

# Workshop 2: Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen (kopal)

## Das Universelle Objektformat und koLibRI

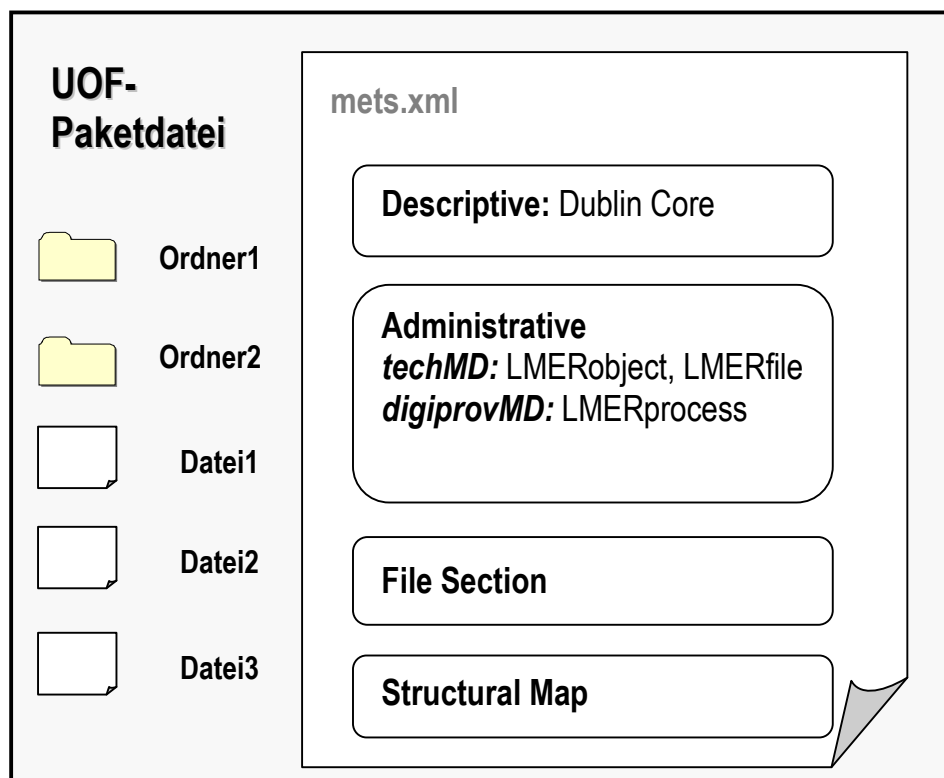
Tobias Steinke, Die Deutsche Bibliothek, Frankfurt/Main

### 1. Das Universelle Objektformat (UOF)

Bei dem Universellen Objektformat (UOF) handelt es sich um eine Beschreibung eines Paketaufbaus für die Archivierung und den Austausch von Objekten und Metadaten. Es wurde im Projekt kopal für das System DIAS von IBM entwickelt, kann aber durch eine sehr generische Ausrichtung im Zusammenhang mit jedem Archivsystem benutzt werden, das sich am OAIS-Referenzmodell orientiert.

Ein Objekt im Sinne des UOF ist eine logische Einheit, die aus beliebig vielen Dateien von beliebiger Art bestehen kann. Diese werden in eine Datei gepackt, z. B. im Format .zip, .tar oder .tar.gz. Jedes Paket muss genau eine XML-Datei mit Namen mets.xml enthalten, die dem Format METS 1.4 (Metadata Encoding and Transmission Standard) entspricht. In den administrativen Abschnitten der METS-Datei müssen sich Metadaten in LMER 1.2 (Langzeitarchivierungsmetadaten für elektronische Ressourcen) befinden. Zusätzlich können optional beliebige Inhaltsmetadaten (z. B. Dublin Core, MODS, MABxml, MARC-XML) enthalten sein.

Eine genaue Spezifikation des UOF ist bei <http://kopal.langzeitarchivierung.de/> erhältlich. Außerdem ist ein offizielles METS-Profil für UOF in Arbeit.



### 2. koLibRI

Die kopal Library for Retrieval and Ingest (koLibRI) ist eine Java-Bibliothek, die im Projekt kopal von Der Deutschen Bibliothek und der Niedersächsischen Staats- und U-

Universitätsbibliothek entwickelt wurde, um ihre lokalen Systeme an das Archivsystem DIAS von IBM anzubinden. koLibRI wurde dabei von Anfang an modular und konfigurierbar entworfen, damit die Software komplett oder in Teilen von anderen Institutionen nachnutzbar ist – auch außerhalb des Kontexts DIAS.

Die einzelnen Module definieren zusammen einen Workflow zur Objektgenerierung, der über XML-Dateien gesteuert wird (sog. Tasklists und Policies). Zu den Modulen gehört ein Metadatenextraktor, der wiederum verschiedene Module für verschiedene Quellen hat (z. B. XMetaDiss, SpringerLink), und ein Metadatengenerator auf Basis des Open-Source-Werkzeugs JHOVE. Des Weiteren generiert ein Modul aus den Metadaten und Objektdateien das Universelle Objektformat (UOF) und ein anderes Modul steuert die Kommunikation mit dem Ingest des Archivsystems (im Fall von kopal ist das DIAS).

Für den Retrieval sieht koLibRI einen lokalen Cache zur Performanzsteigerung vor und gleichzeitig kapselt es den Zugriff zum Archivsystem gegenüber dem Bereitstellungssystem.

Die Open Source koLibRI ist als Beta-Version zusammen mit einer ausführlichen Dokumentation als freier Download bei <http://kopal.langzeitarchivierung.de/> erhältlich.

