

H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, M. Jehn (Hrsg.)

# nestor Handbuch

Eine kleine Enzyklopädie  
der digitalen Langzeitarchivierung

Version 2.0

Kapitel 12.2

Workflow in der Langzeitarchivierung  
– Methode und Herangehensweise

nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung  
hg. v. H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, M. Jehn  
im Rahmen des Projektes: nestor – Kompetenznetzwerk Langzeitarchivierung und  
Langzeitverfügbarkeit digitaler Ressourcen für Deutschland  
nestor – Network of Expertise in Long-Term Storage of Digital Resources  
<http://www.langzeitarchivierung.de/>

Kontakt: [editors@langzeitarchivierung.de](mailto:editors@langzeitarchivierung.de)  
c/o Niedersächsische Staats- und Universitätsbibliothek Göttingen,  
Dr. Heike Neuroth, Forschung und Entwicklung, Papendiek 14, 37073 Göttingen

Die Herausgeber danken Anke Herr (Korrektur), Martina Kerzel (Bildbearbeitung) und  
Jörn Tietgen (Layout und Formatierung des Gesamttextes) für ihre unverzichtbare  
Unterstützung bei der Fertigstellung des Handbuchs.

Bibliografische Information der Deutschen Nationalbibliothek  
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen  
Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter  
<http://www.d-nb.de/> abrufbar.

Die Inhalte dieses Buchs stehen auch als Onlineversion  
(<http://nestor.sub.uni-goettingen.de/handbuch/>)  
sowie über den Göttinger Universitätskatalog (<http://www.sub.uni-goettingen.de>) zur  
Verfügung.

Die digitale Version 2.0 steht unter folgender Creative-Commons-Lizenz:  
„Attribution-Noncommercial-Share Alike 3.0 Unported“  
<http://creativecommons.org/licenses/by-nc-sa/3.0/>



Einfache Nutzungsrechte liegen beim Verlag Werner Hülsbusch, Boizenburg.  
© Verlag Werner Hülsbusch, Boizenburg, 2009  
[www.vwh-verlag.de](http://www.vwh-verlag.de)  
In Kooperation mit dem Universitätsverlag Göttingen

Markenerklärung: Die in diesem Werk wiedergegebenen Gebrauchsnamen, Handelsnamen,  
Warenzeichen usw. können auch ohne besondere Kennzeichnung geschützte Marken sein und  
als solche den gesetzlichen Bestimmungen unterliegen.

Druck und Bindung: Kunsthaus Schwanheide

Printed in Germany – Als Typoskript gedruckt –

ISBN: 978-3-940317-48-3



*Gewidmet der Erinnerung an Hans Liegmann (†), der als Mitinitiator und früherer Herausgeber des Handbuchs ganz wesentlich an dessen Entstehung beteiligt war.*

## 12.2 Workflow in der Langzeitarchivierung: Methode und Herangehensweise

*Reinhard Altenböner*

Die allmähliche Einführung der Langzeitarchivierung in das reguläre Auftragsportfolio von Bibliotheken und anderen Kulturerbeeinrichtungen mit immer höheren Bindungsquoten von Personal und anderen Ressourcen erzeugt(e) zunächst neue, häufig isolierte und händisch durchgeführte Abläufe, verändert aber auch in einer ganzheitlichen Betrachtung Arbeitsabläufe und die sie modellierenden Geschäftsprozesse. So ist schon für sich die Einspielung von Daten in ein Langzeitarchiv ein komplexer Vorgang, in dem eine ganze Reihe von aufeinander bezogenen bzw. voneinander abhängenden Aktivitäten ablaufen. Vor allem aber die zunehmende Relevanz der technischen und operativen Bewältigung der Aufgabe verlangt nach einer systematischen Modellierung der Geschäftsprozesse, also dem Einstieg in ein systematisches (technisches) Workflowmanagement. Es gilt allerdings festzustellen, dass selbst in Einrichtungen, die bereits seit einigen Jahren Erfahrungen mit dem Betrieb von Langzeitarchiven und ihrer Integration in die jeweilige Systemlandschaft gesammelt haben, häufig noch isolierte Bearbeitungsketten ablaufen, die zudem keinesfalls wirklichen Vollständigkeitsgrad haben, also alle Anforderungs- /arbeitsfelder abdecken und außerdem vielfach noch manuelle Eingriffe erfordern, insbesondere auf dem Gebiet des Fehlermanagements.

Diese Feststellung bedeutet aber auch, dass der Erfahrungshorizont zum technischen Workflow und insbesondere zum Management insgesamt noch gering ist, also hier noch konkrete Erfahrungen vor allem im Umgang mit großen Mengen und insbesondere auch im automatisierten Qualitätsmanagement gewonnen werden müssen. Insofern hat die Beschäftigung mit dem technischen Workflow derzeit noch viele theoretische Elemente und ‚propädeutischen‘ Charakter.

In einer Situation, in der verschiedene (bereits existente und neu entwickelte) Arbeitsprozesse ineinander greifen und auch verschiedene Organisationseinheiten an ein und demselben Vorgang beteiligt sind, ist die Modellbildung im Sinne der Geschäftsprozessmodellierung ein Beitrag zu einer umfassenden Optimierung. Damit befinden sich Bibliotheken, Archive und Museen in einer Situation, die man mit den Anstrengungen der Privatwirtschaft Anfang der 1990er Jahre vergleichen kann, als dort die Modellierung von Geschäftsprozessen unter verschärften Wettbewerbs- und Kostendruckbedingungen systematischer als zuvor angegangen wurde. Auch wenn im öffentlich finanzierten Um-

feld in besonderem Maße historisch geprägte Organisationsformen gegeben sind, die eine vorgangsbezogene Sicht erschweren, führt an der grundsätzlichen Anforderung der Neu-Modellierung aus systematischer Sicht kein Weg vorbei. Diese wird im Umfeld des technischen Workflow immer stark auch von der informationstechnischen Entwicklungsseite getrieben sein, denn Ziel der Geschäftsprozessmodellierung ist letztlich ihre technische Abbildung.

Übergeordnete Ziele dieses Herangehens, also der systematischen Modellierung und eines methodenbewussten Workflowmanagements sind<sup>6</sup>:

- Verbesserung der Prozessqualität
- Vereinheitlichung der Prozesse
- Schnellere und zuverlässigere Bearbeitung von Aufträgen (extern und intern)
- Reduzierung der Durchlaufzeiten
- Kostenreduktion
- Verbesserte Verfügbarkeit von Information / Dokumentation
- Erhöhte Prozessflexibilität
- Erhöhung der Transparenz der Prozesse (Statusermittlung, Dokumentation von Entscheidungen), Qualitätssicherung
- Automatische Eingriffsmöglichkeiten: Dokumentation, Eskalation bei Zeitüberschreitungen, Verteilung von Aufgaben und Verantwortlichkeiten
- Vermeidung von Redundanz, mangelnder Aktualität und Inkonsistenz durch Mehrfachschritte

Natürlich lassen sich kleine isolierte Prozesse oder Prozesselemente durch individuelle Programmierung jeweils neu umsetzen. Dies geschah in der Vergangenheit vielfach für einzelne Objektklassen oder auch einzelne Datenübergabe- oder -tauschprozesse. Aber schon beim Zusammenführen bzw. Hintereinandersetzen der einzelnen Teilschritte bedarf es einer Gesamtlogik für das Management des Ablaufs dieser Schritte. Fehlt diese Logik, entstehen letztlich viele immer wieder manuelle neu anzustoßende Teilkonstrukte mit dazu häufig proprietären „Konstruktions“elementen. Schon insofern ist die systematische Analyse verschiedener wiederkehrender Arbeitsabläufe ein sinnvoller Ansatz, um so zur Modellierung auch komplexer Vorgänge im Bereich der Langzeitarchivierung zu kommen.

---

6 Die folgende summarische Zusammenstellung betrifft sowohl organisatorische wie technische Aspekte des Workflowmanagements. Eine Trennung ist theoretisch zwar möglich, praktisch aber nicht sinnvoll.

Ziel dieses systematischen Ansatzes ist es, Services zu definieren, die auch in anderen Kontexten (wieder) verwendbar sind. Sie bilden Arbeitsschritte granular ab, die so in verschiedenen Umfeldern vorkommen (können), beispielsweise das Aufmachen eines Bearbeitungsfalls für ein Objekt und die IT-gestützte Verwaltung verschiedener Be-/Verarbeitungsschritte dieses Objekts. In dieser Perspektive entsteht der Geschäftsprozess für eine Klasse von Objekten aus der Zusammenfügung verschiedener Basisservices, die miteinander interoperabel sind. Dass diese Herangehensweise sehr stark mit dem Modell der Serviceorientierten Architektur (SOA) bei der Entwicklung IT-basierter Lösungen korrespondiert, ist dabei kein Zufall. Voraussetzung dafür ist aber die systematische Modellierung der Arbeits- oder Geschäftsprozesse, die vorgeben, welche Services wann und wie gebraucht werden. Die Prozessmodellierung bildet also die Basis für die Implementierung, die Prozesse selbst dienen der Orchestrierung, dem Zusammenspiel und der Aufeinanderabstimmung der Services. In einem optimalen (Infrastruktur)-Umfeld können so die Arbeitsschritte als kleinere Einheit eines Geschäftsprozesses verschiedene Services lose zusammenbringen.

Der Ansatz, Services nachnutzbar zu gestalten, bezieht sich in der Regel auf eine Organisation. Zwar wird immer wieder versucht, Geschäftsprozesse aus einem institutionellen Umfeld auf ein anderes zu übertragen, allerdings erweist sich dies in der Praxis als außerordentlich schwierig: Zu stark sind die Abweichungen der einzelnen Arbeitsschritte voneinander und zu unterschiedlich die jeweiligen Prioritäten und Schwerpunktsetzungen in den einzelnen Institutionen. Hinzu kommt außerdem noch, dass der Prozess der Modellierung und Ausgestaltung von Geschäftsprozessen selbst erhebliche Erkenntnisgewinne in der jeweiligen Organisation mit sich bringt, die für eine erfolgreiche Einführung neuer oder veränderter Geschäftsprozesse unverzichtbar sind.

Die Informatik hat für die Modellierung und Notation von Geschäftsprozessen verschiedene methodische Herangehensweisen entwickelt, zum Beispiel die Ereignisgesteuerten Prozessketten (EPK), eine von Scheer und Mitarbeitern entwickelte Sprache zur Modellierung von Geschäftsprozessen<sup>7</sup> und vor allem die Unified Modeling Language (UML) der Object Management Group (OMG), die in der Praxis heute dominierende (technische) „Sprache“ für die Modellierung von Daten, Verhalten, Interaktion und Aktivitäten.<sup>8</sup>

Legt man also UML als Syntax fest, sind noch methodische Festlegungen für die Herangehensweise zu treffen und es liegt nahe, sich für die vorbereitende Modellierung von technischen Abläufen in der Langzeitarchivierung am OAIS-

---

7 Keller (1992)

8 OMG Infrastructure (2007) und OMG Superstructure (2007)

Modell zu orientieren, das die prinzipiellen Aufgaben im Umfeld der Langzeitarchivierung in funktionaler Perspektive beschreibt und an anderer Stelle dieser Enzyklopädie ausführlich beschrieben wird.<sup>9</sup> Für den Bereich des Ingest einzu-beziehen ist der Producer-Archive Interface Methodology Abstract Standard<sup>®</sup> (CCSDS 651.0-B-1), der insbesondere Validierungsmechanismen betrachtet.<sup>10</sup>

Einzelne Funktionen lassen sich so vor der Folie bisher bereits gemachter Erfahrungen allgemein beschreiben. Beispiele für diese übergreifenden Basisprozesse sind (ich nenne nur Beispiele für unmittelbar aus dem Kontext der Langzeitarchivierung heraus relevante Prozesse):

- Plattform- und Systemübergreifendes Taskmanagement
- Daten- und Objekttransfer-Mimik (z.B. OAI, ORE)
- Extraktion und Generierung von Metadaten (METS, LMER)
- Validierung von Dokumentformaten (z.B. JHOVE)
- Persistente Adressierung und Zugriffsmanagement auf Objektebene
- Speicherprozesse
- ID-Management
- Inhaltsauswahl / Basisrecherche
- Migrationsprozesse / Formatkonvertierungen
- On-the-fly-Generierung einer Bereitstellungsumgebung

---

9 Vgl. hierzu den entsprechenden Artikel von Nils Brübach / Manuela Queitsch / Hans Liegmann (†) in dieser Enzyklopädie als Kapitel 4: „Das Referenzmodell OAIS - Open Archival Information System“

10 Vgl. hierzu <http://public.ccsds.org/publications/archive/651x0b1.pdf>