

## Maßnahmen zur digitalen Nachhaltigkeit in Langzeitprojekten – Das Beispiel Capitularia

**Schulz, Daniela**

schulzd1@uni-koeln.de

Bergische Universität Wuppertal; Universität zu Köln

**Fischer, Franz**

franz.fischer@uni-koeln.de

Cologne Center for eHumanities

**Geißler, Nils**

nils.geissler@uni-koeln.de

Cologne Center for eHumanities

**Gödel, Martina**

mgoedel@uni-koeln.de

Cologne Center for eHumanities

Bei der Planung und Durchführung von langfristigen (Editions-)Projekten stellen sich hinsichtlich der Nachhaltigkeit besondere Herausforderungen, da sich die technischen Entwicklungen der nächsten Jahre und Jahrzehnte eben nur bedingt voraussehen lassen. Gerade in der Anfangsphase müssen aber bereits zahlreiche, oftmals richtungsweisende Entscheidungen getroffen werden, die den zukünftigen Erfolg oder Misserfolg, potenziell auftretende Probleme und Lösungsmöglichkeiten determinieren. Allerdings scheint es auch überhaupt nur in Projekten mit langer Laufzeit möglich, jenseits von reinen Willensbekundungen umfassende Strategien zur Nachhaltigkeit zu entwickeln und entsprechende Maßnahmen zu ergreifen. Damit leitet sich gleichzeitig die Pflicht ab, dies auch zu tun.

Am Beispiel des Projektes *Capitularia*, einem Langzeitprojekt (2014-2029), welches mit einer Hybrid-Edition frühmittelalterlicher Herrschererlasse befasst ist, sollen die verschiedenen Ebenen digitaler Nachhaltigkeit, damit verbundene Herausforderungen sowie erste Lösungsansätze präsentiert werden. Der von uns vorgeschlagene und in Teilen bereits umgesetzte Maßnahmenkatalog betrifft die Ebenen

- Datenmodellierung und Textauszeichnung,
- Datenhaltung und Dokumentation,
- Infrastrukturen (technisch und institutionell),
- Webseite und verwendete (Web-)Technologien,
- Präsentation, Zugänglichkeit und Nachnutzbarkeit der Forschungsergebnisse.

Bei der **Textauszeichnung** sollte auf etablierte Standards zurückgegriffen werden, um die programm- und plattformunabhängige Weiterverarbeitung der Daten langfristig zu gewährleisten. Bei *Capitularia* werden Transkriptionen, Handschriftenbeschreibungen, Register und bibliographische Daten gemäß der im Projekt erarbeiteten Richtlinien in *TEI-XML* codiert und durch ein den projektspezifischen Anforderungen angepasstes, restriktives Schema sowie den Einsatz von *Schematron* (ISO/IEC 19757-3) kontrolliert, um damit über die gesamte Projektlaufzeit hinweg und auch bei wechselndem Personal die Konsistenz und Einheitlichkeit und damit die Qualität der erstellten Daten zu sichern. (Hedler u.a. 2011: 11-12)

Dies bedingt von Beginn an eine **umfassende Dokumentation**, die nicht nur z.B. in Form eines Wikis vor allem für interne Zwecke verwendet wird, sondern auch möglichst viele Informationen öffentlich zugänglich macht, so dass andere Projekte von den Erfahrungen profitieren können. Die Entwicklung und Rationalisierung von Arbeitsprozessen sowie deren genaue Darlegung gewährleisten dabei neuen Mitarbeitern einen möglichst einfachen Einstieg. Für die öffentliche Dokumentation bieten sich neben der eigenen Projektwebseite beispielsweise Dienste wie *GitHub* an, auf denen gleichzeitig eine **transparente Datenhaltung** mit Versionierung möglich ist, und auch eigene technische Entwicklungen einfach zur Nachnutzung bereitgestellt werden können. Dies alles setzt natürlich die Verwendung entsprechender *Open Access* Lizenzen voraus.

Die **Sicherung der Langzeitverfügbarkeit** einer Ressource lässt sich generell nur durch eine entsprechende **technische Infrastruktur** (Daten-, Kompetenz-, Rechenzentren) in Kombination mit **institutioneller Anbindung** (Universitäten, Forschungsbibliotheken, kulturbewahrende Institutionen) gewährleisten. (Wissenschaftsrat 2011) Zu unterscheiden ist dabei zwischen der langfristigen Archivierung und Vorhaltung der Forschungsdaten und dem möglichst langen Erhalt der Webressource insgesamt, die ja ebenfalls mit allen Funktionalitäten über die Projektlaufzeit hinaus verfügbar sein soll. (Oßwald u.a. 2012: 13-15) Die Voraussetzungen in Köln erscheinen ideal, da dem Projekt mit dem *CCeH* ein im Bereich der DH ausgewiesener technischer Partner zur Verfügung steht, der durch enge Kooperation mit weiteren universitären Einrichtungen wie dem *Data Center for the Humanities* und dem Kölner Rechenzentrum, ideale strukturelle Voraussetzungen für Aufbau und langfristige Erhaltung digitaler Ressourcen schaffen konnte. Zusätzlich nimmt das *Capitularia*-Projekt auch am Webarchivierungsprogramm der Bayerischen Staatsbibliothek teil.

Für die Webpräsentation wurden in enger Zusammenarbeit mit dem technischen Partner **geeignete (open source) Technologien** ausgewählt, deren zukünftiger Erhalt sowie deren Weiterentwicklung von einer breiten Community getragen werden, und die sich in anderen Projekten bereits bewährt haben. Im Fall von *Capitularia* wird aktuell das PHP-basierte Content

Management System *WordPress* zur Verwaltung der Webpräsenz verwendet. Dabei werden die Standardfeatures (Blog, Suche, Mehrsprachigkeit etc.) möglichst breit genutzt und vereinzelt um weitere Eigenentwicklungen ergänzt, die insbesondere den Bereich des XSLT-Pipelining betreffen. Diese im Kontext des Projektes entwickelten *Plugins* sollen auch der Nachnutzung durch andere Projekte zur Verfügung stehen. WordPress erfüllt durch die freie Zugänglichkeit des Quellcodes, der aktiven und heterogenen Entwickler-Gemeinschaft, des Ökosystems, welches sich um diese Software etabliert hat, sowie der verwendeten Lizenzmodells wichtige Voraussetzungen digitaler Nachhaltigkeit. (Stürmer 2015, S. 36-37) Die bewusst gewählten *low-tech*-Lösungen gewährleisten weiterhin langfristig die leichte Wartbarkeit des Systems sowie plattform- und auch personelle Unabhängigkeit. Bei Bedarf könnte somit ohne größere Probleme auf ein anderes Präsentationsframework umgezogen werden.

Die **Präsentation der Forschungsergebnisse** findet auf mehreren Ebenen statt. Die angefertigten Transkriptionen werden zusammen mit weiteren Materialien projektbegleitend auf der Webseite veröffentlicht. Die Inhalte sind dort über Permalinks adressierbar und die zugrunde liegenden Daten (*XML*) stehen zum Download bereit. Um die Nachnutzung der Daten weiter zu erleichtern, ist die Implementierung von Schnittstellen (z.B. REST) vorgesehen. Die kritische Edition erscheint in Druckfassung in der *Leges*-Reihe der *Monumenta Germaniae Historica* und wird somit auch langfristig über deren Online-Angebot (dMGH) verfügbar sein. Vorabversionen der kritischen Editionstexte werden aber bereits zeitnah in digital aufbereiteter Form auf der Webpräsenz zur Verfügung gestellt.

Wenn auch finanzielle Ausstattung und Laufzeit von Projekten und damit auch deren Handlungsmöglichkeiten hinsichtlich Nachhaltigkeitsstrategien durchaus unterschiedlich sind, so erscheinen doch die Bereiche, in denen Maßnahmen getroffen werden können, allgemeingültig zu sein. Mit den hier vorgestellten Ansätzen soll ein aktiver und vor allem praxisorientierter Beitrag zur Diskussion um digitale Nachhaltigkeit geleistet werden.

## Fußnoten

1. Die freie Bereitstellung von im Rahmen von Digitalen Editionen entstandenen (XML-)Daten lässt aktuell noch zu wünschen übrig. Nur durch diese wird aber Nachnutzung und damit integrative und innovative Forschung erst ermöglicht. (Turska u.a. 2016: 1) Technische Schnittstellen können die Nachnutzung befördern und stellen daher ein Kriterium einer aktuellen Ansprüchen genügenden digitalen Edition dar. (Sahle u.a. 2014).

## Bibliographie

**Hedler, Marko / Montero Pineda, Manuel / Kutscherauer, Nico** (2011): *Schematron. Effiziente Business Rules für XML-Dokumente*. Heidelberg: dpunkt-Verlag.

**Oßwald, Achim / Scheffel, Regine / Neuroth, Heike** (2012): „Langzeitarchivierung von Forschungsdaten. Einführende Überlegungen“, in: Neuroth, Heike / Strathmann, Stefan / Oßwald, Achim / Scheffel, Regine / Klump, Jens / Ludwig, Jens (eds.): *Langzeitarchivierung von Forschungsdaten: Eine Bestandsaufnahme*. Boizenburg: Verlag Werner Hülsbusch 13–23 [http://www.nestor.sub.uni-goettingen.de/bestandsaufnahme/nestor\\_lza\\_forschungsdaten\\_bestandsaufnahme.pdf](http://www.nestor.sub.uni-goettingen.de/bestandsaufnahme/nestor_lza_forschungsdaten_bestandsaufnahme.pdf) [letzter Zugriff 9. November 2016].

**Sahle, Patrick / Vogeler, Georg / IDE** (eds.) (2014): *Kriterien für die Besprechung digitaler Editionen*. Version 1.1 Juni 2014 <http://www.i-d-e.de/publikationen/weitereschriften/kriterien-version-1-1/> [letzter Zugriff 12. November 2016].

**Stürmer, Matthias** (2015): „Wann sind Open Source Projekte digital nachhaltig?“, in: swissICT und Swiss Open Systems User Group (eds.): *Open Source Studie: Schweiz 2015* 36–37 <http://www.swissict.ch/fileadmin/customer/Publikationen/OSS-Studie2015.pdf> [letzter Zugriff 9. November 2016].

**Turska, Magdalena / Cummings, James / Rahtz, Sebastian** (2016): „Challenging the Myth of Presentation in Digital Editions“, in: *Journal of the Text Encoding Initiative* 9 <http://jtei.revues.org/1453> [letzter Zugriff 12. November 2016].

**Wissenschaftsrat** (ed.) (2012): *Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsstrukturen in Deutschland bis 2020*. Drs. 2359-12. Berlin <http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf> [letzter Zugriff 09. November 2016].