

Historisches Forschungsnetz. Ein offener Werkstattbericht

Thomas Meyer

Institut für Geschichtswissenschaften

Humboldt-Universität zu Berlin

Sitz: Friedrichstraße 191-193

D-10099 Berlin

Zusammenfassung: H-Soz-Kult, eines der zentralen geschichtswissenschaftlichen Fachinformationsangebote im deutschsprachigen Raum, hat eine umfassende Überarbeitung erfahren. Gefördert durch das DFG-Förderprogramm zu Virtuellen Forschungsumgebungen wurden basierend auf bibliothekarischen Standards neue Datenstrukturen, Schnittstellen und Suchtechnologien implementiert. Neben der Umsetzung von Nutzerwünschen und redaktionellen Anforderungen musste das Projekt sich in seiner technologischen Überarbeitung mit exemplarischen Problemen der Infrastruktur- und Softwareentwicklung im akademischen Umfeld auseinandersetzen.

Schlagworte: Virtuelle Forschungsumgebungen, Fachinformation, Software(entwicklung)

Historisches Forschungsnetz. An open „workshop“ report

Abstract: H-Soz-Kult is a information and communication forum for the humanities, which has undertaken a fully technological reorganization, which was funded by the DFG-Programme for Virtual Research Environments. The data model and interface were redesigned and completely moved into a structure based on library standards. The project was confronted not only with requirements both of the editorial boards and the user community, but also with the pitfalls of infrastructure and software development projects.

Keywords: Virtual research environments; Fachinformation?; software development projects

Der Einfluss neuer Technologien auf Forschungsmethoden ist in jüngster Zeit auch eines der großen Themen des Kinos und der Literatur. Sieht man vom teils fragwürdigen historischen „Wahrheitsgehalt“ von Filmen wie „Imitation Game“ oder der Mythenbildung in beispielsweise George Dysons Bestseller „Turings Kathedrale“ ab, zeigen diese immerhin die Ambivalenzen von Wissenschaft, ihren Akteuren und der „Einführung“ neuer Technologien. Nicht ganz so mythisch

aufgeladen kreisten in den vergangenen Jahren Diskussionen um das scheinbar neue Konzept der Virtuellen Forschungsumgebungen (VFU). Diese VFUs sollten – möglichst umfassend – IT-gestützte Forschungs-„Tools“ in Umgebungen abbilden, die einer breiten wissenschaftlichen Nutzergemeinde bzw. „Community“ zugänglich sind, Forschungsdaten unterschiedlicher Herkunft integrieren und mittels darin eingebetteter Analysetools zu neuen methodischen Ansätze führen. Ein Paradigma, welches mittlerweile auch die Digital Humanities an und für sich beanspruchen. Nun können Technologien sicherlich als Innovationsmotor eine wesentliche Rolle spielen, ein flüchtiger Blick in die Wissenschafts- und Technikgeschichte zeigt eindrucksvolle Beispiele. Allein durch Proklamationen sind diese nicht zustande gekommen: Gerade in der Wissenschaft bedarf es vielmehr williger Akteure, die neue Forschungstechnologien – über Disziplinengrenzen hinweg – auch durch soziale Interaktion und Innovation verbreiten.¹

Ein Projekt, welches im Rahmen der DFG-Förderprogramme zu VFUs zwar nicht neu für eine „Community“ geschaffen wurde, sondern eher selbst technischer Innovation bedurfte, ist H-Soz-Kult, das Online-Forum für die Geschichtswissenschaften. Aus dem einstigen Mailinglistenableger des Humanities-Net – kurz H-NET – der Michigan-State-University im Jahr 1996 mit der ursprünglichen Ausrichtung auf Sozial- und Kulturgeschichte, gestartet mit einer Handvoll Subskribenten, hat sich mittlerweile eines der größten geisteswissenschaftlich-interdisziplinären Foren im deutschsprachigen Raum entwickelt. Informationen über das Geschehen in Forschung und Lehre, über Karrieremöglichkeiten bis hin zu wissenschaftlichen Publikationen sind über konfigurierbare Mailinglisten, neue Suchtechnologien sowie verschiedene Informationskanäle wie RSS-Feeds oder neuerdings twitter zugänglich.

Zum Zeitpunkt, als das Thema VFU um 2009 herum einen ersten Höhepunkt in Deutschland erreicht hatte, befand sich H-Soz-Kult aufgrund seiner überholten technologischen Basis gegenüber jüngeren Projekten der Fachinformation und –kommunikation sichtlich im Rückstand. Mit der Durchsetzung des World Wide Web hatte H-Soz-Kult frühzeitig seine Präsenz genau dort gefunden, mit den technologischen Möglichkeiten, die seit Ende der 1990er Jahre zur Verfügung standen: SQL-Datenbanken und einfache Skriptsprachen beherrschten damals die Szenerie, komplexe Content-Management-Systeme steckten noch in den Kinderschuhen, von mobilen Endgeräten ganz zu schweigen. Die Inhalte der Mailingliste wurden zwar von Beginn an im WWW zugänglich gemacht, nur war die Automatisierung und Flexibilisierung von Arbeitsabläufen/Workflows, die Abbildung komplexer Datenstrukturen/Metadaten und Rezeptionsmöglichkeiten über neue Endgeräte, vor

¹ Unter anderem zuletzt diskutiert auf einer technikgeschichtlichen Jahrestagung, siehe Tagungsbericht: Research technologies - Forschungstechnologien. 94. Jahrestagung der Deutschen Gesellschaft für Geschichte der Medizin, Naturwissenschaft und Technik, 23.09.2011 – 25.09.2011 Stuttgart, in: H-Soz-Kult, 22.10.2011, <<http://www.hsozkult.de/conferencereport/id/tagungsberichte-3857>>

allem aber die Einbettung sowohl externer Daten wie die Vernetzung mit den eingesetzten IT-Technologien der früheren Jahre kaum noch zu leisten.²

Die Website von Soz-Kult zählt neben den mittlerweile mehr als 23.000 Mailinglisten-Abonnements zum Kern des Projekts: Mehr als monatlich 250.000 Leser auf Website und RSS-Feed und 1 Mio. Seitenabrufe in 2014 zeugen vom anhaltenden Wachstum der Leserschaft. Der Umfang an Beiträgen hatte ab Mitte der 2000er Jahre bereits ein Maß erreicht, welches die Möglichkeiten der bis dahin gängigen Rezeptionsmöglichkeiten vor allem in Form der Einzelmails überschritt: im Bereich der Veranstaltungsankündigungen, Tagungsberichte und Stellenausschreibungen allein hat sich seitdem das Beitragsvolumen verdoppelt. Auf technischer Ebene wurden bereits 2005 erste Abo-Optionen im Kontext des Clio-online-Projekts zur Ausdifferenzierung des Angebots entwickelt. Auf redaktioneller Ebene wurden Fachredaktionen für einzelne Themengebiete oder Epochen gegründet, um die Breite des wissenschaftlichen Publikationsgeschehens abzudecken. Dem Wachstum vor allem ab 2007/2008 war damit allein nicht zu begegnen, waren doch die Zunahme an Beiträgen und die wachsende Leserschaft nicht nur durch die historische Zunft bedingt. Vielmehr wird H-Soz-Kult durch weitere geisteswissenschaftliche Disziplinen genutzt. Interdisziplinäre Vorhaben, sicherlich auch beflügelt durch die Exzellenzinitiative des BMBF oder zahlreiche interdisziplinäre SFBs, haben ihren Teil zur weiterhin steigenden Popularität beigetragen. Zudem fehlte oder fehlt immer noch in anderen geisteswissenschaftlichen Disziplinen ein entsprechendes Pendant, sodass die Attraktivität von HSK auch in den Philologien oder den Sozialwissenschaften ungebrochen ist.

Eine weitere Diversifikation redaktioneller Strukturen, die stärkere Einbeziehung der Nutzer und ihrer Bedürfnisse sowie eine technische Generalüberholung standen somit auf der Agenda. Nutzer und Redaktion waren zielgerichteter in die Erschließung und Veröffentlichungsprozesse einzubinden, die Arbeitswerkzeuge auf verteilte Redaktionen an verschiedenen Standorten auszurichten und schließlich die Erschließung und Zugänglichkeit zu verbessern. Letzteres zum Beispiel mittels neuer Konfigurationsoptionen oder facettierter Suchen. Letztlich wuchs nach fast zwanzig Jahren des Bestehens ein Archiv des geschichtswissenschaftlichen Geschehens heran, dessen Informationen bisher unzureichend strukturiert waren, um sie später beispielsweise digital auswerten zu können. Eine technologische Neuaufstellung beinhaltete auch die Beseitigung infrastruktureller Desiderata bzw. eine stärkere arbeitsteilige Verantwortung für Basisdienste gemeinsam mit Rechenzentren oder die bessere Anbindung an Archiv- und Bibliotheksdienste. Bereits im Clio-online-Projekt waren solche Ansätze erprobt und in Teilen umgesetzt worden, u.a. mit verschiedenen Verzeichnisdiensten oder einer Metasuche. Die Idee der VFUs – verschiedene Arbeitswerkzeuge und Datenstrukturen

² Die ersten, damals noch statischen HTML-Seiten von H-Soz-Kult können im Web-Archive eingesehen werden unter <https://web.archive.org/web/19980130083245/http://hsozkult.geschichte.hu-berlin.de/> (25.06.2015)

zusammenzuführen – bot sich somit für eine strukturelle und technische Neuaufstellung von H-Soz-Kult an, mit einem Ansatz, der die Kerntätigkeiten der Historiker – das Schreiben, Erschließen, Kommunizieren und Publizieren – in den Mittelpunkt stellte.

Ziel des DFG-geförderten Projekts „Historisches Forschungsnetz“ war daher die Entwicklung eines modularen Frameworks basierend auf Open-Source-Komponenten, in dem Bausteine für die Erschließung und Veröffentlichung gängiger Fachinformationsformate sowie weitere Tools für deren Distribution bereitgestellt werden. Ein neues Datenmodell, welches die Modularität der einzelnen Informationstypen, deren Verknüpfung, die Anbindung an gängige Informationssysteme von Archiven und Bibliotheken unterstützt, wurde mit in der Bibliothekswelt gängigen Standards geschaffen: mit FRBR und RDA wurden die wesentlichen Metadatenstrukturen und Verfahren etabliert, mit der sich bibliografische Einheiten erschließen und in automatisierte Abläufe einpassen lassen, diese bilden nunmehr das Grundgerüst des Datenmodells von H-Soz-Kult.

```

<type name="frbr:Manifestation">
  <title>FRBR Manifestation</title>
  <parent>frbr:Expression</parent>
  <properties>
    <property name="frbr:complementOf">
      <title>A work or expression that is complemented</title>
      <type>d:text</type>
      <index enabled="true">
        <tokenised>both</tokenised>
      </index>
    </property>
    <property name="rdaRoles:author">
      <title>Authors Name</title>
      <type>d:text</type>
      <mandatory>false</mandatory>
      <multiple enforced="false">true</multiple>
      <index enabled="true">
        <atomic>true</atomic>
        <stored>true</stored>
        <tokenised>FALSE</tokenised>
      </index>
    </property>
    <property name="rda:titleManifestation">
      <title>Title</title>
      <type>d:text</type>
      <index enabled="true">
        <tokenised>both</tokenised>
        <!-- to make it orderable -->
      </index>
    </property>

    [...]

  </properties>
</type>

```

Abbildung 1. Ausschnitt aus dem H-Soz-Kult Datenmodell mit FRBR, RDA und MODS

<Datei: datenmodell_hfn.png>

Neben bibliografischen Metadaten waren außerdem Personen, Organisationen und Ereignisse abzubilden. Durch Anreicherung mit weiteren Vokabularien (u.a. bibo, FOAF, geonames, Linked Data Service der DNB, MarcOnt, vcard) unterschiedlicher Provenienz konnten diese Entitäten und ihre Eigenschaften modelliert werden, nur wenige Ausnahmen wurden mit eigenen neuen Eigenschaftsmodellen aufgefüllt. Ergänzt wurde das Modell durch Anleihen bei gängigen Industrie-Standards, um Informationen zu Arbeitsabläufen/Workflows abzubilden. So steht mit dem BPMN – Business Process Model and Notation³ – ein Modell bereit, in dem Zustände von Objekten, die damit betrauten Personen sowie deren ihre Einbettung in Abläufe beschrieben und Prozesse automatisiert

³ <http://www.bpmn.org/>

werden können. Eine Vielzahl an Tools bzw. Workflowengines für gängige Programmiersprachen ermöglicht die Einbettung in eigene Applikationen. Allein mit dem „Vokabular“ von BPMN lassen sich eigene Informationen für die Abbildung von Workflows in

```
<type name="cliowf:WorkflowTask">
  <parent>bpmn:Task</parent>
  <properties>
    <!-- Task Context -->
    <!-- e.g. Space, Document -->
    <property name="cliowf:context">
      <type>d:noderef</type>
    </property>
    <!-- Task Outcome -->
    <property name="cliowf:outcome">
      <!-- NOTE: Transition id -->
      <type>d:text</type>
    </property>
    <!-- Items within package marked as complete -->
    <property name="cliowf:completedItems">
      <type>d:noderef</type>
      <multiple>true</multiple>
    </property>
  </properties>
</type>
```

Datenmodellen unterbringen.

Abbildung 2. Metadaten für Workflows: Eigene Task-Objekte auf BPMN-Basis

<Datei: workflowmodell_hfn.png>

Für das HFN-Projekt waren damit Formate und Vokabularien eingeführt, in die vorhandene Daten und Abläufe übertragen, teils sogar in ihrer Abbildung verbessert werden konnten. So werden Rezensionen im neuen Datenmodell nicht als singuläre Einheiten aufgefasst, sondern in einzelne Entitäten „Buch / Website“ – „Rezensent“ – „Verlag“ – „Besprechung“ und ihre jeweiligen Relationen zueinander zerlegt. Was für Bibliothekare hier auf den ersten Blick banal klingen mag – die Zerlegung in einzelne informatorische Einheiten – ist in der Praxis vieler wissenschaftlicher Projekte und Publikationen nicht unbedingt selbstverständlich.

Neben einer weitaus performanteren Behandlung der Entitäten ist diese Zerlegung Grundvoraussetzung für zukünftige Kooperationen und arbeitsteilige Vorgehensweisen. Denkbar sind einfachere Zugriffe auf Erschließungsinformationen aus Bibliotheken, ergänzende Materialien von Verlagen („Organisationen“) und die Verknüpfung mit strukturierten Personendaten (Autoren, Rezensenten). Mit der GND sowie Personennormdatei stehen schließlich zwei Werkzeuge zur

Verfügung, die mit den Entitäten des neuen Datenmodells in Beziehung gesetzt werden können. Eine Anreicherung mit bspw. Kataloginformationen entlastet Abläufe in einer Redaktion, das Redaktionsgeschäft kann stärker auf die Redaktionsbelange ausgerichtet werden, da weniger Zeit und Personal in die Erschließung und Aufbereitung der bibliografischen Informationen der besprochenen Werke zu investieren ist – auch dies leider eine noch oft anzutreffende Praxis in Zeitschriftenredaktionen.

Für die Vernetzung selbst wurden im Umfeld von Bibliotheks- und Content-Management-Systemen (CMS) gängige Schnittstellen evaluiert. Neben bekannten Standard wie OAI, Z39.50 oder SRU steht mit CMIS (Content Management Interoperability Services) eine weitere Schnittstellendefinition aus der kommerziellen Softwarewelt bereit, die den Austausch von Inhalten (Dokumente/Objekte) aus unterschiedlichen Datenspeichern über normierten Verfahren ermöglicht. Über Such- und Browsingverfahren können CMS wie z.B. Sharepoint, Alfresco, Liferay oder Drupal, um nur einige repräsentative Beispiele zu nennen, miteinander vernetzt werden. Der Austausch ist dabei nicht auf das Retrieval beschränkt, sondern bezieht das „Schreiben“ mit ein. Eine Umsetzung für H-Soz-Kult wurde in einem Alfresco-Repository prototypisch implementiert: Eine Java-Bibliothek (core) übernimmt als Serviceschnittstelle die Kommunikation zwischen CMIS-Server am Dokumentenrepository und dem jeweiligen Client (z.B. ein Plugin für die Suche und Anzeige in einem CMS wie z.B. Drupal oder ein separates administratives System für die Verwaltung von Beiträgen und Metadaten).

Repositorien

Nun sind Repositorien keineswegs eine Erfindung der Bibliothekswelt, die Terminologie ist auch in der Welt der Content-Management- und Dokumenten-Management-System allgegenwärtig. Die „Repositories“ der Bibliothekswelt fokussieren wie ihre Pendants im Content-Management auf bestimmte Anwendungsfälle, meist sind sie technisch auf den Bedarf von Projekten ausgelegt, die „Sammlungen“ erschließen und präsentieren. Eines der gängigen Repositorien – FEDORA – wurde im HFN-Projekt für die Nutzung untersucht. Auch Fedoras Schwerpunkt liegt, wie man heute noch auf der Website lesen kann, auf „Collections“. Fachinformationsforen wie H-Soz-Kult, die sehr heterogene und vor allem auch fluide Formate anbieten, lassen sich in solche Sammlungsschemata nur bedingt einfügen, wie u.a. Tests mit größeren Datenbeständen ergaben. Nachdem zuerst eine kleine Auswahl eigener Beiträge in ein Repository importiert und mit den gängigen Fedora-Clients darauf CRUD-Operationen (create, read, update, delete) ausgeführt wurden, ist der Testlauf auf größere Datenmengen ausgeweitet worden. Diese ergaben auffällige Performanceverluste – von bspw. mehr als 80.000 zu importierenden Rezensionen der Rezensionsdatenbank „Historische Rezensionen Online“ waren nach 12 Stunden Testlauf gerade einmal 25% der Daten eingefügt (Ziel

des HFN-Projekts war es, alle Teildatenbanken, die im Umfeld von Clio-online und H-Soz-Kult genutzt werden, in eine gemeinsame Struktur zu überführen). Abwägungen, ob die zur Verfügung stehenden zeitlichen und personellen Ressourcen in die Behebung von Performanzproblemen zu investieren sind, verbunden mit weiterer Einarbeitung und Expertisebildung, mündeten schließlich in der Entscheidung, stattdessen auf bekannte bzw. erprobte Verfahren zurückzugreifen, die von allen Projektbeteiligten beherrscht werden.

Um die Tür zu FEDORA allerdings nicht zu verschließen, wurde weiterhin die Entwicklung des Projekts Bamboo in den USA beobachtet, welches mit „HUBZero“ eine Webplattform entwickelte, die einige der im HFN-Projekt definierten Schnittstellen unterstützte. In „HUBZERO“ wurde u.a. die Implementierung des Dokumentenaustausch über eine CMIS-Schnittstelle mit einem FEDORA-Repository vorangetrieben; das hätte zu diesem Zeitpunkt Architekturen ermöglicht, die sowohl CMS-/DMS-Systeme nutzen, als auch Schnittstellen und Speichersysteme der Bibliothekswelt einschließen. Fertiggestellt wurde diese Funktionalität in Bamboo leider nicht, mittlerweile ist das Projekt auch abgeschlossen.⁴

Ein weiterer Exkurs auf gängige NOSQL-Datenbanken, welche in letzter Zeit durch das Thema „BIG-DATA“ recht populär geworden sind, sollte deren Verwendbarkeit evaluieren. Die Abbildung der ‚realen Welt‘ in NOSQL-Datenbanken ist eingeschränkt, sie enthalten keine Relationen bzw. geben eine Datenstruktur vor, in der Relationen eher durch sog. horizontale Verknüpfungen von Datensätzen abgebildet werden. Diese Verknüpfungen werden in einer Baumstruktur abgebildet, deren Verarbeitung mit den Mitteln der Informatik jedoch recht aufwändig ist. Vor allem Applikationen mit großen Datenmengen stoßen schnell an ihre Leistungsgrenzen, sodass auch diese Datenspeicher für eine skalierbare Umgebung im Projekt verworfen wurden.⁵

Weitaus performanter und in ihrer Wartung und Erweiterbarkeit nachhaltiger haben sich im HFN-Projekt klassische SQL-Datenbanken erwiesen. Deren Verbreitungsgrad, ihre Verfügbarkeit in universitären Rechenzentren und die Nutzung in akademischen wie kommerziellen Projekten ist ungebrochen hoch. Verknüpft mit performanten Indexierungstools wie beispielsweise SOLR-basierten Suchmaschinen und gängigen Programmierframeworks lassen sich mit SQL zügig einfache Anwendungen erstellen und in vorhandene Dokumenten- und Content-Management-Systeme integrieren. Für das HFN-Projekt wurden daher einige Entwicklungsstränge verändert – ein oftmals

⁴ Ein Großteil der Dokumentation ist unter <http://www.projectbamboo.org/> archiviert.

⁵ NOSQL-Datenbanken können zwar hinsichtlich der Menge sehr große Datenbestände verwalten, schwer aber nur stark strukturierte Daten bzw. in viele kleine einzelne Entitäten gegliederte Daten, die miteinander in Relationen stehen. Ein detaillierter Bericht hierzu findet sich auf dem HFN-Blog unter <https://www2.hu-berlin.de/historisches-forschungsnetz/2012/03/jcr-repositorien-und-nosql-datenbanken/> sowie <https://www2.hu-berlin.de/historisches-forschungsnetz/2012/03/datenbanken-und-repositorien-teil-2-fedora-commons-nuxeo-und-alfresco/>.

vernachlässigter und ausbleibender Schritt in akademischen Projekten.⁶ Die Implementierung des Datenmodells und der Schnittstelle zur Verknüpfung von Datenspeicher und Anwendungsoberflächen (Websites, Redaktionsumgebung) wurde innerhalb kurzer Zeit in eine neue Zwischenschicht auf der alten Datenbank verlagert. Die von vornherein geplante lose Verknüpfung von Datenschicht und Anwendungsschichten über eine CMIS-Struktur erwies sich in dieser Stelle als richtig, da in dieser offenen Architektur nun fortlaufend einzelne Bausteine (Datenspeicher, Öffentliche Anwendungen/Websites, Internes Bearbeitungs- und Redaktionssystem) schrittweise weiter ausgebaut werden können, unabhängig von der Entscheidung, welches Content- oder Dokumenten-Management-System oder Datenbank als Datenspeicher endgültig zu verwenden sein wird.

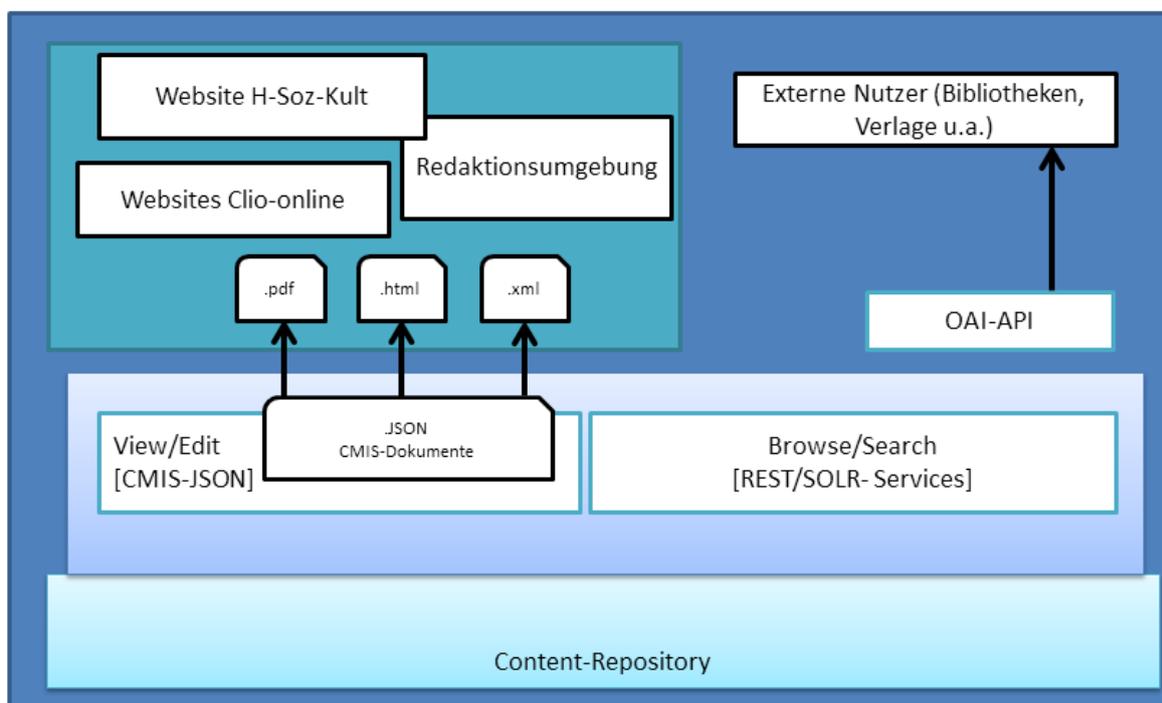


Abbildung 3. Architektur der HFN-Module

<Datei: architektur_hfn.png>

Im Herbst 2014 wurde die Website von H-Soz-Kult mit neuen Funktionalitäten freigeschalten. Die Site basiert auf dem neuen Datenmodell, als auch einer Auswahl der geplanten Schnittstellen, wie beispielsweise der umfangreich implementierten SOLR-Suchmaschine, auf der sämtliche Browsing und Suchfunktionalitäten aufsetzen, wie auch die Ausgabe der Dokumente in einer datenbankunabhängigen JSON-CMIS-Struktur, die in beliebige weitere Ausgabeformate (xml, html,

⁶ Selbst Untersuchungen zu Erfolgskriterien entsprechender Softwareprojekte gehen leider immer noch von einer statischen Systematisierung von Entwicklungsprozessen und Arbeitsweisen von Wissenschaftler/innen aus. Stattdessen ist eine fortlaufende Evaluierung einzelner Umsetzungsschritte notwendig, die rechtzeitig eine Neuausrichtung von Projekten ermöglicht. Im Softwarebereich besteht mit „SCRUM“ ein Modell für eine solche iterative Vorgehensweise.

pdf) konvertierbar ist. Ein noch weiter bestehendes Desiderat ist die geplante Einführung von XML-Dokumententypen für die Volltextfassung bzw. –erschließung. Bisher existiert weder auf dem kommerziellen Markt noch in der akademischen Welt eine zufriedenstellende Lösung (Editor), mittels derer sich wie in gängigen HTML-Editoren wie TinyMCE oder CKEditor leicht die Auszeichnung von Volltexten nach wissenschaftlichen Anforderungen für die Erschließung und für die automatisierte Verarbeitung vornehmen lässt. Gängige Implementationen wie beispielsweise von Eclipse-Umgebungen oder XML-Editoren wie Oxygene sind weder einer mehr als fünfzigköpfigen Redaktion noch mehreren tausend Autoren zuzumuten.⁷

Ausblick

Der Prozess der technologischen Neuaufstellung ist damit nicht abgeschlossen. Derzeit befinden sich Teile des Redaktionssystems noch in Überarbeitung, vor allem die Unterstützung konfigurierbarer Distributionsformate und –kanäle unterliegt weiteren Veränderungen: Der „Vertrieb“ der Beiträge als Emails war und bleibt eine der zentralen Säulen von H-Soz-Kult, auch wenn mittlerweile Beiträge in twitter oder per RSS-Feed abgerufen werden können. In den vergangenen Jahren jedoch verschärfen kommerzielle wie auch akademische Email-Hoster/Provider ihre SPAM-Schutzmechanismen aus nachvollziehbaren Gründen. Das wiederum hat Auswirkungen auf den „Vertrieb“ von Emails gerade über Mailinglisten. Dieses Gemengelage einschließlich der weiterhin stark wachsenden Subskribentenzahlen und der Menge an Beiträgen führt zu einem erhöhten Konfigurations- und Wartungsaufwand, die Kommunikation der Listenserver muss viel stärker auf einzelne Empfänger(server) angepasst werden. Für das H-NET rückt Email daher derzeit als primärer Distributionskanal in den Hintergrund, das „Announcement“ auf der Website gewinnt stärker an Bedeutung, Emails werden in erster Linie als „Alerts“ zu neuen Beiträgen versendet (für H-Soz-Kult bleibt Email einer der zentralen Distributionskanäle).

Schon dieses Beispiel zeigt, dass Technologien einem Nutzungswandel und teils eingeschränkten Lebenszyklen unterliegen. Gleiches gilt für Datenbanksysteme, Indexierungstools und Suchmaschinensoftware bis hin zu Design- und Layoutstandards im Print- und Onlinebereich. Fortlaufende Anpassungen richten sich nicht nur an den „Bedürfnissen der Community“ aus, sondern sind auch durch technische Rahmenbedingungen determiniert. An dieser Stelle gewinnt die Webstatistik an Bedeutung, um Funktionalitäten und Inhalte hinsichtlich ihrer realen Nutzung und Nachfrage zu überprüfen. Die bei H-Soz-Kult anonymisierte Erhebung von Zugriffszahlen der einzelnen Rubriken wird nun durch Erhebungen zur Nutzung der neu eingeführten Facettierungen der Suchen ergänzt: Diese werden erfreulicherweise in den Suchoptionen nachweislich am stärksten

⁷ Eine Aufstellung der evaluierten Doctypes ist zu finden unter <https://www2.hu-berlin.de/historisches-forschungsnetz/2012/08/ubersicht-xml-doctypes-fur-historisches-forschungsnetz/>

genutzt, im Bereich Rezensionen vor allem die Aufstellung der thematischen Redaktionsbereiche, der thematischen Klassifikation sowie der besprochenen Publikationstypen und –formate (Monografie, Sammelband, Quellensammlung, Nachschlagewerk etc.)

Zuletzt misst sich die Entwicklung digitaler Werkzeuge natürlich an den Bedürfnissen der „Community“. Für H-Soz-Kult sind das einerseits die wachsende Leserschaft und deren Bedarf an Inhalte-geleiteten Rezeptionsmöglichkeiten und andererseits die Redaktion, welche mit möglichst einfachen Mitteln den Workload redaktioneller Arbeit – die im Übrigen unentgeltlich neben der eigentlichen Tätigkeit in Forschung und Lehre stattfindet! – bewältigen können muss. Hierfür sind einfache und transparente Verfahren notwendig, die nicht zwangsläufig auf komplexen Funktionen aufsetzen müssen. Umfassende Workflow-Frameworks wie das im Alfresco-Repository enthaltene Activity sind zu komplex und für den redaktionellen Gebrauch überdimensioniert. Stattdessen wurden in H-Soz-Kult überschaubare Optionen und Funktionen im Redaktionssystem implementiert, mit denen sich schnell ein Überblick über das Redaktionsgeschehen erfassen lässt, die aber auf dem BPMN-Workflow-Modell basieren. Diese, wie auch die gängigen Basistechnologien (SQL, Suchmaschinenindexierung, Schnittstellen wie OAI und CMIS) waren leicht zu implementieren. Schließlich sind bereits das Know-How rund um diese Technologien und die Technologien selbst bei Projektträgern und Partnereinrichtungen (wie z.B. der Computer- und Medienservice der HU und dessen Datenbankhosting) vorhanden.

Nachhaltigkeit

Es war einer der Grundannahmen im HFN-Projekt, von vornherein an Lösungen anzuknüpfen, die nicht nur im akad. Umfeld, sondern großflächig verbreitet sind, sodass die Pflege von Updates, die Implementierung von Features usw. über mehrere Jahre hinweg auch gewährleistet bleibt. Trotz aller Bestrebungen zur Etablierung von Kompetenzzentren wird Software im akademischen Umfeld nur selten über umfassende Strukturen gepflegt, die mit ausreichend finanziellen und personellen Mitteln ausgestattet sind für viele Jahre mit Wartung und Ausbau beschäftigen können. Verschiedene Finanzierungs- und Beteiligungsmodelle sind zwar erprobt und eingeführt (siehe z.B. MyCoRe oder Textgrid). Eine Vergleichbarkeit mit freien Open-Source-Frameworks hinsichtlich Funktionalität, Wartung usw. ist aber kaum gegeben. Für Projekte wie die „Virtuelle Forschungsumgebung für die Geisteswissenschaften – FuD“⁸ oder eben auch H-Soz-Kult haben deren „Heimat“-Universitäten Teile der Trägerschaft übernommen. Nur geht die technische Entwicklung in der Regel mit Schulungen von Anwendern, der Weiterentwicklung von Funktionalitäten und Wartungsaufwand einher; Aufgaben,

⁸ <http://fud.uni-trier.de/de/>

deren finanziellen Bedarfe bzw. die dafür notwendigen Ressourcen nicht zu unterschätzen sind.⁹ Letztlich lässt sich die zukünftige Nutzung neuer digitaler Arbeits- und Kommunikationsformen nur schwer prognostizieren, vor allem in welcher Art und Weise technologische Veränderungen und einhergehende soziale Praktiken das Informations- und Kommunikationsverhalten verändern. Augenscheinlich ist bspw. die wachsende Nutzung von twitter, facebook und blogs auch im wissenschaftlichen Bereich angekommen. Allerdings mehren sich auch Stimmen, die für eine Rückbesinnung auf analoge Formen plädieren. So hat kürzlich die TED-Conference – eine Institution im Bereich Nutzung digitaler Medien in der Wissenschaftskommunikation – auf ihrer letzten Tagung nach einem klaren Votum ihrer Teilnehmer/innen für einen vollständigen Ausschluss von Smartphones, twitter und Co. auf der Konferenz plädiert.¹⁰ Außerhalb der akademischen Welt beginnen Tageszeitungen ausufernde „Debatten“ in ihren Kommentarfunktionen auf ein redaktionell zu bewältigendes Maß zurückzufahren.¹¹ Wie eingangs erwähnt, hat die Technikgeschichte in den vergangenen Jahren einige interessante „Beobachtungen“ gemacht: „Research technologies“ setzen sich in anderen Fächern durch, wenn sie durch einige Akteure, die in unterschiedlichen Disziplinen aktiv sind, zwischen den Disziplinen übertragen werden. In der Geschichtswissenschaft sind dies in der Regel „technikaffine“ Wissenschaftler/innen, die nicht den Mehrwert von Tools in ihrer „Community“ preisen müssen, sondern Anforderungen definieren und selbst den Nutzen IT-gestützter Arbeitstechniken in ihrem Arbeitsumfeld praktizieren (und somit vermitteln) können. Jede Forschungsrichtung, jede Lehranwendung beansprucht zudem für sich eigene Anforderungen, insofern dürfte weiterhin von der Prämisse auszugehen sein, dass wegen dieser „Heterogenität [...] ein allgemeingültiges Set von Erfolgskriterien als unrealistisch erscheint.“¹²

⁹ Über die Fallstricke der Softwareentwicklung im akademischen Rahmen wurde nicht zuletzt auf einer Tagung Ende 2014 diskutiert. Vgl. Tagungsbericht: Redaktionssysteme und Virtuelle Forschungsumgebungen, 20.10.2014 – 21.10.2014 München, in: H-Soz-Kult, 04.12.2014, <http://www.hsozkult.de/conferencereport/id/tagungsberichte-5716>.

¹⁰ Ted Conference - Freigang für Pawlows Hunde. In: Süddeutsche Zeitung, 19.06.2015, <http://www.sueddeutsche.de/kultur/ted-conference-freigang-fuer-pawlows-hunde-1.2526924>

¹¹ Neue zeitliche Regelung in der Community. In: tagesspiegel, 09.06.2015, <http://www.tagesspiegel.de/medien/digitale-welt/in-eigener-sache-neue-zeitliche-regelung-in-der-community/11890646.html>

¹² Vgl. Buddenbohm, Stefan; Enke, Harry; Hofmann, Matthias u.a.: Erfolgskriterien für den Aufbau und nachhaltigen Betrieb Virtueller Forschungsumgebungen, 2014, <http://webdoc.sub.gwdg.de/pub/mon/dariah-de/dwp-2014-7.pdf>.