Clifford Geertz.

Dieser Unterschied zwingt dazu, die zuvor gewählte Begrifflichkeit noch etwas weiter zu differenzieren. In der Forschungsliteratur werden drei verschiedene Begriffe für möglichen Varianten der Dokumentation von Forschungsaktivitäten verwendet (Hunter 2006). Diese sind:

1. Workflow
2. Provenienz
3. Lineage

Die Begriffe können auch als präskriptive, inskribierende, bzw. deskriptive Verfahren bezeichnet werden. Sie unterscheiden sich durch den Zeitpunkt von dem aus die Darstellung eines Forschungsprozesses erfolgt. SDM und NeMO realisieren somit die Workflow und die Lineage Perspektive. Was fehlt ist eine wirkliche Provenienz Perspektive im Kontext der methodischen Evaluation von Forschungsprozessen in den Digital Humanities. Provenienzdaten sind Daten die während einer Aktivität aufgezeichnet werden, indem zum Beispiel bestimmte Aktionen als Trigger für das Abspeichern von Informationen über diese Aktionen dienen.

Der Vorteil einer solchen Vorgehensweise gegenüber den anderen Ansätzen bei der Methodeevaluation ist zweierlei. Die Granularität mit der ein Forschungsprozess beschrieben werden kann ist höher als in SDM und NeMO. Der Anteil inhaltlicher Vorentscheidungen bei der Erfassung der Daten ist geringer. In der Beispielanwendung des SDM soll die Erweiterung der Praxis des Annotierens evaluiert werden. Gleichzeitig nimmt die Entscheidung was wann mit dem im SDM Vokabular bereitgestellten Begriff 'Annotieren' beschrieben wird das Ergebnis der Evaluation schon ein Stück weit vorweg. In einem Provenienz Ansatz reicht es aus zu definieren was aussagekräftige Ereignisse für die Dokumentation des Forschungsprozesses sind. Die Interpretation einer Ereigniskette kann später erfolgen.

Der Vortrag stellt eine beispielhafte Realisierung für ein inskribierendes Verfahren zur Evaluation von Forschungspraxis in den Digital Humanities im zuvor beschriebenen Sinn vor. Die Arbeiten sind Bestandteil der dritten Förderphase des DARIAH-DE Projekts. Der Ausgangspunkt für die Entwicklung des genannten Verfahrens bildet der *Wissensspeicher* ("Digitaler Wissensspeicher" 2016) der *Berlin-Brandenburgischen Akademie der Wissenschaften*.

Der Wissensspeicher ist ein Infrastrukturprojekt, dass eine inhaltliche und technisch Interaktion mit den heterogenen digitalen Ressourcen an der BBAW ermöglicht. Ziel des Wissensspeichers ist es diese Interaktion nicht mit der Objekt-, sondern der hinter den Objekten stehenden Inhaltsebene zu ermöglichen. Die Interaktionen können daher als wissensverarbeitende Prozesse verstanden werden, die es bei entsprechender

Evaluation erlauben Strategien der Wissensgenerierung zu identifizierbar.

Verfahren der Benutzer-Interaktionsanalyse im Kontext der Dokumentation von Forschungspraxis

Da die Aufzeichnung der Provenienz Daten die Beschreibung von wissensverarbeitenden Prozessen an Hand von bedeutungsvollen Handlungen zum Ziel hat und nicht einen Transformationsprozesses von Daten, ist die Dokumentation dieser Praxis schwieriger als zum Beispiel bei Hunter und anderen Implementierungen von Provenienz Modellen. Allerdings gibt es einen Forschungsbereich der sich tatsächlich mit einer ähnlichen Fragestellung beschäftigt. Dieser Bereich ist die *Benutzer-Aktions-Analyse*, beziehungsweise *User-Activity-Analysis*. User-Activity-Analysis zeichnet die Mensch-Computer Interaktion mit dem Ziel auf, diese vor dem Hintergrund einer spezifischen Fragestellung zu analysieren. In den meisten Fällen findet User-Activity-Analysis im Browser statt, dies ist jedoch nicht zwingend.

Die zwei Bereiche in denen User-Activity-Analysis hauptsächlich durchgeführt wird sind *e-commerce* und *online-social-networks*. Im Kontext des ersten Bereichs werden solche Analysen zum Beispiel für den Betrieb von Empfehlungssystemen durchgeführt (Plumbaum, Stelter, und Korth 2009). Bei Online-Social Networks steht häufig die Identifikation von Verhaltensmustern von Menschen in sozialen Interaktionen im Vordergrund (Dang et al. 2016)

Im Kontext akademischer Dienste und Umgebungen ist die User-Activity-Analysis noch nicht so weit verbreitet. Eine Ausnahme bilden spezielle Suchmaschinen für Forschungsliteratur. Beiträge, die dem hier vorgestellten Szenario noch am nächsten kommen sind die von Vozniuk et al. (2016) und Suire et al. (2016), die User-Activity-Analysis im Kontext von *e-learning* und im *cultural heritage* Bereich verorten.

Zwei Aufgabenstellungen sind zunächst einmal konzeptuell von einander zu trennen. Zum einen stellt sich die Frage was, überhaupt als Ausgangspunkt zur Erhebung von Daten in einer spezifischen Mensch-Computer Interaktion dienen kann. Dieser Bereich kann auch als *User-Activity-Tracking* bezeichnet werden. Der andere Bereich umfasst die Frage mittels welchen Verfahren den Ereignissen und auf deren Grundlage gewonnenen Daten Bedeutung zugeschrieben werden kann.

Viele unterschiedliche Strategien der Erhebung von Nutzeraktivitätsdaten sind seit der Entstehung des Web vorgeschlagen worden (Calzarossa, Massari, und Tessera 2016). Die am weitesten verbreitetste ist die sogenannte *Click-Stream-Analyse* bei der die Server-Logdateien, bzw. HTTP-Requests ausgewertet werden, die das Klicken auf Links durch den Benutzer erzeugen.

Dieser am einfachsten zu realisierende Ansatz ist aus verschiedenen Gründen problematisch. So enthalten Server-Logdateien keine Informationen über sogenannte 'leise Interaktionen' (Benevenuto et al. 2009), also Klicks, die nicht mit einem Request einhergehen und Interaktionen wie zum Beispiel Mausbewegungen, die gar keinen Klick beinhalten. Darüber hinaus enthalten die erzeugten Daten keinerlei Angaben über den Inhaltskontext in dem die Interaktion stattgefunden hat. Aus diesem Grund wurden Verfahren entwickelt, die auf der Basis von Browser Plugins oder mittels JavaScript (Dhawan und Ganapathy 2009) eine detailliertere Dokumentation von Benutzerinteraktion ermöglichen. Ein weiterer Ansatz versucht mittels Parsing- und Miningverfahren Interaktionsspuren oder Inhalte im Kontext von Interaktionen in die Analyse mit einzubeziehen (Vozniuk et al. 2016).

Wesentlich komplexer als die Frage danach welche User-Activity-Daten wie erhoben werden können ist die Frage wie ihnen und den Ereignissen die sie erzeugen Bedeutung beigemessen werden kann. Hier existieren ebenfalls eine Reihe unterschiedlicher Ansätze. Grundsätzlich lassen sich 5 verschiedene Ansätze unterscheiden. Dazu gehören solche, die die Bedeutung von Ereignissen

- im Vorfeld festlegen,
- vor dem Hintergrund von Durchschnittswerten aus den Daten, die für ein bestimmtes Ereignis erhoben wurden ermitteln
- durch differentielle Verfahren wie zum Beispiel Clusteranalysen ermitteln (Wang et al. 2016)
- unter Einbeziehung des Zustandes der Ereignisumgebung zum Ereigniszeitpunkt wie dem Inhalt der Website bestimmen.
- zu erfassen zu suchen in dem Ereignisse mit externen Quellen wie zum Beispiel Nutzerprofilaten gestellt werden.

Das einzig sinnvolle Verfahren im Kontext des zuvor beschriebenen Ziels ist eine Kombination mehrerer Ansätze sowohl auf der Ebene des Tracking als auch der Analyse. Ausgewertet werden soll wie zuvor umschrieben die Interaktion mit Wissen, also einem Gegenstand, der sich nicht mit einer Trägergröße allein wie z.B. der Website in Übereinstimmung bringen lässt und der per Definition kontextuell konstituiert ist. Wie eine solche Kombination aussehen kann, hängt natürlich von der technischen aber auch sozialen Umgebung ab innerhalb der evaluiert wird.

Eine Architektur zu Evaluation von Forschungsaktivität im Wissensspeicher

Konkret operiert der Use-Case Wissensspeicher im Bereich des Trackings mit einer Kombination aus:

- in das User-Interface hart kodierten expliziten Feedback-Möglichkeiten,

- ins User-Interface integrierte JavaScript snippets, die bei Interaktionen getriggert werden und diese dokumentieren,
- sowie Server-Log Dateien.

Ereignisse werden spezifiziert im Hinblick auf Ereignisgruppen (User Interface, Browser, Request und andere). User Interface Ereignisse werden weiterhin dahingehend kategorisiert auf welchem Seitentyp (Such Interface, Ressourcen Interface und andere) und welchem Layoutbereich sie angesiedelt sind. Schließlich werden für jedes Ereignis variable Eigenschaften bestimmt, die bei der Aufzeichnung dokumentiert werden müssen.

Die Bedeutungsgebung soll ebenfalls durch einen mehrstufigen auf einander aufbauenden Prozess stattfinden. Das Konzept sieht vor, Ereignisse in einem *Task-Model* (Yadav et al. 2015) einzuordnen, welches antizipierte Interaktionsprozesse innerhalb des Wissensspeicher formalisiert. Parallel dazu wird eine erste Auswertung von User-Activity-Daten wiederkehrende Ereignissequenzen identifizieren, die mit dem Task-Model verglichen werden. Dabei auftauchende Fragestellungen lassen sich dann in einem letzten Schritt mittels des *Thinking-Aloud* Verfahrens (Kuusela und Paul 2000) aus dem Bereich des *Usability Testing* bearbeiten. Hierbei arbeitet der Benutzer mit dem entsprechenden Angebot auf dem Computer und beschreibt was er tut, während er es tut.

Der Vortrag wird die vorangegangene Argumentation für den gewählten Ansatz nachzeichnen. Er wird darüber hinaus die ausgewählten Konzepte für die Generierung von User-Activity-Daten und ihrer Bedeutungszuschreibung im Kontext einer methodischen Evaluation von Forschungspraxis im Use-Case diskutieren. Zu guter Letzt wird ein Ausblick auf ein Modell gegeben werden, dass eine gegenseitige Bereicherung zwischen Ansätzen wie SDM und NeMO und dem vorgestellten aufzeigt.

Bibliographie

Benevenuto, Fabrício / Rodrigues, Tiago / Cha, Meeyoung / u. a. (2009): „Characterizing user behavior in online social networks“, in: *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*. ACM 49–62.

Calzarossa, Maria Carla / Massari, Luisa / Tessera, Daniele (2016): „Workload Characterization: A Survey Revisited“, in: *ACM Comput. Surv.* 48 (3), 48:1–48:43 10.1145/2856127.

Constantopoulos, Panos / Dallas, Costis / Bernadou, Agiatis (2016): „Digital Methods in the Humanities: Understanding and Describing their Use across the Disciplines“, in: Schreibman, Susan / Siemens, Ray / Unsworth, John (ed.): *A New Companion to Digital Humanities*. 1. Aufl. Chichester, West Sussex, UK: John Wiley & Sons.

Dang, Anh / Moh'd, Abidalrahman / Milios, Evangelos / u. a. (2016): „What is in a Rumour: Combined

Visual Analysis of Rumour Flow and User Activity“, in: *Proceedings of the 33rd Computer Graphics International*. ACM 17–20.

Dhawan, Mohan / Ganapathy, Vinod (2009): „Analyzing information flow in JavaScript-based browser extensions“, in: *Computer Security Applications Conference, 2009. ACSAC'09. Annual*. IEEE 382–391.

Gradmann, Stefan / Hennicke, Steffen / Tschumpel, Gerold / u. a. (2015): „Beyond Infrastructure! Modelling the Scholarly Domain“.

Hunter, Jane (2006): „Scientific models: a user-oriented approach to the integration of scientific data and digital libraries“, in: *VALA2006* 1–16.

Kuusela, Hannu / Paul, Pallab (2000): „A Comparison of Concurrent and Retrospective Verbal Protocol Analysis“, in: *The American Journal of Psychology* 113 (3): 387–404 10.2307/1423365.

o. A. (o. J.): DCU OnTo - NeDIMAH Ontology Navigation. <http://nemo.dcu.gr/> [Letzter Zugriff 26. August 2016].

o. A. (o. J.): Digitaler Wissensspeicher. <http://wissensspeicher.bbaw.de/> [Letzter Zugriff 26. August 2016].

Plumbaum, Till / Stelter, Tino / Korth, Alexander (2009): „Semantic Web Usage Mining: Using Semantics to Understand User Intentions“, in: Houben, Geert-Jan / McCalla, Gord / Pianesi, Fabio / u. a. (ed.): *User Modeling, Adaptation, and Personalization*. Berlin / Heidelberg: Springer (Lecture Notes in Computer Science) 391–396.

Suire, Cyrille / Jean-Caurant, Axel / Courboulay, Vincent / u. a. (2016): „User Activity Characterization in a Cultural Heritage Digital Library System“, in: *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries*. ACM 257–258.

Vozniuk, Andrii / Rodríguez-Triana, María Jesús / Holzer, Adrian / u. a. (2016): „Combining Content Analytics and Activity Tracking to Identify User Interests and Enable Knowledge Discovery“, in: *Proceedings of the 6th Workshop on Personalization Approaches in Learning Environments (PALE 2016)*. 24th conference on User Modeling, Adaptation, and Personalization (UMAP 2016), CEUR workshop proceedings.

Wang, Gang / Zhang, Xinyi / Tang, Shiliang / u. a. (2016): „Unsupervised Clickstream Clustering For User Behavior Analysis“, in: *SIGCHI Conference on Human Factors in Computing Systems*.

Yadav, Piyush / Kalyani, Kejul / Mahamuni, Ravi (2015): „User Intention Mining: A Non-intrusive Approach to Track User Activities for Web Application“, in: Tan, Ying / Shi, Yuhui / Buarque, Fernando / u. a. (ed.): *Advances in Swarm and Computational Intelligence*. Springer International Publishing (Lecture Notes in Computer Science) 147–154.