

nestor Handbuch:
**Eine kleine Enzyklopädie
der digitalen Langzeitarchivierung**

10.1.3 PREMIS

Herausgeber

Heike Neuroth
Hans Liegmann †
Achim Oßwald
Regine Scheffel
Mathias Jehn
Stefan Strathmann

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Im Auftrag von

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

c/o

Niedersächsische Staats- und Universitätsbibliothek Göttingen

Dr. Heike Neuroth

Forschung und Entwicklung

Papendiek 14

37073 Göttingen

Tel. +49 (0) 55 1 39 38 66

Der Inhalt steht unter folgender Creative Commons Lizenz:
<http://creativecommons.org/licenses/by-nc-sa/2.0/de/>



10.1.3 PREMIS

Olaf Brandt

PREMIS steht für „PREservation Metadata: Implementation Strategies“. Diese von der OCLC (Online Computer Library Center) und RLG (Research Library Group) im Jahre 2003 ins Leben gerufene Initiative betreibt die Entwicklung und Pflege des international anerkannten gleichnamigen PREMIS-Langzeitarchivierungsmetadatenstandards.

Die Mitglieder von PREMIS sind Akteure aus dem Umfeld von Gedächtnisorganisationen wie Archive, Bibliotheken und Museen, sowie der Privatwirtschaft. Diese befassen sich in internationalen Arbeitsgruppen mit Problemen der digitalen Langzeitarchivierung.

Das Hauptziel von PREMIS ist die Entwicklung von Empfehlungen, Vorschlägen und best-practices zur Implementierung von Langzeitarchivierungsmetadaten. Dazu gehört die Schaffung eines Kerns von Langzeitarchivierungsmetadaten mit größtmöglicher Anwendbarkeit innerhalb unterschiedlichster Langzeitarchivierungskontexte.

Die Arbeit von PREMIS baut auf den Ergebnissen der Preservation-Metadata-Working-Group auf. Diese Arbeitsgruppe wurde 2001 zur Entwicklung eines Rahmenkonzeptes für Langzeitarchivierungsmetadaten gebildet. Eine wichtige Grundlage für PREMIS ist das Referenzmodell des Open-Archival-Information-Systems (OAIS, ISO Standard 14721:2003). Dieses behandelt v.a. organisatorische und technische Fragen der digitalen Langzeitarchivierung.

Die Zielsetzung der so genannten Core Elements Group war bis Anfang 2005 die Entwicklung eines Kerns von Langzeitarchivierungsmetadaten, die Erstellung von Mappings und die Anbindung an andere Standards sowie der Aufbau eines Langzeitarchivierungsmetadatenlexikons. Die Ergebnisse dieser Gruppe sind in einem Abschlussbericht im Mai 2005 veröffentlicht worden.

Der Bericht beinhaltet das sogenannte PREMIS Data Dictionary 1.0, welches von einem ausführlichen Kommentar begleitet wird. Hierin sind der Kontext, das Datenmodell und die PREMIS-Grundannahmen aufgeführt. Zudem enthält der Bericht Erklärungen und Erläuterungen zu im Bericht erwähnten Themen, ein Glossar und erläuternde Beispiele. Das PREMIS-Data-Dictionary ist die Grundlage für die praktische Implementierung von Langzeitarchivierungs-

metadaten in digitalen Archiven.

Die zweite Arbeitsgruppe widmete sich den eher praktischen Fragen der realen Implementierung von Langzeitarchivierungsmetadaten. Untersucht wurden Fragen wie ‚Wie ist der Entwicklungsstand?‘ und ‚Was wird in welcher Weise implementiert?‘. Darüber hinaus werden Themen über Datenhandling, eingesetzte Software und rechtliche Fragen erörtert. Erzielt wurden Empfehlungen zu best-practices auf Basis einer Reihe von Systemumgebungen. Die Ergebnisse flossen in einen im September 2004 veröffentlichten Untersuchungsbericht ein.

Implementierung

Aufbauend auf den Ergebnissen der Arbeitsgruppen stehen XML-Schemas zur Verfügung, welche in Langzeitarchivsysteme implementiert werden. Weiterhin sind die Schemas in Metadaten-Container-Formate (z.B. METS) integriert. Zu den nächsten Schritten zählen die maschinelle Erzeugung und Verarbeitung von PREMIS-Metadaten sowie die Integration in Workflows. Eine Liste von Institutionen, die PREMIS implementieren findet sich auf den PREMIS-Maintenance-Activity-Seiten der Library of Congress in den Vereinigten Staaten von Amerika. Eine rege Community in der Mailingliste der PREMIS-Implementors-Group diskutiert viele Fragen rund um die Implementierung von PREMIS und um unterschiedliche Themen der digitalen Langzeitarchivierung.

Die PREMIS-Maintenance-Activity übernimmt die weitere Koordination der Aktivitäten. Ein Teil davon, das PREMIS-Editorial-Committee, widmet sich der Verbreitung von PREMIS und der weiteren Pflege des Standards. Dazu gehören z.B. notwendige Anpassungen im Data-Dictionary oder den XML-Schemas. Diese Anpassungen werden gerade unter dem Eindruck der ersten praktischen Erfahrungen vorgenommen. Zur Verbreitung von PREMIS werden international unterschiedliche Veranstaltungen angeboten.

Den PREMIS-Aktivitäten wird im Kontext der Langzeitarchivierung übereinstimmend große Bedeutung im Bereich der Zusammenarbeit und des Datenaustausches beigemessen. Das schlägt sich auch in zwei internationalen Auszeichnungen nieder: für das Data-Dictionary wurde der PREMIS-Gruppe Ende 2005 den Digital-Preservation-Award der Digital-Preservation-Coalition und im August 2006 den Preservation-Publication-Award der Society of American Archivists verliehen.

Datenmodell

Das PREMIS-Datenmodell kennt einen vielseitigen Objektbegriff. Ein Objekt (Object) kann entweder eine Datei (File), ein Datenstrom (Bitstream) oder eine Repräsentation (Representation) sein. Ein Datenstrom ist dadurch gekennzeichnet, dass er sich nicht ohne zu ergänzende Daten oder einer Umformatierung in eine selbstständige Datei wandeln lässt. Eine Repräsentation ist eine Menge von Dateien, welche nur zusammenhängend eine sinnvolle und vollständige Darstellung einer intellektuellen Einheit (Intellectual Entity) liefern. Neben intellektuellen Einheiten und Objekten existieren im Datenmodell noch Rechte (Rights), Agenten (Agents) und Ereignisse (Events). Ereignisse und Rechte stehen in direkten Beziehungen zu Objekten und/oder Agenten. Zwischen Objekten können Beziehungen bestehen, die strukturelle Zusammengehörigkeit, Ableitungen oder Abhängigkeiten kennzeichnen.

Object Entity

Zu den Metadaten des Objekts gehören eindeutige Kennungen, Charakteristiken der Datei wie Größe und Format, Beschreibungen der Systemumgebungen (Software, Hardware), eine Auflistung der relevanten Eigenschaften der Objekte, sowie die Beziehungen zu Events und Rechteinformationen.

Event Entity

In der Ereignis-Entität können Aktionen, die in Verbindung mit Objekten oder Agenten stehen, dokumentiert werden. Dazu gibt es eindeutige Kennungen für Ereignisse und Aktionen, sowie Informationen über deren Resultate.

Agent-Entity

Ein Agent ist eine Person, eine Organisation oder Software, die Aktionen mit Objekten durchführt. Agenten werden durch eine eindeutige Kennung beschrieben.

Rights-Entity

Bei den Rechten werden Genehmigungen zur Durchführung von Aktionen von Agenten mit Objekten genau definiert.

Für PREMIS gibt es für jeden Entity-Typ ein eigenes XML-Schema, sodass eine modulare Einbindung in andere Schemas wie METS möglich ist.

Literatur:

Webseite der PREMIS Arbeitsgruppe: <http://www.oclc.org/research/projects/pmwg/>

Webseite der PREMIS Maintenance Activity: <http://www.loc.gov/standards/premis/>

Abschlußbericht der PREMIS Arbeitsgruppe inkl. Data Dictionary for Preservation Metadata: <http://www.oclc.org/research/projects/pmwg/premis-final.pdf>

PREMIS Survey Implementing Preservation Repositories for Digital Materials, Current Practice and Emerging Trends in the Cultural Heritage Community (survey report):

<http://www.oclc.org/research/projects/pmwg/surveyreport.pdf>

Digital Preservation Award 2005 der DPC:

<http://www.dpconline.org/graphics/advocacy/press/award2005.html>

Preservation Publication Award 2006 der Society of American Archivists:

<http://www.archivists.org/recognition/dc2006-awards.asp#preservation>

Preservation Metadata Working Group (PMWG 2002) Framework:

http://www.oclc.org/research/projects/pmwg/pm_framework.pdf