

DIMAG

Kurzbeschreibung

August 2020

Herausgegeben vom DIMAG-Entwicklerverbund

Landesarchiv Baden-Württemberg

Hessisches Landesarchiv

Generaldirektion der Staatlichen Archive Bayerns

Kooperationsverbund Digitales Archiv Nord

Bearbeiter: Christian Keitel, Rolf Lang, Elisabeth Klindworth

Stand: August 2020

Inhalt

1. Grundinformationen	3
2. Eigenschaften von DIMAG	4
3. Metadatenkonzept	6
4. Technische Angaben zum System.....	7
5. Schnittstellen.....	9
6. Rechte.....	9
7. Literatur	10
8. Kontakt	11

1. Grundinformationen

Das Archivierungssystem DIMAG wurde 2006 vom Landesarchiv Baden-Württemberg im Rahmen des Projekts „Konzeption für ein digitales Landesarchiv“ und nach den Vorgaben des OAIS-Standards (ISO 14721), des PREMIS-Standards und des Nestor-Kriterienkatalogs „Vertrauenswürdige digitale Langzeitarchive“ (jetzt DIN 31644) entwickelt. DIMAG läuft produktiv seit dem 7. Juli 2006. Seitdem wird das System kontinuierlich weiterentwickelt. Aktuell werden beim Landesarchiv in 1.7 Mio. Dateien über 90.000 Archivalien mit mehr als 265 Mio. Datensätzen archiviert.

DIMAG kann historisch wertvolle Informationen sicher importieren, technisch dokumentieren, authentisch erhalten und für die Benutzung ausgeben. Das System ist zur Archivierung sämtlicher denkbarer Typen digitaler Objekte geeignet.

Das Akronym DIMAG steht für „Digitales Magazin“. DIMAG wird derzeit vom Landesarchiv Baden-Württemberg, dem Landesarchiv Hessen, den Staatlichen Archiven in Bayern und dem Digitalen Archiv Nord¹ eingesetzt. Diese Archivverwaltungen teilen sich nun die Arbeit an der Weiterentwicklung von DIMAG.

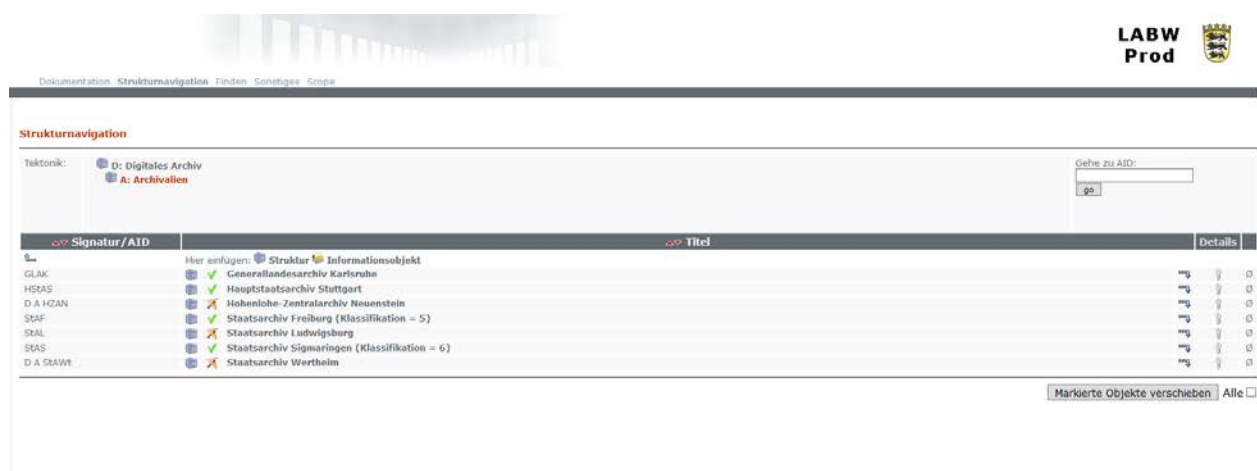


Abb. 1: DIMAG Kernmodul

Seit 2016 übernimmt das Landesarchiv auch die digitalen Grundakten der baden-württembergischen Justizverwaltung und archiviert diese mit der DIMAG-Software. Das G-DIMAG („G“ für Grundakte) stellt sicher, dass die Grundakten auf Dauer inhaltlich unverändert in lesbarer Form zur Verfügung stehen. Die besondere Herausforderung für G-

¹ Das Digitale Archiv Nord wird getragen von den Archiverwaltungen von Hamburg, Mecklenburg-Vorpommern, Sachsen-Anhalt, Schleswig-Holstein und Bremen.

DIMAG besteht in der großen Anzahl der zu speichernden digitalen Objekte und den häufigen Zugriffen auf die gespeicherten Objekte. Aktuell beinhaltet die Datenbank des G-DIMAG über 80 Mio. Einträge.

2. Eigenschaften von DIMAG

DIMAG ist heute ein Programmpaket, das von den DIMAG-Entwicklungspartnern gemeinsam weiterentwickelt wird:

- DIMAG-Kernmodul
- IngestList
- IngestTool
- DIMAG-Accessmodul
- DIMAG-Bestandserhaltungsmodul

Die Module werden nur als Gesamtpaket abgegeben. Sie werden ergänzt durch das OpenSource-Werkzeug IngestList, das frei verfügbar im Internet steht.

Die sichere und dauerhafte Erhaltung der Archivalien in DIMAG wird durch verschiedene Faktoren unterstützt:

- DIMAG orientiert sich an den für die digitale Archivierung einschlägigen Standards ISO 14721 (OAIS), PREMIS, DIN 31644 und DIN 31645.
- DIMAG verwaltet zwar die digitalen Archivalien. Die Ablage erfolgt jedoch außerhalb von DIMAG im Dateisystem. Selbst bei einem sofortigen, nicht vorhersehbaren Ausfall der DIMAG-Software könnten daher alle Metadaten und Primärdokumente der digitalen Archivalien aufgerufen und rekonstruiert werden.
- An der fachlichen Weiterentwicklung von DIMAG sind derzeit vier Partner beteiligt: Die Archivverwaltungen von Baden-Württemberg, Hessen und Bayern und die im digitalen Archiv Nord (DAN) zusammengeschlossenen Archivverwaltungen der nördlichen Bundesländer. Derzeit gehören zu DAN die Archivverwaltungen von Hamburg, Mecklenburg-Vorpommern, Sachsen-Anhalt, Schleswig-Holstein und Bremen. Der Beitritt weiterer Archivverwaltungen ist möglich. Damit ist sichergestellt, dass die Weiterentwicklung stets „state of the art“ erfolgt.
- Als browserbasiertes System ist DIMAG – nach erfolgter Freischaltung – von jedem beliebigen Rechner aufrufbar.

- Intuitive, einfache Bedienung des Systems
- Da DIMAG beliebige Formate und Strukturierungen aufnehmen kann, ist auch eine Ablage neuer Dokumente in heute unbekanntem künftigen Dateiformaten und Objektstrukturen antizipiert.
- DIMAG besitzt ein eindeutiges, systemunabhängiges Identifizierungskonzept, das auch nach der Ablösung von DIMAG durch eine zukünftige Software weitergeführt werden kann.
- Eine Versionsverwaltung gewährleistet in DIMAG, dass auch bei erneuter Ablage eines Dokuments oder seiner Metadaten die vorhandenen Dateien nicht überschrieben werden. Selbst bei dem absichtlichen oder unabsichtlichen erneuten Abspeichern gehen die ursprünglichen Informationen nicht verloren.
- DIMAG besitzt ein speziell auf Anforderungen der Archivierung zugeschnittenes Metadatenkonzept.
- Das IngestTool ermöglicht die einfache, graphisch unterstützte Zuordnung von Informationen aus sehr unterschiedlichen Quellen (Dateiattribute, Verzeichnisnamen, CSV-/TXT-/ oder XML-Dateien) in frei zu definierende XML-Ausgabeformate zu übertragen, die wiederum in archivische Erschließungssysteme importiert werden können. Vordefiniert und auswählbar ist die Ausgabe einer XML-Datei für den Import von Metadaten in DIMAG, weitere XML-Ausgabeformate können durch den Anwender selbst definiert werden.

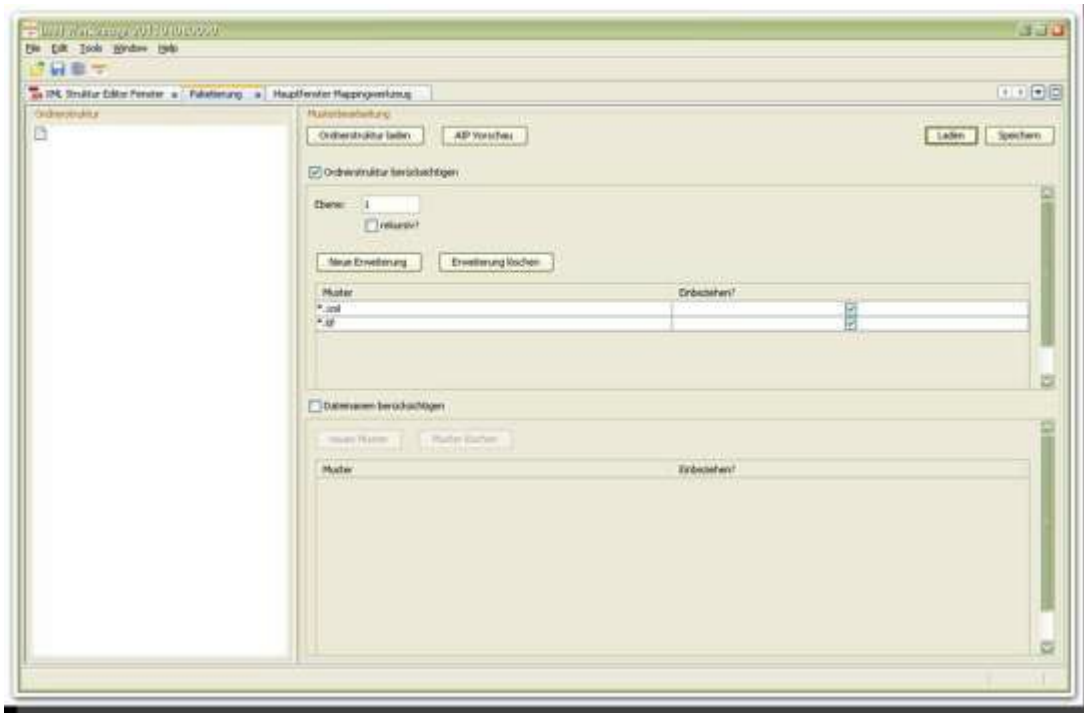


Abb. 2: DIMAG-IngestTool: Erstellung von Archivpaketen (AIPs)

- Die Paketierungskomponente des IngestTools erlaubt es, aus einem Übergabepaket (SIP) musterbasiert AIPs zu bilden und diese anschließend in DIMAG zu importieren. Die Paketierung kann gemäß Ordnerstrukturen oder Dateinamen und -endungen erfolgen.
- Das Accessmodul eröffnet unterschiedliche Möglichkeiten bei der Generierung eines Nutzungspakets.

3. Metadatenkonzept

Das Metadatenkonzept von DIMAG beruht auf folgenden Grundsätzen und Elementen:

- Es wird zwischen den Elementen Primärdatendatei, Metadatendatei, Repräsentation, Digitales Objekt (Archivale) und Struktureintrag unterschieden.
- Die einzelnen Elemente sind eindeutig in einer monohierarchischen Struktur verortet.
- Der konkrete Ausbau der hierarchischen Struktur (Klassifikation, Tektonik) kann frei gestaltet werden.

- Es gibt nur sehr wenige Pflichtmetadaten (führt zu besserer Akzeptanz bei den archivischen Benutzern).
- Die Elemente werden durch persistente Identifikatoren dauerhaft eindeutig miteinander verknüpft.
- Durch das in DIMAG angelegte Element der Repräsentation ist eine eindeutige Abbildung und Integration der durch künftige Dateiformatmigrationen entstehenden Dateien sichergestellt.
- Sämtliche relevanten Prozesse in DIMAG werden in Protokolle geschrieben. Die Protokolle werden in XML gespeichert, sodass der Lebensweg der digitalen Archivalien auch noch nach Jahrzehnten nachvollziehbar bleibt. Unwesentliche Informationen wie z.B. die täglichen Backupläufe werden an anderer Stelle festgehalten.

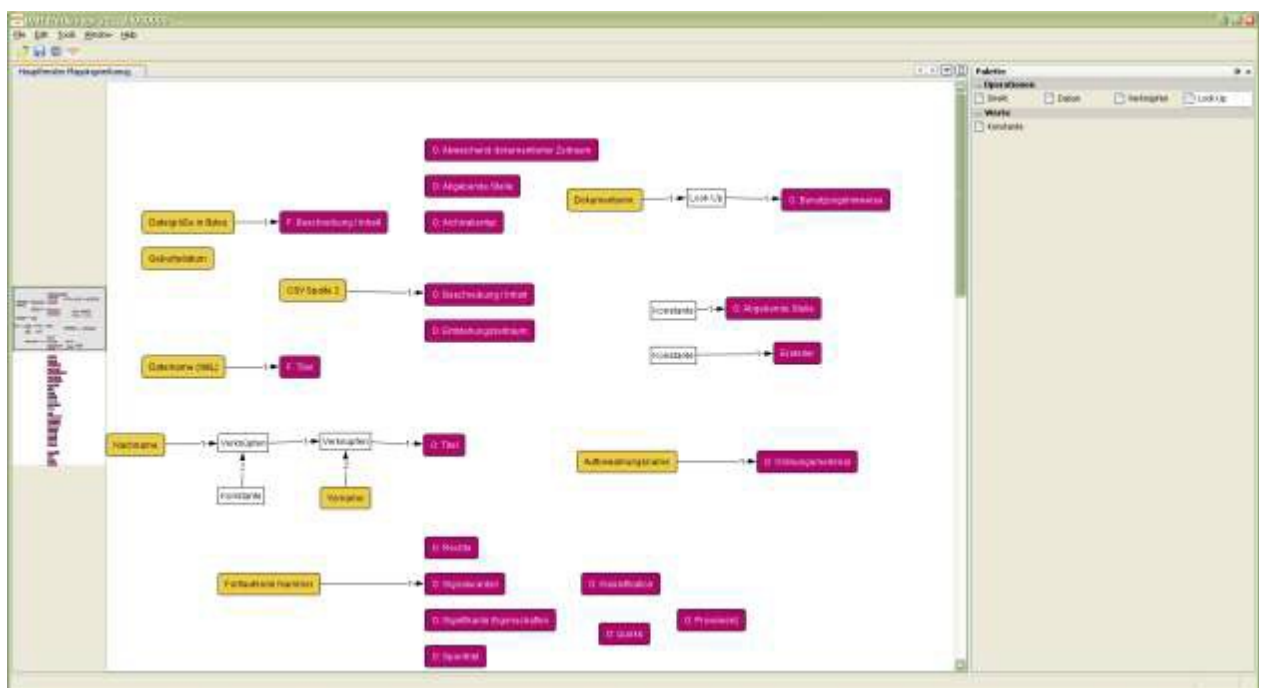


Abb. 3: DIMAG-IngestTool: Zuordnung der Metadaten

4. Technische Angaben zum System

Die in DIMAG archivierten Archivalien und Metadaten können von speziell freigeschalteten Rechnern über einen der gängigen Internetbrowser abgerufen werden. Die Kommunikation erfolgt stets verschlüsselt über https und sftp. Der Zugang ist an eine Rechteverwaltung

gekoppelt.

Sämtliche Metadaten und Primärdaten werden im Dateisystem abgelegt. Außerdem speichert DIMAG die Metadaten zusätzlich für Recherchezwecke in einer Datenbank ab. Die verwendeten Felder sind frei konfigurierbar in einer Datenbank-Tabelle.

Um die Integrität des Dateisystems sicherzustellen sind folgende Maßnahmen umgesetzt:

- Das Dateisystem basiert auf einem journaling Dateisystem, welches alle Änderungen vor dem eigentlichen Schreiben in einem dafür reservierten Speicherbereich, dem Journal, aufzeichnet.
- Ein RAID System wird eingesetzt, damit die Daten gegen Plattenausfälle geschützt werden.
- Alle Dateien werden zum Zeitpunkt der Ablage mit einer zusätzlichen md5-Hashwert-Datei gesichert.
- Das Dateisystem wird an zwei weiteren Standorten mit zwei verschiedenen Sicherungsmethoden abgelegt.
- Vor jeder Sicherung wird das Dateisystem auf die md5-Konsistenz überprüft, um einen möglichen Zerfall zu erkennen.
- Vor jedem Zugriff wird die Unversehrtheit der Primärdaten ebenfalls geprüft.

DIMAG stellt sicher, dass sämtliche digitale Archivalien noch vorhanden und benutzbar sind, auch wenn die Datenbank oder das Dateisystem ausfallen sollte.

Die Ablage der Dateien erfolgt in einer wohl strukturierten Form. Dabei wird eine technisch eindeutige ID verwendet (aID). Diese ID bleibt persistent, auch wenn Objekte innerhalb des Systems oder nach extern verschoben werden.

DIMAG wurde vom Landesarchiv Baden-Württemberg auf der Basis von Linux, Apache, MySQL und PHP (LAMP) programmiert, es läuft auf RAID-5. Die Protokolle SCP und SFTP werden unterstützt.

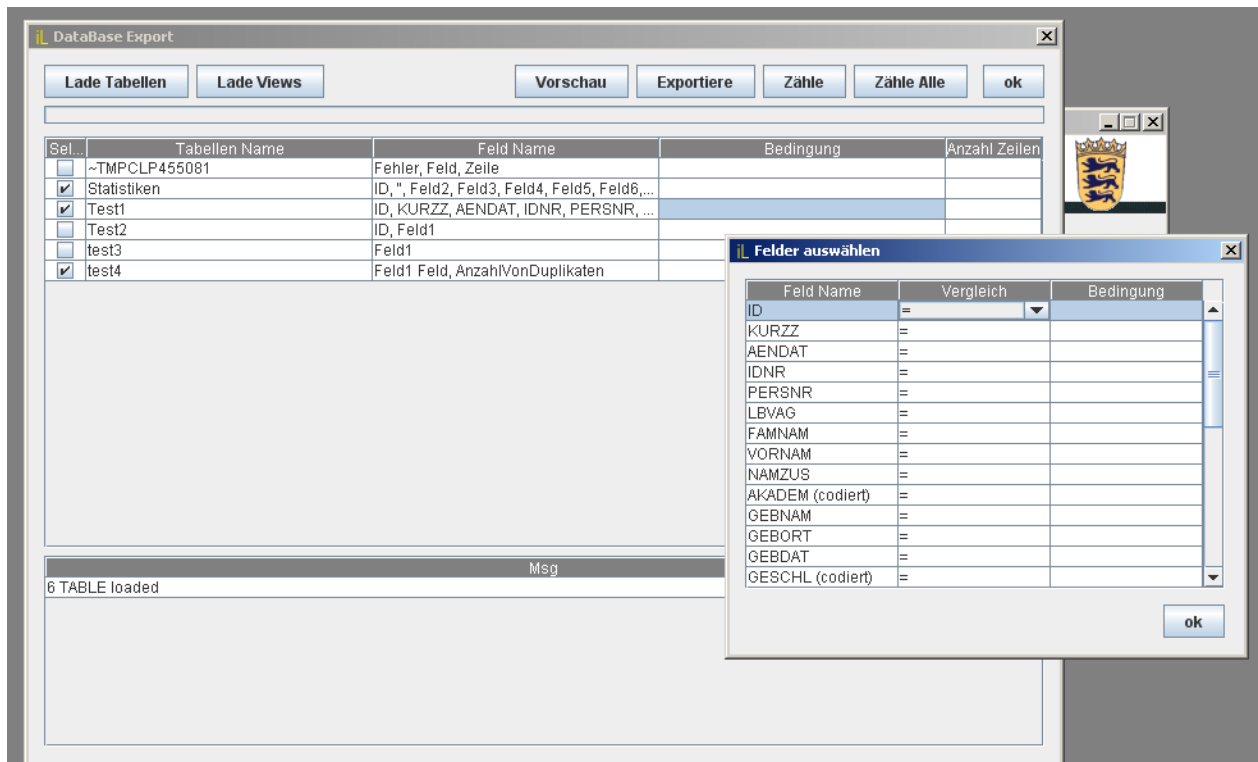


Abb. 4: Datenbankeexport durch IngestList

5. Schnittstellen

Neben der webbasierten Oberfläche gibt es in DIMAG eine standardisierte SOAP² basierte Schnittstelle, welche für die Anbindung an Ingest-Systeme und das Access-System gedacht ist. Eine weitere REST-basierte Schnittstelle ist zur Kopplung des archivischen Fachinformationssystems vorhanden.

6. Rechte

Die DIMAG-Entwicklungspartner sind Inhaber sämtlicher Rechte an dem Programmpaket DIMAG.

² Ursprüngliche Bezeichnung für Simple Object Access Protocol. Ist ein Netzwerkprotokoll, mit dessen Hilfe Daten zwischen Systemen ausgetauscht und Remote Procedure Calls durchgeführt werden können.

7. Literatur

Zu DIMAG

- DIMAG - Das digitale Magazin des Landesarchivs Baden-Württemberg (application/pdf 165.6 KB). Online verfügbar unter: http://www.landesarchiv-bw.de/sixcms/media.php/120/42596/Dimag_Archivtag2007.pdf
- Christian Keitel und Rolf Lang: DIMAG und IngestList – Übernahme, Archivierung und Nutzung von digitalen Unterlagen im Landesarchiv Baden-Württemberg. In: Gerald Maier und Thomas Fritz (Hrsg.): Archivische Informationssysteme in der digitalen Welt. Aktuelle Entwicklungen und Perspektiven. Stuttgart 2010, S. 53 – 63.
- Christian Keitel, Kai Naumann & Rolf Lang: Konzeption und Aufbau eines digitalen Archivs. Von der Skizze zum Prototypen. Online verfügbar unter: http://www.landesarchiv-bw.de/sixcms/media.php/120/42600/aufsatz_labw_aufbau.pdf
- Christian Keitel: Das digitale Landesarchiv Baden-Württemberg. Eine Standortbestimmung. In: *Archivar*, 2015, 68. Jg. (Heft 4), S. 335–341. Online verfügbar unter http://www.archive.nrw.de/archivar/hefte/2015/Ausgabe_4/Archivar_Heft_2015_4-Internet.pdf
- Christian Keitel: DIMAG-Kooperation (Stand 2017). In: Archiv-Nachrichten Niedersachsen. Mitteilungen aus niedersächsischen Archiven, 2017, 21. Jg., S. 50-58. Online verfügbar unter https://www.vna-online.de/wordpress/wp-content/uploads/2019/04/A-NN-21-2017_Internet.pdf
- Christian Keitel: Das Projekt DIMAG. Sachstand 2019. In: Karolína Šimůnková, Milan Vojáček (Hrsg.): 23. Tagung des Arbeitskreises Archivierung von Unterlagen aus digitalen Systemen. 12. und 13. März 2019 in Prag. Prag 2020, S. 21-31. Online verfügbar unter: <https://www.nacr.cz/vyzkum-publikace-akce/publikace/detail-publikace/23-tagung-des-arbeitskreises-archivierung-der-unterlagen-aus-digitalen-systemen-am-12-und-13-marz-2019-in-prag>
- Susanne Laux: DIMAG – Programmpaket zur digitalen Langzeitarchivierung im Verbund. In: *ABI Technik*, 2019, 39. Jg. (Heft 3), S. 202-2012.

Zum Metadatenmodell

- Christian Keitel, Das Repräsentationenmodell des Landesarchivs Baden-Württemberg, in: Susanne Wolf (Hrsg.), Neue Entwicklungen und Erfahrungen im Bereich der digitalen Archivierung: von der Behördenberatung zum Digitalen Archiv. 14. Tagung des Arbeitskreises „Archivierung von Unterlagen aus digitalen Systemen“, München 2010, S. 69 – 82, auch:
http://www.staatsarchiv.sg.ch/home/auds/14/_jcr_content/Par/downloadlist/DownloadListPar/download_8.ocFile/Text%20Keitel.pdf
- DIMAG-Entwicklungsverbund (Hrsg.): DIMAG. Metadatenreferenz für die Archivierung digitaler Unterlagen. [Stand 23. April 2015] https://www.landearchiv-bw.de/sixcms/media.php/120/58735/DIMAG-Metadatenreferenz_2015-04-23.pdf

8. Kontakt

Landesarchiv Baden-Württemberg Eugenstraße 7
70182 Stuttgart
landesarchiv@la-bw.de, 0711/212 4272

Weitere Informationen Prof. Dr. Christian Keitel
christian.keitel@la-bw.de, 0711/212 4276