

Samuel Müller

Vorstellung NIE-INE, Düsseldorf, den 17. September 2018

Sehr geehrte Damen und Herren

ich freue mich sehr, Ihnen hier und heute zeigen zu dürfen, mit welchen Mitteln und welchem Vorgehen wir in der Schweiz versuchen, die Problematik der Langzeitsicherung und -verfügbarkeit digitaler Editionen zu lösen. Um die Komplexität der Aufgabe und die Facetten der Lösung benennen zu können, möchte ich zunächst einige grundsätzliche Erwägungen über das Umfeld, in dem eine Lösung stattfinden muss, anstellen, und im Nachgang in kurzen Strichen unsere konkrete Arbeit präsentieren.

Wenn wir über die Langzeitsicherung und -verfügbarkeit editionswissenschaftlicher Arbeit sprechen, bewegen wir uns in einem integrierten Umfeld, das verschiedene Akteure umfasst und von hohem wissenschaftspolitischem Gewicht ist. Nur wenn wir es schaffen, uns mit allen Akteuren in diesem Umfeld auf ein gemeinsames Vorgehen zu einigen und die einzelnen Zuständigkeiten zu benennen, werden wir erfolgreiche Lösungen anbieten können. Lassen Sie mich deshalb zunächst einen Blick auf dieses Umfeld werfen, um zu klären, welche Ansprüche und Pflichten hier gelten können.

Als Editionsphilologe nenne ich als ersten Akteur den **Forscher**, der eine Edition erarbeitet: den Editionswissenschaftler. Wie bei anderen Philologen besteht seine Arbeit eigentlich primär darin, seinen Lektüreprozess voranzubringen und diesen exakt zu dokumentieren. Dabei muss er theoretische Fragen nach Werk und Text stellen und diese allenfalls auch lösen können. Wie jeder Forscher besteht er darauf, in der Wahl seiner Mittel und seiner Argumente frei zu sein, um mit der nötigen Präzision seine Erkenntnisse entwickeln und präsentieren zu können. Diese Freiheit nutzt er in herkömmlichen Buchausgaben, indem er in unwahrscheinlicher Kreativität die üblichen Vorgaben des Buchsatzes erweitert und sich über eine Vielfalt von Marginalienspalten, Apparaten und Diakritika enorme Ausdrucksmöglichkeiten schafft. Dabei verhält es sich wie mit anderen wissenschaftlichen Publikationen auch: Die Verständlichkeit muss bisweilen hinter der Exaktheit des Ausdrucks zurücktreten. Leserfreundlichkeit kann nicht das vordringliche Ziel eines Editors sein. Der gewählte Ausdruck ist dabei aber immer Bedingung des Verstehens. Das bedeutet - und das ist eine zentrale Erkenntnis -, dass die gewählte Darstellung integrierender Bestandteil jeder Edition ist. Es ist dabei wichtig festzuhalten, dass der Prozess, der mit der Definition eines Corpus beginnt und

mit der Darstellung der Ergebnisse endet, nicht nur Corpus und Darstellung definiert, sondern seinerseits von diesen beiden abhängig ist. Das bedeutet, dass die für die Darstellung eingeforderte Freiheit gleichermassen eine Voraussetzung für den Arbeitsprozess darstellt.

Die digitale Welt ist für einen Editor deshalb eine Verheissung: Sie bietet ihm in beiden genannten Bereichen neue Freiheiten. Bereits seit etwa 70 Jahren nutzt er diese Freiheiten für seinen Arbeitsprozess. Und während der letzten drei Jahrzehnte hat er vermehrt dazu gefunden, sich von den materiellen Zwängen der seitengebundenen Darstellung zu lösen, um durch die enorme Flexibilität der Bildschirmpräsentation neue Blicke auf die beforschten Phänomene adäquat wiedergeben zu können. Diese beiden Freiheiten der Erarbeitung und der Darstellung sind wie gezeigt unabdingbar für jede editorische Tätigkeit. Nur wenn sie erhalten bleiben, kann digitale Edition funktionieren. Werden sie aber erhalten, erbringt das digitale Umfeld für den Editor einen grossen Mehrwert.

Gedächtnisinstitutionen (Museen, Bibliotheken, Archive), die zweite Gruppe der hier genannten Akteure, haben die Aufgabe, wissenschaftliche Ergebnisse für eine gegenwärtige und zukünftige Öffentlichkeit zu sichern und zugänglich zu machen. Dies gilt natürlich auch für die Ergebnisse der Editionsphilologen. Um die genannten Freiheiten des Editionsphilologen nicht zu beschneiden, dürfen die Gedächtnisinstitutionen ihm keine Vorgaben hinsichtlich seines Arbeitsprozesses und seiner Darstellung machen. Er muss sich darauf verlassen können, dass seine Ergebnisse in der von ihm festgelegten Form langfristig präsentiert werden. Diese schwierige Aufgabe setzt jedoch voraus, dass die Archivare in der Lage sind, selbst über die Bedingungen der Pflege der ihnen anvertrauten Ergebnisse zu wachen. Um ihrer Aufgabe gerecht werden zu können, müssen Sie folglich die Freiheit haben, über die eingesetzten Mittel und den mit diesen verbundenen Aufwand zu entscheiden.

Der Bedeutung dieser Feststellung liegt wiederum eine zentrale Erkenntnis zugrunde: Langzeitsicherung ist immer eine Frage von Institutionen und nicht von Technologien. Obwohl es häufig bestritten wird, gilt das auch für das Medium des Buches. Dass Bücher vergleichsweise stabile Datenträger sind, verdanken sie jahrtausendelanger Forschung und Entwicklung. Aber auch ihre ausgereifte Form setzt zum Erhalt institutionelle Strukturen in Form von Bibliotheken und Archiven voraus, die Umwelteinflüsse kontrollieren und die einzelnen Objekte pflegen. Und wenn wir statt auf die Bücher auf die in ihnen enthaltenen Texte blicken, wird sichtbar, dass die Pflege der materiellen Textträger für die Überlieferung hinter den Tätigkeiten des Abschreibens

und neu Darstellens zurücktritt: Die Stabilität des Datenträgers Buch allein reichte bei weitem nicht aus, um die Texte Homers und Platons, aber auch aller möglichen heiligen Schriften auf uns kommen zu lassen: Dazu brauchte es Institutionen, die sich des Erhalts dieser Texte annahmen und die Textträger bei Bedarf durch Abschrift oder Druck reproduzierten.

Auch der dritte Akteur im genannten Umfeld verlangt grosse Freiheiten. Es handelt sich um den **Nutzer** einer Edition. Sie gestatten mir die Verwendung dieses durchaus problematischen Begriffs. Auch ich bevorzuge als Editionsphilologe die Lektüre als adäquate Antwort auf meine Arbeit. Gleichzeitig habe ich es nicht in der Hand, zu steuern, was ein Nutzer mit meinen Produkten anstellt. Und ob bei einer analogen Publikation mein Buch gelesen wird oder dazu dient, das zu kurz geratene Bein eines Tisches zu unterlegen, kann ich nicht beeinflussen. Ob es mir passt oder nicht: daraus folgt die grundlegende Freiheit des Nutzers, mit meinen Produkten anzufangen, was und wie er es möchte.

Das bedeutet natürlich nicht, dass der Nutzer gegenüber dem Editor Ansprüche formulieren kann. Es ist nur so, dass sich die Freiheit des Editors, seine Produkte in der Form herzustellen, wie sie ihm angemessen erscheint, sich in der Freiheit des Nutzers spiegelt, einen angemessenen Umgang mit diesen Produkten selbst definieren zu können und zu müssen. Um ihn in die Lage zu versetzen, diese Freiheit zu nutzen, müssen ihm unsere Produkte an die Hand gegeben werden. Der Nutzer ist also auf die Vorhaltung und die Distribution unserer Produkte durch die Gedächtnisinstitutionen angewiesen. Gerade im Bereich der Distribution vermag die Digitalisierung viel zu leisten. Keine Bibliothek verschafft mir im analogen Bereich den gleichen instantanen und umfassenden Zugriff auf die Welt des Wissens, wie das World wide Web. Dieser elektronische Zugriff funktioniert aber nur, wenn der Nutzer in der Lage ist, die dargebotene Information seinerseits darzustellen und allenfalls weiter zu verwerten. Besonders die Frage der Nachnutzung stellt hier sehr hohe Anforderungen an die Datenqualität und -formate. Da diese Nutzer ihrerseits wieder Forscherinnen und Forscher sind und somit die Ergebnisse der Editionswissenschaftler eine Voraussetzung ihrer weiteren Arbeit darstellen, gerieren sich die Forschungsförderungsinstitutionen als Anwälte der Nachnutzer. Das führt zu einem institutionellen Druck auf die Form der präsentierten Ergebnisse im Sinne einer angenommenen weiteren Forschung.

Fassen wir also zusammen: die Problematik der Langzeitsicherung und -verfügbarkeit digitaler Editionen umfasst die Sicherung von Daten, Darstellung und Nachnutzung, wobei für die Erarbeitung, die Pflege und Vorhaltung und die Nachnutzung

gegebenenfalls unterschiedliche Technologien zum Einsatz kommen müssen. Jede unnötige Einschränkung und Standardisierung läuft Gefahr, die Waage in eine Richtung ausschlagen zu lassen und das Beziehungsnetz zwischen Editor, Gedächtnisinstitution und Nutzer empfindlich zu stören. Eine solche Störung hätte zwingend Qualitätseinbussen zur Folge.

Wie kann nun aber eine Lösung aussehen, die diesem heiklen Rahmen Rechnung trägt? Ich will Ihnen im folgenden unseren Schweizer Ansatz einer nationalen Infrastruktur für Editionen vorstellen. Die Nationale Infrastruktur für Editionen, NIE-INE, ist eine Instanz, die technologische und inhaltliche Vermittlung innerhalb des dargestellten Dreiecks vornimmt. Es handelt sich um eine von Vertretern aller drei Gruppen von Akteuren getragene und entwickelte Softwarelösung, die mit dem Service einhergeht, die nötige Vermittlung zu leisten. Die primären Ziele der NIE-INE können vor diesem Hintergrund wie folgt umrissen werden:

- a) die Bereitstellung einer quelloffenen, mehrsprachigen, individuell anpassbaren Arbeits- und Publikationsumgebung für Editionsprojekte vermittelt
 - der Erarbeitung generischer, aber hochdifferenzierter editionsspezifischer Basis-Ontologien,
 - der Entwicklung generischer und spezifischer Module zur Präsentation und Bearbeitung von Forschungsdaten innerhalb der Plattform und
 - der Gewährleistung des Im- und Exports von Daten aus in der Editionswissenschaft weit verbreiteten Datenformaten (z.B. TEI / XML),
- b) die Garantie langfristiger Sicherung und Verfügbarkeit der über die NIE – INE erarbeiteten und präsentierten Forschungsdaten und -ergebnisse,
- c) die Beratung und Betreuung der Editionsprojekte in Belangen der Sicherung und Präsentation digitaler Forschungsdaten und -ergebnisse und
- d) allgemeine Services, wie Use Cases und individuelle Anpassungen an Frontend und Ontologien.

Das an der Universität Basel angesiedelte Forum für Edition und Erschliessung (FEE) koordiniert die Entwicklung einer Arbeits- und Publikationsplattform zur Gewährleistung der langfristigen Sicherung, Verfügbarkeit und Publikation digitaler Forschungsergebnisse und der zugrundeliegenden -daten komplexer wissenschaftlicher digitaler Editionen. Diese Entwicklung fusst auf der engen Zusammenarbeit von 13

grossen Editionsprojekten, die von der germanistischen Mediävistik über den Atharvaveda, naturwissenschaftliche und musikwissenschaftliche Editionen bis hin zu modernen Literaturwissenschaft reichen. Diese inhaltliche und methodische Breite der beteiligten Editionsprojekte soll zu einem möglichst generischen Ansatz führen und es uns inhaltlich ermöglichen, eine allgemeine maschinenlesbare Theorie von Edition als Grundlage zu erarbeiten.

Zur Garantie der langfristigen Datenverfügbarkeit und -serialisierung nutzt die NIE – INE die bestehende Infrastruktur des Data and Service Center for the Humanities (DaSCH), einer Unternehmung der Schweizerischen Akademie der Geisteswissenschaften (SAGW) mit dem Auftrag, geisteswissenschaftliche digitale Forschungsergebnisse langfristig zu sichern und permanent zugänglich zu halten. Um den Kreis der Gedächtnisinstitutionen weit genug ziehen zu können, beteiligen sich zudem die Universitätsbibliothek Basel und die Zentralbibliothek Zürich beratend und unterstützend an den Entwicklungen.

Die Forschenden als Editionswissenschaftler und Nachnutzer werden durch die Universitäten Basel, Bern und Zürich und durch die Beteiligung des schweizerischen Nationalfonds zur Förderung der Wissenschaften (SNF) vertreten.

Für die Kommunikation mit den Editionsprojekten und insbesondere mit deren IT-Verantwortlichen sind in Basel, Zürich und Bern Toolkoordinatorinnen und -koordinatoren zuständig, die gleichzeitig die Anforderungen und Bedürfnisse in Entwicklungen am Frontend umsetzen.

Softwarearchitekten in Basel und Zürich kümmern sich um die Systemarchitektur, die Datenmodellierung, den Datenimport und die Koordination mit dem DaSCH.

Für Fragen der Langzeitarchivierung und des Datenmanagements wurde gemeinsam mit der Universitätsbibliothek Basel eine Beratungsstelle geschaffen.

Doch was sind die technischen Grundlagen einer solchen Entwicklung?

Die Entscheidung des DaSCH, zur Datensicherung auf eine Semantic-Web-Umgebung zu setzen, ermöglicht es uns, die erwähnten vielfältigen Ansprüche an Daten und Darstellung gewährleisten zu können. Die Daten der Editionsprojekte werden auf einem RDF-Triplestore als semantisch verknüpfte Informationen abgelegt. (Für das DaSCH handelt es sich standardmässig um die dortige Eigenentwicklung Knora, die Plattform der NIE-INE arbeitet jedoch – allenfalls mit Anpassungen – auch in anderen Umgebungen). Zu diesem Zweck wurden umfassende NIE-INE Basisontologien erarbeitet. Diese Ontologien beziehen sich auf Standardumgebungen von Cidoc und

FrBr. Das führt dazu, dass jede im Rahmen der NIE-INE erarbeitete Edition auf eine Vielzahl und bereits in semantischen Webumgebungen abgelegten Informationen zurückgreifen kann. Gleichzeitig können die Ergebnisse dieser Editionen ebenfalls leicht nachgenutzt werden. Die Nachnutzung schliesst explizit aktuelle Ansätze wie zum Beispiel Machine reasoning ein. Die Voraussetzung dieser Arbeit kann nicht der Wunsch nach einfacher Standardisierung sein. Die Erarbeitung der Ontologien muss immer auf einen Konsens hin zielen. Es handelt sich dabei um einen interaktiven Prozess, der nie zu einem endgültigen Abschluss kommen kann. Ergänzend zu diesem Vorgehen werden etwaige Bilddateien auf einem IIIF-Server gespeichert, von wo sie leicht eingebunden und weiter genutzt werden können.

Auf dieser Grundlage setzt das NIE – OS genannte Frontend auf. Aus verschiedenen Modulen kann der Nutzer sich hier eine eigene Arbeitsumgebung einrichten, mit deren Hilfe sich die Forschungsdaten flexibel zusammenstellen, bearbeiten und annotieren lassen. Weitere Module dienen der Web-Publikation. So können z.B. die Bilddateien in der Arbeitsumgebung und der Web-Präsentation angezeigt werden und diverse, auswählbare Suchoberflächen ermöglichen zum einen einfache Abfragen über verschiedene, flexibel anpassbare Suchfelder und zum anderen komplexe Suchmöglichkeiten über benutzereigene Abfragen. Bei Bedarf können Daten auch wieder in anderen, für spezifische Nachnutzungen erforderlichen Formaten (Text z.B. in TEI / XML) ausgegeben werden.

NIE – OS stellt projekteigene Workspaces und Webseiten zur Präsentation der Forschungsdaten und -ergebnisse zur Verfügung. Hierdurch wird die NIE – INE in die Lage versetzt, auf effiziente Weise projektspezifische Services anzubieten und eine durch zentrale Updates und Erweiterungen optimierte Wartung der Projekte vornehmen zu können.

Diese Infrastruktur steht grundsätzlich allen wissenschaftlichen Editionsprojekten in der Schweiz zur Verfügung. Die internationale Nachnutzung der erarbeiteten Software und der zugrundeliegenden Ontologien sind ohne Hürden nutzbar, da alle Entwicklungen Open Source erfolgen.

Was heisst das nun für ein einzelnes Editionsprojekt? (vgl. hierzu nie-ine.ch)

Ein einzelnes Editionsprojekt legt seine Arbeitsabläufe, seine Bedürfnisse an die Strukturierung der Daten und der Darstellung fest. Zur Gewinnung der Daten kann es je nach seinen Bedürfnissen direkt innerhalb der Arbeitsumgebung der NIE – INE arbeiten oder auf bestehende, übliche Werkzeuge ausserhalb der NIE – INE

zurückgreifen. Hierzu suchen wir die Kooperation mit möglichst vielen Entwicklerteams im Bereich der digitalen Edition. Anfang dieses Jahres wurde zum Beispiel eine enge Kooperation mit dem Trier Center for Digital Humanities (TCDH) vereinbart, um die dortige FUD-Umgebung einfach im NIE – INE-Kontext nutzen zu können. Im Fall, dass die Erarbeitung einer Edition weitgehend ausserhalb der NIE-INE erfolgt, werden die Daten mit den NIE – INE Basis-Ontologien sowie, falls erforderlich, mit projektspezifischen Ontologien gemappt und mit Hilfe dieses Mappings in den RDF-Triplestore importiert. Analog dazu werden Bilddateien über einen IIIF-Server eingebunden, wodurch sie auf einfache und variable Weise abrufbar sind. Über NIE – OS können sich die Projektmitarbeiter eine eigene Entwicklungsumgebung einrichten, in der Daten durchsucht, bearbeitet, annotiert und verlinkt werden können. Neben Modulen zur Datenbearbeitung stehen auch Präsentationsmodule zur Verfügung, mit deren Unterstützung eine Webseite zur online-Publikation der Daten erstellt werden kann. Die Daten und deren Darstellung können dann – soweit sie von den Eigentümern freigegeben wurden – auch von Dritten nachgenutzt und eingebunden werden.

Wo stehen wir in der Umsetzung?

Der Schwerpunkt in der ersten Projektphase liegt primär auf der Entwicklung einer Präsentationsplattform. Im ersten Projektjahr wurden hierzu die Grundlagen gelegt (u.A. Ermittlung der Bedürfnisse der beteiligten Editionen, Erstellung von Use Cases, Generierung von Basisontologien) und anhand des Pilotprojekts der Historisch-kritischen Online-Edition von Kuno Raebers Lyrik der Import und die Präsentation der Daten einer Modelledition umgesetzt. Parallel zur Website der Kuno Raeber Edition existiert nun die nahezu identische und teils erweiterte Präsentation der Daten durch NIE – INE (vgl. <http://raeber.nie-ine.ch>).

Bereits zum jetzigen Zeitpunkt sind die Datenstrukturen aller Projekte soweit modelliert, dass der Import eines Grossteils der Daten aller Projekte in diesen Wochen erfolgen kann. Parallel werden das NIE – OS und generische Module zur Präsentation der Daten weiterentwickelt. Eine Umgebung mit Benutzerverwaltung, in der diese Module zur Verfügung gestellt und angewendet werden können, ist im Aufbau, ebenso wie eine an die Bedürfnisse von Editionen angepasste und für den Einzelanwender spezifisch anpassbare Suchoberfläche. (Vgl. github.com/nie-ine)

Sollte ich mit meinen Ausführungen ihr Interesse an den Entwicklungen der NIE-INE geweckt haben, stellt sich natürlich die Frage in welcher Form Sie von unsere Arbeit



profitieren können. Inhaltlich kann man bei uns schon jetzt viel holen. Unsere Ontologien und das NIE-INE-Frontend (NIE-OS und die dazugehörigen Apps) können frei heruntergeladen werden (Das Frontend steht vermutlich Ende des Jahres zur Nutzung über Browser öffentlich bereit). Wir freuen uns über jede Nachnutzung und die damit verbundenen Weiterentwicklungen der Technologie. Die Ontologien lassen bereits heute wenig Wünsche offen und werden schnell weiter vervollkommen. (Wer keinen Aufwand mit Anpassungen möchte, kann übrigens auch die zugrunde liegenden Entwicklungen des DaSCH, KNORA und Sipi gleichfalls frei beziehen.) Zentraler aber scheint es mir zu sein, eine international integrierte Umgebung für digitale Editionsprojekte zu schaffen. Wir freuen uns deshalb über jede Kooperation und bieten gerne Hand zu grenzüberschreitenden Entwicklungen.