

Virtuelles Treffen der nestor AG SIP-Konkretisierung am 19. Juni 2024

Teilnehmende: Tobias Steinke, Maximilian Maede, Jens Steidl, Ziliana Efken

Wiki-Seite

Der Abschnitt „Tools“ wurde erweitert und inhaltlich breiter gefasst. Dazu können auch spezielle Tools für Teilaspekte in Ingest-Workflows gehören. In diesem Sinne werden BagIt-Tools differenzierter aufgelistet und auch Tool-Bibliotheken für Rosetta ergänzt. Der Abschnitt „Standards“ sollte als Tabelle mit Spalten zur Erklärung des jeweiligen Zwecks insbesondere bei Metadatenstandards versehen werden.

Berichte über aktuelle Entwicklungen

Migration von Rosetta zu Archivematica

Bei der SLUB wird derzeit angesichts der Cloud-Pläne von Rosetta der Umstieg auf Archivematica geprüft. Dabei geht es auch um die dann nötige Migration der Informationspakete. Eine direkte Migration von AIP zu AIP ist nicht einfach möglich. Es würde daher über eine Zwischendatenbank als SIP in das neue LZA-Archiv eingespielt. Das bisherige Rosetta-METS wäre Teil des SIP-Inhalts. Die SIP-Spezifikationen, wie sie im Beispiel-Dokument der SLUB veröffentlicht wurden, sind durch die abstrakte Schicht so, dass sie auch bei einer Migration weiterhin gültig bleiben. Archivematica wurde bereits ergänzt um Eigenschaften, die die SLUB für einen Umstieg bräuchte. Die endgültige Entscheidung für den Umstieg ist jedoch noch nicht gefallen.

Weitere Informationen hierzu:

<https://slubarchiv.slub-dresden.de/technische-standards-fuer-die-ablieferung-von-digitalen-dokumenten>

-> Ergänzende Dokumente

-> RFC Workflow Rosetta Archivematica Migration (v0.11)

(ganz unten rechts auf der Seite)

Webarchivierungspraxis im Sächsischen Staatsarchiv

Bei einer virtuellen Veranstaltung am 8.5.2024 stellte das Sächsische Staatsarchiv ihre aktuelle Praxis bei der Webarchivierung vor. Dabei wurde auch ausdrücklich die SIP-Generierung erwähnt, die vergleichsweise ungewöhnlich ist. Die vom verwendeten Tool Offline Explorer erzeugten Verzeichnisse und Dateien eines Crawls werden als ISO-Diskimage archiviert. Dies erfolgt für eine einfachere Bereitstellung. Für die Qualitätssicherung werden vorher die Dateien manuell korrigiert und verändert, wobei dies durch die Tatsache vereinfacht wird, dass etwa Zweidrittel der gesammelten Webseiten aus dem gleichen CMS von sachsen.de kommen.

Metadaten und Normalisierung als neue Schwerpunktthemen der AG

Bei der letzten Sitzung wurden als neue Themen für Dokumente, die von der AG erstellt werden könnten, Empfehlungen für Metadaten in SIPs und zu Dateiformaten ggf. per Normalisierung genannt. Allerdings stellt sich bei beiden Themen die Frage, inwiefern dies verallgemeinerbar ist. Zudem müsste die veröffentlichte Handreichung zur SIP-Bildung noch mal überprüft werden zu den dortigen Aussagen in den Bereichen und ob eine weitere Konkretisierung sinnvoll und möglich erscheint. Dies soll beim nächsten Termin noch mal diskutiert werden.