Evaluating Research Practices in the Digital Humanities by Means of User Activity Analysis

Niels-Oliver Walkowski
walkowski@bbaw.de
Berlin-Brandenburgische Akademie der Wissenschaften
Germany

Approaches to the Evaluation of DH Research Processes

The documentation of digital research processes has been a heavily discussed topic for some years now. It is most often addressed by the term *provenance*. In most cases, provenance data is created to record the chain of production of digital research results, in order to increase transparency in research and make such results reproducible.

In more recent times, a slightly different version of this topic has appeared in the field of digital humanities. Accordingly, digital research processes are modeled and documented with the aim to identify methodologies and practices of digital research in the arts and humanities on a broader scale. Two models have been introduced in this respect. One is the *Scholarly Domain Model* (SDM) (Gradmann et al, 2015) the other is the *NeDiMAH Method Ontology* (NeMO) (Constantopoulos, Dallas, and Bernadou, 2016). Both models provide formal semantics for the description of research processes as well as for their methodological evaluation.

The project environment in which these models were defined was dominated by European infrastructure projects, specifically *DARIAH* in the case of NeMO and *Europeana* in the case of SDM. Accordingly, such models aim to identify user needs and the qualitative use of infrastructure services as a first priority. However, it is also easily possible to refer to them in a broader perspective of *laboratory research* and *science studies*. For the case of the digital humanities community such a perspective corresponds with this community's wish to develop a unique methodological self-awareness.

Upon closer observation, the two models take up a different approach to accomplish their goals. In terms of NeMo, applications of the model describe research processes after they have taken place. In contrast, SDM makes reference to the concept of "modeling for" by Clifford Geertz and describes research processes in advance.

Such difference calls for a more concise evaluation of the terminology that was used before. In the research literature three concepts are used to distinguish between three possible viewpoints from which research processes can be described (Hunter 2006). These concepts are: workflow, provenance and lineage. As it has been indicated in the evaluation of SDM and NeMO the difference of these viewpoints is marked by the place in time from which they look at a research process. As such, it is also possible to call these concepts "prescribing", "inscribing", and "describing".

In accordance with this systematization SDM defines workflows while NeMO presents lineages. What is missing however, is real provenance data that is created and modeled in order to systematically evaluate digital humanities practices. More specifically, this means data which is recorded during the research process. The main goal of this presentation is to introduce an approach for how such provenance data can be created and modeled meaningfully to reach its goal by way of example. The example is the Wissensspeicher at the Berlin-Brandenburg Academy of Sciences and Humanities. The Wissensspeicher connects all digital resources of the academy in a way that lets the user interact not just with metadata but with parts of the content itself. The work which will be presented is part of the evaluation phase of the DARIAH-DE project that started in March 2016.

When evaluating research practices provenance data has some advantages in comparison to workflow or lineage data. On the one hand, it is easier to obtain very detailed data. On the other hand, semantic implications which might predefine the results of the evaluation are less necessary. For instance, in the SDM primer the concept of annotating exists before it is applied to an activity in a specific situation. In consequence, research results about digital annotation practices are biased. They depend on personal decisions of someone who classifies activities as annotating, or on a normative concept of annotating. For the identification of new research practices in digital environments, both aspects are problematic. Provenance data does not have the same risk because it is mostly created before classification takes place. The only aspect which is predefined is the fact that recordable actions take place and that these actions form part of a broader research context.

Nevertheless, the implementation of technological solutions for the creation of provenance data in these circumstance is more complicated than in common situations where provenance data is created. It is not enough to record which software component modifies a digital resource at a certain point of time as it happens in e-science. The "resource" is the research process itself, the actions take place on multiple levels and such actions are carried out not only by software but also by humans.

User Activity Analysis and Digital Humanities Research Processes

In fact, there is one field of research which addresses a comparable situation and this field is **user activity analysis.** In user activity analysis, human-computer-interaction is recorded in order to evaluate the user with respect to a specific research interest. The approach is used in areas like **e-commerce** and **online social networks** research in order to create services like recommendation systems (Plumbaum, Stelter, and Korth 2009) or to analyze social behavior (Dang et al. 2016). There are few examples of user activity analysis in academic digital environments. Suire et al. (2016) use this approach in the cultural heritage domain while Vozniuk et al. (2016) applies it to model learning processes in e-learning environments.

Having said that, no ready-made solution exists which can be easily used in the present context. Instead different approaches to user-activity analysis have to be evaluated in order to decide which ones can be adopted. Nevertheless, under the circumstances of evaluating digital research practices these decisions remain contingent. Digital research takes place in very different digital environments and under different conditions. Thus, in every situation in which digital research practice should be evaluated a different selection from the existing set of options might be the best. An overview of these options will be published in a DARIAH-DE report in the future.

The advantage of the Wissensspeicher use-case is the fact that it is a web platform– most user activity analysis takes place on websites and in web environments. There are two major tasks which need to be distinguished. The first task is user activity tracking and concerns how the data is created. The second task is the actual analysis. It demands to evaluate in which sense the created data constitute meaningful events and how to make sense out of these events.

Use–Case: Wissensspeicher

The Wissensspeicher implements user activity tracking by combining three different strategies: http-request logging, browser-event parsing and user annotations. Http-requests are logged by virtue of the *Django* request object and the *logger* library in the Python Django app that creates the website. Thereby request information can be pre-processed when it is detected. When a page is loaded in the browser a *JavaScript* client registers event listeners for page elements and certain user actions. Each event that is triggered causes the client to parse relevant information in the *DOM* of the *HTML* including microdata which has been created in the Django app in advance. Additionally, the user is able to directly give feedback in some situations. The created data is stored in a *MongoDB* database.

User activity analysis is also realized by virtue of three steps. In a certain way these steps resemble the three angles of workflow, provenance and lineage. First, events are evaluated in a so called *task model*. This task model describes ideal sequences of actions and user goals as conceived by the project employees. Second, users are evaluated by applying the **thinking aloud** technique from the field of **usability testing.** Finally, existing data will be evaluated to identify common event sequences by computing its clusters. A systematization of the results from these evaluations will enables researchers to associate certain meanings with events in such a way that the data can be analyzed to permit insights into research practices within the use case.

A Dialogue of Approaches

This presentation will summarize activities to evaluate research practices and methods in the digital humanities. It will outline a unique and complementary approach and indicate how this approach can be used in conjunction with existing digital humanities research practices. Finally, the implementation and results will be described up to the point that such results are present after two-thirds of the project time has elapsed.

Bibliography

Constantopoulos, P., Dallas, C., and Bernadou, A. (2016). "Digital Methods in the Humanities: Understanding and Describing Their Use Across the Disciplines." In *A New Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and

John Unsworth, 1st ed. Chichester, West Sussex, UK: John Wiley & Sons.

Dang, A., Moh'd, A., Milios, E., and Minghim, R. 2016. "What Is in a Rumour: Combined Visual Analysis of Rumour Flow and User Activity." In *Proceedings of the 33rd Computer Graphics International*, 17–20. ACM.

Gradmann, S., Hennicke, S., Tschumpel, G., Dill, K., Thoden, K., Pichler, A., and Morbidoni, C. (2015). "Beyond Infrastructure! Modelling the Scholarly Domain."

Hunter, J. (2006). "Scientific Models: A User-Oriented Approach to the Integration of Scientific Data and Digital Libraries." In *VALA2006*, 1–16.

Plumbaum, T., Stelter, T., and Korth, A. (2009). "Semantic Web Usage Mining: Using Semantics to Understand User Intentions." In *User Modeling, Adaptation, and Personalization*, edited by Geert-Jan Houben, Gord McCalla, Fabio Pianesi, and Massimo Zancanaro, 391–96. Lecture Notes in Computer Science 5535. Springer Berlin Heidelberg. doi:*10.1007/978-3-642-02247-0_42*.

Suire, C., Jean-Caurant, A., Courboulay, V., Burie, J.-C., and Estraillier, P. (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*, 257–58. ACM.

Vozniuk, A., Rodriguez-Triana, M.J., Holzer, A, and Gillet, D. (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, this volume.