

museumdat – das Metadaten-Harvestingformat für Museumsobjektdaten und seine Anwendung in der Praxis

museumdat – the metadata harvesting format for museum object data
and its practical application

Regine Stein, Angela Kailus
Bildarchiv Foto Marburg
Biegenstraße 11, D-35037 Marburg
Tel.: 0049/(0)6421/282-3666, 282-2345, Fax: 0049/(0)6421/282-8931
E-mail: r.stein@fotomarb.de, kailus@fotomarb.de, Internet: www.fotomarb.de

Zusammenfassung:

museumdat ist ein für Recherche und Publikation optimiertes Metadaten-Harvestingformat zur Herstellung der Interoperabilität von Sammlungsdaten aus den Bereichen Kunst, Architektur, Kulturgeschichte, Naturkunde, Technikgeschichte, entwickelt von der Fachgruppe Dokumentation im Deutschen Museumsbund. Das Format trägt den fachspezifischen Anforderungen der Datenmodellierung Rechnung und gestattet differenzierte Suchmöglichkeiten.

Die Publikation von Daten verschiedener Einrichtungen in Internetportalen wird mit Hilfe von museumdat wesentlich vereinfacht, da die Daten in einem standardisierten Format bereit gestellt werden, das auch über Schnittstellen wie OAI-PMH oder SRU ausgewertet werden kann.

museumdat basiert wesentlich auf dem vom J. Paul Getty Trust entwickelten Harvestingformat CDWA Lite. Durch Erweiterungen, die dem ISO-zertifizierten CIDOC-CRM folgen, bietet museumdat aber deutlich verbesserte Möglichkeiten, Daten aus über die bildende Kunst hinausgehenden Sammlungsbeständen abzubilden. museumdat und CDWA Lite werden aktuell zu einem gemeinsamen internationalen Referenzformat weiterentwickelt.

Abstract:

museumdat is a metadata harvesting format specially designed to allow interoperability of collection data from fine art, architecture, science, technology, and cultural or natural history, developed by the Special Interest Group Documentation of the German Museum Association. Its data model is well adapted to expert requirements and allows sophisticated methods for searching and presenting the data.

Using museumdat, the publication of data in multi-source internet portals becomes much easier, as data can be provided in a standardized format that may also be processed through interfaces like OAI-PMH or SRU.

museumdat is based essentially on CDWA Lite, the harvesting format developed by the J. Paul Getty Trust. By modifications in accordance with ISO-certified CIDOC-CRM, museumdat provides considerably improved options to represent data from collections other than the field of fine art. By now, museumdat and CDWA Lite are being developed into a joint international reference format.

Museen und andere Sammlungen messen dem Internet als Präsentationsplattform und Aufmerksamkeitsfaktor zunehmende Bedeutung bei. In den letzten Jahren ist daher eine Reihe von Institutionen dazu übergegangen, ihre Sammlungsbestände in eigenen Online-Datenbanken zu publizieren. Für den Nutzer sind diese Angebote nicht einfach zugänglich: wertvolle Bestände sind im

Netz verstreut, ohne einschlägige Vorkenntnis schlecht auffindbar und sehr unterschiedlich strukturiert.

Vor diesem Hintergrund wächst die Bedeutung von Portalen, die die Recherchierbarkeit von Beständen über Sammlungsgrenzen hinaus gewährleisten. Sie bieten den Sammlungen die Möglichkeit, ihre Bestände nicht-exklusiv zu zeigen und damit an mehreren Orten im Netz präsent zu sein.

Die Integration eines Datenbestandes in ein Portal war allerdings bislang mit hohen Aufwänden für maßgeschneiderte Übernahmelösungen verbunden, die viele potentielle Interessenten abschreckten.

Mit museumdat¹ gibt es nun einen im XML-Format beschriebenen Metadatenstandard, der die Übernahme von Objektdaten aus verschiedensten Sammlungsbereichen wie Kunst, Architektur, Kulturgeschichte, Naturkunde und Technikgeschichte wesentlich vereinfacht und die Interoperabilität von Beständen ganz unterschiedlicher Provenienz gewährleistet.

Das von der AG Datenaustausch der Fachgruppe Dokumentation im Deutschen Museumsbund² entwickelte XML-Schema basiert auf internationalen Standards, im Folgenden wird das Format in seinen Grundzügen vorgestellt und seine Verwendung anhand von Praxisbeispielen illustriert.

museumdat baut in weiten Teilen auf das Harvestingformat CDWA Lite³ auf, welches in den USA vom J. Paul Getty Trust gemeinsam mit ARTstor genau mit dem Ziel entwickelt wurde, Kerndaten für museumsübergreifende Repositorien auf Basis des Open Archives Initiative Harvesting Protokolls OAI-PMH bereitzustellen. Damit sollte eine Alternative mit reichhaltigeren Recherchemöglichkeiten im Vergleich zu dem vielfach in Portalen verwendeten Dublin Core-Standard angeboten werden, da dieser die fachspezifischen Anforderungen der Museumsdokumentation nicht berücksichtigt.

Schon die Referenz im Namen CDWA – Categories for the Description of Works of Art – verdeutlicht jedoch, dass das Format primär auf die Sammlungsbestände in Kunstmuseen ausgerichtet ist. Dies schlägt sich sowohl in der Auswahl der Kerndatenfelder als auch in der Definition der Pflichtfelder nieder: So sind z.B. Angaben zum Künstler/Hersteller eines Objekts nicht nur die einzigen im Format enthaltenen Personenangaben, sondern sie sind auch verpflichtend. Damit erwies sich der Fokus von CDWA Lite als zu eng im Hinblick auf Portale, in denen Objekte aus unterschiedlichsten – etwa kultur-, technik- oder naturhistorischen – Sammlungsbeständen nachgewiesen werden.

museumdat ist nun eine Rekonfiguration des CDWA Lite-Formats, das dessen Konstruktionsprinzipien und Elementenauswahl im Wesentlichen beibehält – damit konnten in Abstimmung mit dem J. Paul Getty Trust insbesondere dessen sehr gut ausgearbeitete Dokumentation und die Schema-Definition genutzt werden. Ein entscheidendes Prinzip von CDWA Lite wie auch museumdat ist die Trennung in die für die Datenanzeige optimierten Elemente einerseits und die für differenzierte Such- und Navigationsmöglichkeiten optimierten Elemente andererseits. Darüber hinaus werden Gruppenstrukturen in Sets und Wiederholelemente in Wrappern („Umschlägen“) zusammengefasst, Referenzen in kontrolliertes Vokabular können über XML-Attribute gebildet werden. Diese Prinzipien beziehen sich vor allem auf die deskriptiven Metadaten-Elemente. Schwerpunkte bei den administrativen Elementen liegen auf der Referenzierung von Belegen, insbesondere Abbildungen, zu einem Objekt sowie auf dem Verweis auf weitere Internetressourcen, dabei insbesondere die Publikation des Objekts in einer eigenen Online-Datenbank der datenliefernden Institution.

Die Rekonfiguration von Elementen beschränkt sich auf den Bereich der deskriptiven Metadaten, als Grundlage galt der Arbeitsgruppe Datenaustausch dabei das vom internationalen Fachkomitee

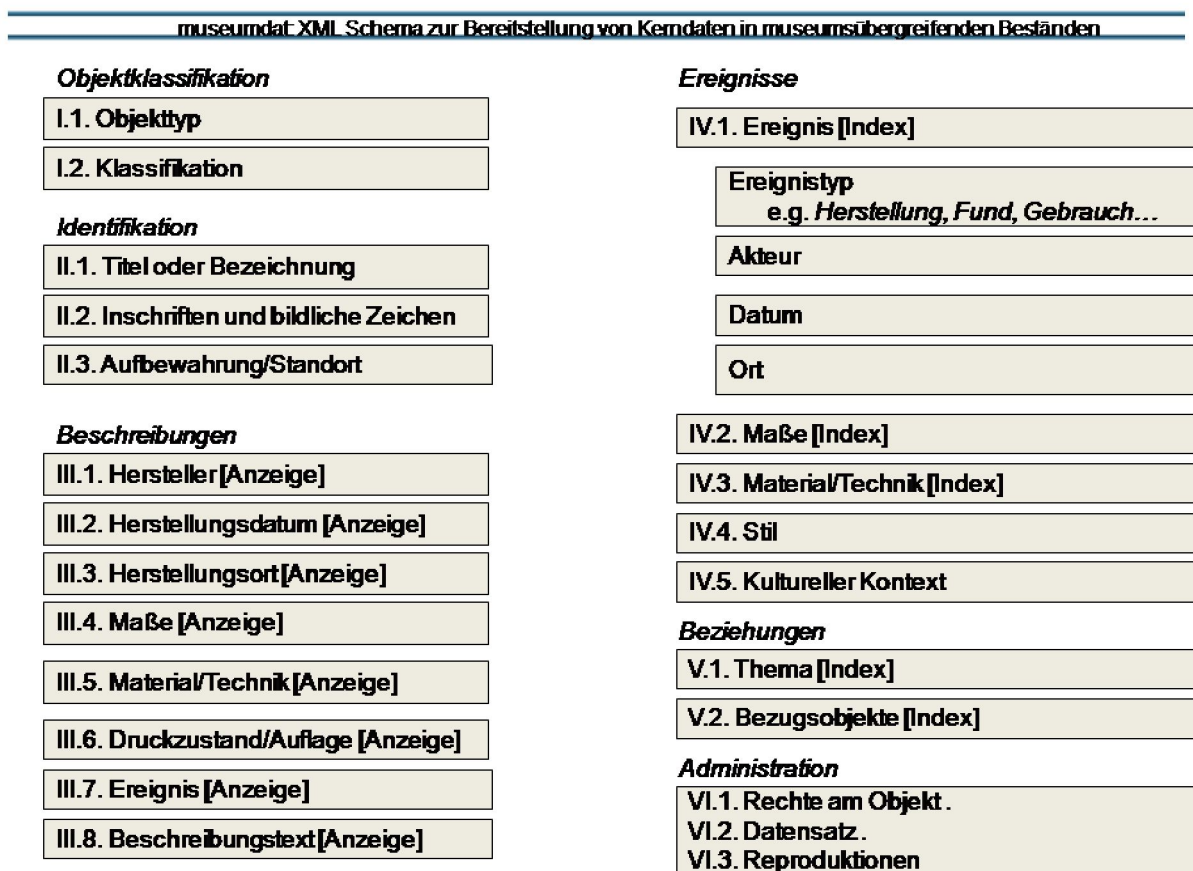
¹ <http://www.museumdat.org>

² <http://www.museumbund.de> > Fachgruppen & Arbeitskreise > Fachgruppe Dokumentation

³ http://www.getty.edu/research/conducting_research/standards/cdwa/cdwalite.html

Dokumentation CIDOC des Internationalen Museumsbundes ICOM entwickelte „Conceptual Reference Model“⁴, das als ISO-Standard 21127 das Referenzmodell für die Entwicklung von Metadatenformaten im Bereich des Kulturerbes bildet. Das CIDOC-CRM beschreibt die in der Museumsdokumentation verwendeten impliziten und expliziten Konzepte und Beziehungen in einem objekt-orientierten Modell, es ist sozusagen der „semantic glue“ zur Integration und zum Datenaustausch zwischen verschiedenen Ressourcen.

Das Ausgangsformat CDWA Lite wurde also auf Basis einer Analyse mit dem CIDOC CRM restrukturiert und entlang der im CRM Core⁵ definierten Metadaten-Elemente in Wrappern organisiert: I. Objektklassifikation, II. Identifikation, III. Beschreibungen, IV. Ereignisse, V. Beziehungen sowie VI. Verwaltungsangaben. Eine Übersicht der Elemente bietet die folgende Abbildung.



Der wesentliche Kunstgriff bei der Rekonfiguration ist die Übernahme des ereignisorientierten Ansatzes des CIDOC-CRM, der die Beziehung zwischen einem Museumsobjekt einerseits und Personen (bzw. Institutionen)-, Datierungs- und Ortsangaben immer als ereignisvermittelt darstellt: Künstler und Objekt stehen etwa über das Herstellungsereignis in Bezug zueinander, für ein naturkundliches Objekt hingegen ist die Angabe des Sammlers als die am Sammelereignis beteiligte Person die wichtigste Information.

⁴ <http://cidoc.ics.forth.gr>

⁵ http://cidoc.ics.forth.gr/working_editions_cidoc.html#crm_core

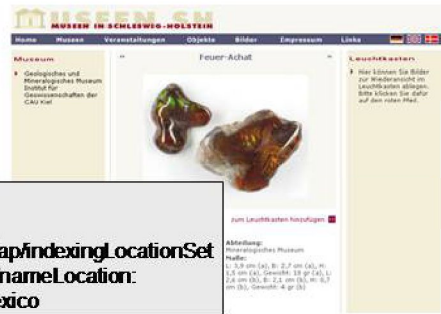
Dementsprechend wurde das CDWA Lite-Element zur Angabe des Künstlers/Herstellers eines Objekts in museumdat zunächst verallgemeinert zu einem Element zur Angabe von Akteuren, also jeglichen Personen oder Institutionen, die über ein Ereignis vermittelt mit dem Objekt in Beziehung stehen (IV.1.1.2. indexingActorSet). Das museumdat-Format unterscheidet darüber hinaus im Gegensatz zum CDWA Lite ein Element zur Angabe von (mit dem Objekt in unterschiedlichster Weise in Beziehung stehenden) Orten (IV.1.1.4. indexingLocationWrap) sowie ein Element zur Angabe von Identifikatoren und Aufbewahrungsinformationen (II.3. repositoryWrap). Schließlich wird die Information zu Akteuren, Datierung und Ort – letztere beiden werden in CDWA Lite durch Angabe eines Typs näher qualifiziert, alle drei Elemente sind aber völlig unabhängig voneinander definiert – in museumdat in einem Ereignis-Element zusammengefasst, welches nun durch einen Ereignistyp wie Herstellung, Gebrauch, Fund, Sammelereignis, Zerstörung etc. näher qualifiziert wird.

Als Pflichtelemente werden nur vier Elemente deklariert: Der Objekttyp, ein Objekttitle resp. Objektname sowie Datensatz-ID und Datensatz-Quelle (i.d.R. die aufbewahrende Institution). Damit wird der Diversität der Objekte und den unterschiedlichen Schwerpunkten der Sammlungsdokumentation Rechnung getragen.

Die folgenden Beispiele veranschaulichen den Mehrwert, der durch die Rekonfiguration mit museumdat erreicht wird:

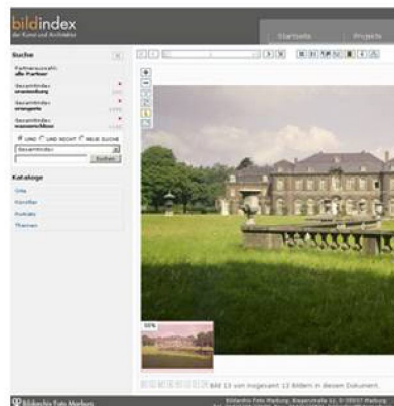


```
indexingEventSet
event Type: Fund
indexingLocationWrap/indexingLocationSet
nameLocationSet/nameLocation:
Hermosillo / Mexico
```



museumdat – Beispiele

```
indexingEventSet
event Type: Herstellung
indexingDates
earliestDate: 1947
latestDate: 1951
indexingLocationWrap/indexingLocationSet
nameLocationSet/nameLocation: Deutschland
nameLocationSet/nameLocation:
Deutsche Demokratische Republik
indexingEventSet
event Type: Gebrauch
indexingActorSet
nameActorSet/nameActor: Pieck, Wilhelm
roleActor: Nutzer
indexingLocationWrap/indexingLocationSet
nameLocationSet/nameLocation: Berlin
indexingDates
earliestDate: 1950
latestDate: 1960
```



```
indexingEventSet
event Type: Bau
indexingActorSet
nameActorSet/nameActor:
Gottfried Laurenz Pictorius
roleActor: Architekt
indexingDates
earliestDate: 1718
latestDate: 1720
indexingEventSet
event Type: Umbau
indexingActorSet
nameActorSet/nameActor:
Johann Conrad Schlaun
roleActor: Architekt
indexingDates
earliestDate: 1725
latestDate: 1733
indexingEventSet
event Type: Erweiterung
indexingDates
earliestDate: 1913
latestDate: 1913
```

- Im Bestand des Deutschen Historischen Museums in Berlin befindet sich der „Bücherschrank aus dem Arbeitszimmer Wilhelm Piecks in Schloss Niederschönhausen“. Dieses Möbelstück wurde von einem unbekanntem Schreiner gefertigt, seine historische Bedeutung resultiert

einzig und allein aus der Tatsache, dass es vom ersten (und einzigen) Präsidenten der Deutschen Demokratischen Republik gebraucht wurde. Der Bezug zwischen diesem Möbelstück und Wilhelm Pieck kann in museumdat in einem Ereignis vom Typ „Gebrauch“ mit dem Akteur „Wilhelm Pieck“ dokumentiert werden.⁶

- Zur Mineraliensammlung des Instituts für Geowissenschaften der CAU Kiel gehört auch ein Feuer-Achat, dessen Fundort in Mexiko bekannt ist, zu dem aber selbstverständlich kein Künstler oder Hersteller anzugeben ist.⁷
- Der Bildindex der Kunst und Architektur weist in seiner Dokumentation die Baugeschichte der Orangerie des Wasserschlosses Nordkirchen mit Bau-, Umbau- und Erweiterungsphase aus – die Information zu beteiligten Architekten und Zeitangaben kann in museumdat in den jeweiligen Ereignis-Sets zusammengeführt werden.⁸

Im Ergebnis ist museumdat ein zum CIDOC-CRM kompatibles Metadaten-Format, das bereits von verschiedenen Portalbetreibern in Deutschland wie dem Bildarchiv Foto Marburg, dem Museumsportal Schleswig-Holstein, dem BAM-Portal, und nicht zuletzt auch für die Europeana für die Zulieferung einer breiten Vielfalt von Objektdaten genutzt wird. Eine Reihe von Herstellern von Sammlungsinventarisationssoftware hat museumdat bereits als Standardexportformat in die Programmpakete integriert oder implementiert dies derzeit. Museen als Datenanbietern steht damit ein Weg offen, ihre Sammlungsdaten in standardisiertem Format verschiedenen Service-Anbietern weiterzugeben, ohne die hauseigene Sammlungsdokumentation zwingend umstellen zu müssen. Der Anbieter entscheidet selbst, welche über die drei Pflichtfelder hinausgehenden Objektinformationen publiziert werden sollen. Dabei mag man sich auf Basisinformationen beschränken, durch seine modulare Struktur ist museumdat aber auch in der Lage, eine sehr detaillierte und ausführliche Dokumentation angemessen abzubilden. Den Aufwand des Mappings, der Übertragung des hauseigenen Datenformats in ein anderes Format, braucht man im Idealfall nur noch ein Mal zu betreiben, um vielfältig verwendbare Standard-Exportdaten verfügen zu können.

Service-Anbieter haben umgekehrt ein Standard-Importformat zur Verfügung – wenn sie nicht gleich museumdat als Portal-Format verwenden.

museumdat bildet somit die Schnittstelle, um den Aufbau von Verbundkatalogen auch für Museums-Datenbestände effizient zu gestalten. Dabei kann museumdat zusammen mit unterschiedlichen technischen Protokollen genutzt werden: Zum maschinellen Metadaten-Harvesting über das Protokoll der Open Archives Initiative OAI-PMH,⁹ zum verteilten Suchen über die SRU-Schnittstelle¹⁰ der Library of Congress. Aber auch technisch weniger anspruchsvolle Lösungen wie z.B. ein einfacher FTP-Upload sind mit museumdat selbstverständlich möglich.

Die Erweiterung und Verallgemeinerung des Fokus von museumdat gegenüber seinem Ausgangsformat CDWA Lite sowohl in inhaltlicher als auch technischer Hinsicht wird mittlerweile auch von den CDWA Lite-Entwicklern reflektiert. Mit der Etablierung einer transatlantischen Arbeitsgruppe von Getty, ARTstor, OCLC und der Fachgruppe Dokumentation seit Frühjahr 2008 wird das Ziel verfolgt, die beiden Formate – CDWA Lite und museumdat – wieder in ein gemeinsames Schema mit dem erweiterten Fokus zusammenzuführen. Damit dürfte sich museumdat, das bereits breite Anwendung im deutschen und europäischen Umfeld findet, auch zu einem echten internationalen Standardformat entwickeln.

⁶ <http://www.dhm.de/datenbank>

⁷ <http://www.museen-sh.de/ml/digicult.php?s=1&t=1&sparte=museen&pid=&ps=1>

⁸ <http://www.bildindex.de>

⁹ <http://www.openarchives.org/>

¹⁰ <http://www.loc.gov/standards/sru/>