

Empfehlungen für die Erfassung

Ebenen der Objektbeschreibung

Bevor man mit der Erfassung eines Datensatzes zur Beschreibung eines Objekts beginnt, sollte man sich darüber im Klaren sein, was genau der Gegenstand des Datensatzes sein soll, denn abhängig davon können die zu erfassenden Informationen stark variieren. Hierbei lassen sich verschiedene Ebenen der Betrachtung voneinander unterscheiden, die im Folgenden anhand einiger Beispiele erläutert werden.

Einzelobjekt- oder objektgruppen-bezogener Datensatz

Bei der Beschreibung eines mehrteiligen Objekts hat man die Wahl: Möchte man es in seiner Gesamtheit beschreiben, z. B. als eine dreiteilige Skulpturengruppe oder ein mehrteiliges Retabel, oder möchte man die einzelnen Teile des Objekts in einzelnen Datensätzen genau beschreiben und diese anschließend aufeinander beziehen? Wenn man pro Teil eine gewisse Anzahl voneinander abweichender Informationen vorfindet (z. B. unterschiedliche Urheber:innen, Themen, Techniken, Formate), dann lohnt es sich, für jeden Teil einen einzelnen Datensatz anzulegen, damit diese Informationen detailliert erfasst werden können. Auch wenn man nur einen oder wenige Teile eines heute aufgelösten Gesamtwerks in der eigenen Sammlung aufbewahrt, sollte man den Fokus der Datensätze jeweils auf den einzelnen Teil richten. Mehrteilige Werke als Objektgruppe zu erfassen, ist hingegen nur sinnvoll, wenn die zu erfassenden Informationen für alle Teile identisch sind, wenn die Bedeutung des Objekts als Werk nur in seiner Gesamtheit erschlossen werden kann oder wenn man aus Zeitgründen in einem ersten Schritt der digitalen Katalogisierung nur zu Gruppen zusammengefasste Objekte in einem Datensatz beschreiben kann (z. B. ein mehrteiliger frühneuzeitlicher Tafelaufsatz).

Je nachdem, ob das Werk in seiner Gesamtheit oder aber ein einzelner Teil des Werks im Fokus des Datensatzes steht, ändert sich der Inhalt einiger Datenfelder.

Empfehlungen für die Erfassung

Beispiel für ein mehrteiliges Werk:

Objekttyp: Gemäldezyklus (<https://www.wikidata.org/wiki/Q16905563>)

Titel: Gemäldezyklus der Maria de' Medici

Objektbeschreibung: Serie von 24 monumentalen Gemälden, beauftragt von Maria de' Medici, Witwe von Heinrich IV. von Frankreich für das Palais du Luxembourg in Paris.

Datensatztyp: Objektgruppe (Katalogisierungsebene), <http://terminology.lido-schema.org/lido00453>

Beispiel für ein Einzelobjekt:

Objekttyp: Ölskizze (<https://www.wikidata.org/wiki/Q1964917>)

Titel: Apotheose Heinrichs IV. und Proklamation der Regentschaft Maria de' Medicis (Skizze zum Medici-Zyklus)

Objektbeschreibung: Höhe- und zugleich Wendepunkt im Medici-Zyklus ist die vorliegende Szene, die sich an der Stirnwand gegenüber dem Eingang in der Galerie befand. Sie leitet vom Leben Maria de' Medicis mit dem König zum Fortleben als verwitwete Regentin über, das die zweite Hälfte des Zyklus bestimmt.

Datensatztyp: Einzelobjekt (Katalogisierungsebene), <http://terminology.lido-schema.org/lido00141>

Objektbezogener oder abbildungsbezogener Datensatz

Ebenso gilt es zu unterscheiden, ob sich der Datensatz auf ein bestimmtes künstlerisches Werk bezieht, das einer entsprechenden bildmedialen Reproduktion zugrunde liegt, oder ob die Reproduktion selbst der Gegenstand des Datensatzes ist. Je nach dem Fokus des Datensatzes – entweder auf das abgebildete Werk oder auf die zumeist analoge Reproduktion – ändern sich auch hier die zu erfassenden Informationen erheblich.

Beispiel für einen objektbezogenen Datensatz:

Objekttyp: Bozzetto, Abguss (Skulptur), Rundplastik

Titel: Kauernde

Urheber: Georg Kolbe

Material/Technik: Kunststein (Stucco)

Maße: Höhe: 45 cm, Breite: 17 cm, Tiefe: 22 cm

Datierung: 1925–1927

Standort: Museum für Kunst und Kulturgeschichte, Philipps Universität Marburg

Inventarnummer: 126

Beispiel für einen abbildungsbezogenen Datensatz:

Objekttyp: Silbergelatineabzug

Titel: Kauernde

Urheberin: Margit Schwartzkopff

Material/Technik: Gelatinesilberverfahren

Maße: 19,8×13,5 cm

Datierung: 1925–1927

Standort: Georg Kolbe Museum, Berlin

Inventarnummer: GKFo-0292_002

Werk- oder exemplarbezogene Informationen bei Auflagen

Als **Werk** wird in diesem Handbuch gemäß des CIDOC CRM die geistige Schöpfung verstanden, die einem physischen Objekt zugrunde liegt oder die sich in der Summe aller Eigenschaften erfassen lässt, die auf alle einzelnen Exemplare eines seriell hergestellten Werks zutreffen. Gerade bei seriell hergestellten Objekten (Auflagen) ist es von entscheidender Bedeutung, bereits während des Erschließens darauf zu achten, worauf sich ein einzelner Datensatz oder aber bestimmte Teilinformationen innerhalb eines Datensatzes beziehen. In diesem Handbuch wird daher immer dann das Wort **Werk** (= E28 Conceptual Object) benutzt, wenn auf die geistige Schöpfung Bezug genommen wird. Als **Objekte** werden hingegen unikale physische Gegenstände verstanden oder bestimmte physische Exemplare eines seriellen Werks (= E19 Physical Object).

Auf einer konzeptuellen Ebene, der sogenannten Werkebene, werden alle Eigenschaften beschrieben, die auf alle physischen Realisierungen des Werks zutreffen. Diese Informationen entsprechen üblicherweise den Einträgen in Werkverzeichnissen und in Werknormdatensätzen. Dazu zählen:

- ▶ der Titel des Werks
- ▶ die Namen und Rollen der an der geistigen Schöpfung beteiligten Personen
- ▶ die Beschreibung des Darstellungsinhalts, der Bildthemen
- ▶ der Verweis auf Werkverzeichnisse
- ▶ der Verweis auf Verwandtschaften zu anderen Werken

Auf der konkreten Ebene eines bestimmten Exemplars sind alle Informationen anzusiedeln, die sich auf dieses eine Objekt beziehen. Merkmale, die nur dieses Exemplar hat und die es kennzeichnen, müssen aus Gründen der Identifizierbarkeit erfasst werden. Dazu gehören zum Beispiel:

- ▶ das Material, aus dem das physische Exemplar besteht und die Technik, mit der es hergestellt wurde
- ▶ handschriftliche Vermerke auf dem Exemplar

Empfehlungen für die Erfassung

- ▶ der physische Standort des Exemplars
- ▶ die Inventarnummer dieses Exemplars in einer bestimmten Sammlung
- ▶ die individuelle Auflagennummer (z. B. 6 von 8)
- ▶ der Erhaltungszustand des Exemplars
- ▶ die Maße des Exemplars
- ▶ die Datierung, wann genau dieses Exemplar hergestellt worden ist (z. B. Datierung des konkreten Abgusses)
- ▶ konkrete Bezüge, die nur für dieses Exemplar gelten (z. B. Teil/Ganzes-Beziehungen)

So sind beispielsweise sämtliche Abgüsse von Georg Kolbes Werk „Die Kauernde“ als individuelle **Exemplare** zu betrachten, die über dieselben Eigenschaften auf der konzeptuellen Werkebene verfügen, so den gemeinsamen Schöpfer „Georg Kolbe“, den Werktitel „Kauernde“ oder den gemeinsamen Darstellungsinhalt eines knienden weiblichen Akts. Sie können sich jedoch voneinander unterscheiden hinsichtlich einiger Eigenschaften auf der Ebene der Auflage, wie Herstellungstechnik, Größe oder Inschriften auf der Plinthe. So etwa sind die Exemplare einer Auflage der „Kauernden“ in Kunststein (Stucco) gegossen, diejenigen aus einer anderen Auflage jedoch als Bronzeguss ausgeführt worden. Dasselbe Werk gibt es also in verschiedenen Realisierungen, die auch auf unterschiedlichen Gipsmodellen beruhen. Ein Gipsmodell (Gi388) befindet sich im Georg Kolbe Museum in Berlin, (<https://sammlung.georg-kolbe-museum.de/de/objekte/skizze-des-entwurfs-der-kauernden/64515>), es bildete aber vermutlich nicht die unmittelbare Vorlage der bekannten Auflagen. Darüber hinaus gibt es Eigenschaften, die nur für ein bestimmtes Exemplar gelten, wie sein Standort oder die Inventarnummer. Ist man sich der Unterscheidung nach konzeptuellen Ebenen und der Ebene der konkreten Objekte bewusst, so trägt dies dazu bei, die Organisation von Informationen auch in sammlungsübergreifenden Portalen zu erleichtern und nach bestimmten Kriterien zusammengehörige Objekte gemeinsam präsentieren zu können.

Es ist auch möglich, Werke, die nicht physisch realisiert worden sind, in einzelnen Datensätzen als eine eigene Entität zu beschreiben. Das ist etwa dann sinnvoll, wenn man über die Verlinkung einzelner Datensätze inhaltliche Beziehungen zwischen den Teilen eines konzeptuellen Ganzen, wie einem Gemäldezyklus, herstellen möchte. In einem solchen Fall kann man einen Datensatz zum Gemäldezyklus als Werkzusammenhang erstellen, z. B. <https://d-nb.info/gnd/1200051734> als Werknormdatensatz für Tizians sechsteiligen Gemäldezyklus „Poesie“, den er während seines Aufenthalts am spanischen Hof in Madrid zwischen 1553 und 1562 für König Philipp II. geschaffen hat. Ein Gemälde aus dem Zyklus befindet sich im Museo del Prado in Madrid, eines in Edinburgh in den National Galleries of Scotland, ein weiteres in Boston im Isabella Stewart Gardner Museum. Würde

nun jede Sammlung jeweils auf den URI (Uniform Resource Identifier) des Datensatzes in der GND verweisen, der den gesamten Zyklus beschreibt, könnte man virtuell alle Teile wieder zusammenfügen, die im Lauf der Geschichte in die verschiedenen Sammlungen Eingang gefunden haben.

Ereignistypen in LIDO

Ein wesentliches Grundprinzip des LIDO-XML-Schemas ist – in Übereinstimmung mit dem zugrunde liegenden Referenzmodell CIDOC CRM – die Zuordnung einzelner Informationen zu Ereignissen, die mit einem Objekt verknüpft sind. Diese können je nach Art des im Datensatz beschriebenen Objekts sehr unterschiedlich sein.

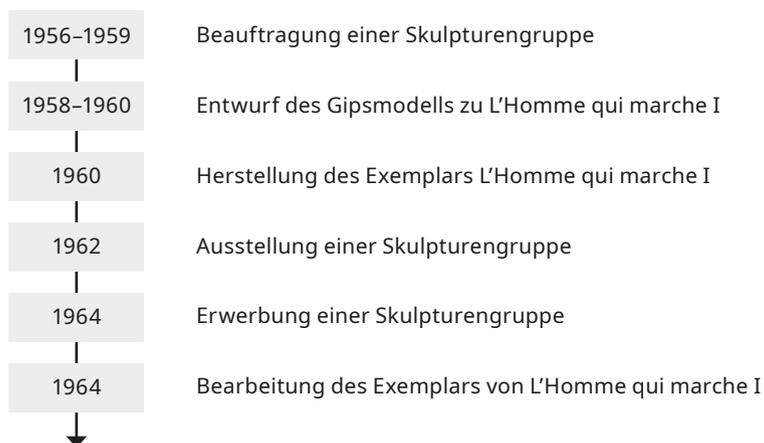
Bei **unikalen Werken** beschreibt man den gesamten Entstehungsprozess in der Regel als ein einziges, zentrales Ereignis, das zur Entstehung eines Objekts geführt hat: Jemand hat an einem Ort zu einem bestimmten Zeitpunkt oder in einem bestimmten Zeitraum unter Verwendung ausgewählter Materialien und Techniken ein Kunstwerk geschaffen. Zum Beispiel hat Peter Paul Rubens in Antwerpen zwischen 1609 und 1610 mit Ölfarben auf Leinwand ein Doppelporträt von seiner Frau Isabella Brant und sich selbst gemalt. Das Gemälde mit dem Titel „Rubens und Isabella Brant in der Geißblattlaube“ befindet sich heute in der Alten Pinakothek in München (<https://www.sammlung.pinakothek.de/de/artwork/o5xrMEr47X>). Dieses eine Ereignis, das zur Entstehung des Gemäldes geführt hat, erhält im LIDO-Datensatz den Ereignistyp Herstellung (<http://terminology.lido-schema.org/lido00007>). Die genannten Informationen zur beteiligten Person, zu Ort, Zeitraum und den verwendeten Materialien werden im Rahmen eines Datenexports im LIDO-XML-Schema innerhalb dieses Event-Sets abgelegt und bleiben so auch im Rahmen der Datenbereitstellung eng miteinander verknüpft.

Bei **seriell hergestellten Objekten**, wie Bronzeabgüssen, können jedoch einzelne Ereignisse im Entstehungsprozess beschrieben werden, die jeweils von verschiedenen Personen an bestimmten Orten und/oder zu verschiedenen Zeitpunkten ausgeführt worden sind. Dies sei an einem Beispiel verdeutlicht: Gordon Bunshaft, der Architekt der New Yorker Chase Manhattan Bank, beauftragte gemeinsam mit dem New Yorker Kunsthändler Pierre Matisse den Künstler Alberto Giacometti im Jahr 1956 damit, eine Skulpturengruppe für die davorliegende Plaza zu entwerfen. Giacometti entwickelte daraufhin viele Modelle für Einzelfiguren, die zueinander in Beziehung gesetzt werden sollten. Eine Vielzahl der Entwürfe zerstörte er selbst wieder. Schließlich scheiterte das Projekt, weil Alberto Giacometti die Dimensionen des realen Umraums nicht kannte und doch

Empfehlungen für die Erfassung

die Figuren so entwickeln musste, dass sie sich gegenüber dem sie hinterfangenden Wolkenkratzer behaupten konnten. Doch hielt Giacometti an den Ideen für die einzelnen Figuren fest und entwickelte sie weiter. Ab 1958 arbeitete er in seinem Pariser Atelier an einem großen Gipsmodell eines schreitenden Mannes, das später unter dem Titel „L’Homme qui marche I“ Teil einer fünfteiligen Skulpturengruppe wurde. Das Gipsmodell befindet sich heute im Kunsthaus Zürich. Erste Abgüsse dieses Gipsmodells wurden bereits 1960 in der Gießerei „Susse Fondeur“ in Paris hergestellt. Dazu wurde das von Alberto Giacometti geschaffene Modell zuerst in insgesamt vier Einzelteile zersägt und dann einzeln abgeformt. Kopf und Füße wurden im Wachsauerschmelzverfahren gegossen, die beiden Beine hingegen im Sandgussverfahren. Die Bronzen wurden anschließend in der Gießerei patiniert. 1962 stellte Giacometti ein Exemplar der Figurengruppe auf der Biennale in Venedig aus und wurde dafür mit dem Großen Preis für Bildhauerei ausgezeichnet. Alberto Giacometti überließ ein Exemplar des „L’Homme qui marche I“ seinem Pariser Kunsthändler Aimé Maeght, da dieser die Herstellungskosten übernommen hatte. Gemeinsam mit einer zweiten Version, dem „L’Homme qui marche II“, zwei Fassungen einer überlebensgroßen stehenden Frau (Grand Femme debout I und II) und einem großen Kopf (Grand Tête) installierte Giacometti das Exemplar 1964 auf der Terrasse der Fondation Marguerite et Aimé Maeght in Saint-Paul-de-Vence nördlich von Nizza. Während des Aufstellens der Figurengruppe stellte Giacometti fest, dass ihm die Farbe der Patina nicht gefiel. So übermalte er die Oberfläche des Abgusses des „L’Homme qui marche I“ mit Ölfarbe. Das Werk ist so berühmt geworden, dass es seit 1998 die 100-Franken-Note ziert. Ein weiterer Abguss, der aus der ehemaligen Sammlung der Dresdner Bank stammt, wurde schließlich 2010 bei Sotheby’s in London zum Preis von etwa 75 Millionen Euro verkauft, dem damals höchsten je für ein Kunstwerk erzielten Auktionsergebnis.

Abb. 3 Darstellung der mit einem Abguss des „L’Homme qui marche I“ verbundenen Ereignisse im Zeitstrahl



Die oben beschriebenen Ereignisse können in LIDO unter Verwendung folgender Ereignistypen abgebildet werden:

1. Zuerst gibt es die **Beauftragung** (<http://terminology.lido-schema.org/lido00226>) 1956 durch Gordon Bunshaft und Pierre Matisse an Alberto Giacometti. Briefe belegen, dass es einen fortwährenden Austausch zwischen dem Vermittler des Auftrags, Pierre Matisse, und dem Künstler gegeben hat, bis das Projekt schließlich in seiner ursprünglichen Form aufgegeben wurde.
2. Alberto Giacometti formt in der Zeit von 1958 bis 1960 das Gipsmodell, welches die Vorlage für alle Abgüsse bildet. Dies entspricht dem Ereignistyp **Entwurf (Aktivität)** (<http://terminology.lido-schema.org/lido00224>). Auch wenn man in einem Datensatz ein Exemplar einer Auflage beschreiben möchte, dessen konkretes Modell nicht mehr erhalten ist, kann man Informationen zur Herstellung des Modells in den Datensatz integrieren, indem man ein solches Entwurfsereignis anlegt. Wenn auf ein konkretes Modell verwiesen werden kann, wie es bei Alberto Giacomettis „L’Homme qui marche I“ der Fall ist, so wird ein Bezug zu diesem Modell hergestellt, siehe auch den folgenden Abschnitt **Beziehungen zu anderen Werken und Objekten**.
3. In der Gießerei werden 1960 erste Exemplare in Bronze gegossen. Das entspricht der **Herstellung des Exemplars** (<http://terminology.lido-schema.org/lido01096>). Auch die Patinierung der Bronze gehört zum Ereignis der Herstellung des individuellen Exemplars.
4. Die Präsentation des „L’Homme qui marche I“ neben weiteren Bronze-güssen und Gemälden auf der Biennale in Venedig 1962 wird in LIDO als **Ausstellung (Aktivität)** (<http://terminology.lido-schema.org/lido00225>) abgebildet.
5. Schließlich wird das Exemplar des „L’Homme qui marche I“ gemeinsam mit vier weiteren Skulpturen von der Fondation Maeght 1964 erworben. Diese Tatsache wird im Rahmen des Ereignisses der **Erwerbung** (<http://terminology.lido-schema.org/lido00001>) beschrieben.
6. In den Fällen, wo Giacometti einzelne Bronzeabgüsse zusätzlich bemalt hat, gehört dieser Arbeitsschritt zum Ereignis der **Bearbeitung** (<http://terminology.lido-schema.org/lido00006>).

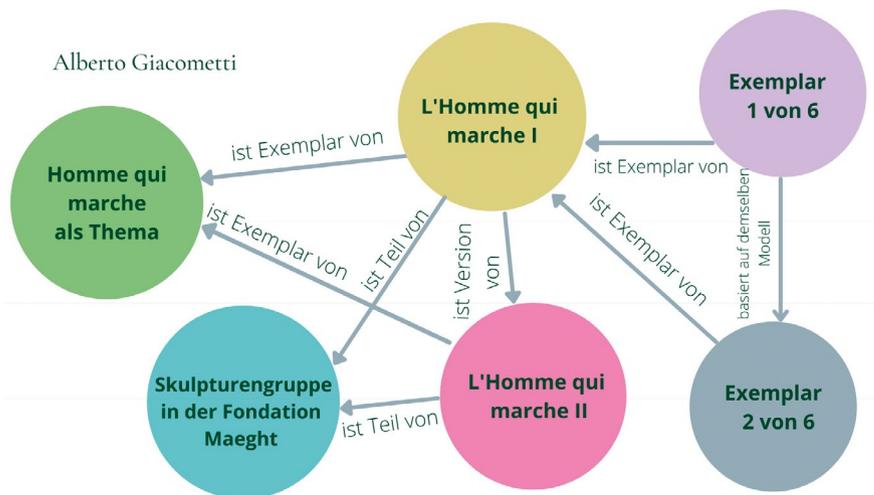
Siehe auch: **Ereignistyp** im 5. Block: Ereignisse in der Objektgeschichte und **Ereignistyp** im 6. Block: Provenienz.

Die Idee des Künstlers, einzelne Figuren zu einer Skulpturengruppe zusammenzufassen, wie die zwei Versionen des „L’Homme qui marche“, zwei Versionen der „Grand Femme debout I und II“ und den „Grand Tête“, kann man in einem eigenen Datensatz auf konzeptueller Werk Ebene beschreiben und dann mit den einzelnen Exemplaren verknüpfen. Die Relation vom Datensatz des Exemplars zum werkbezogenen Datensatz über

Empfehlungen für die Erfassung

die Skulpturengruppe lautet in diesem Fall „hat größeren konzeptuellen Kontext“ (<http://terminology.lido-schema.org/lido00575>, siehe auch: **Art des Bezugs**). Die Tatsache, dass das Werk auf Geldscheinen dargestellt ist, wird ebenfalls nicht als eigenes Ereignis beschrieben, es handelt sich vielmehr um ein neues Werk, das mit Giacomettis „L'Homme qui marche I“ in einer inhaltlichen Beziehung steht und demzufolge in einem eigenen Datensatz beschrieben wird. Auf die Darstellung auf der 100-Franken-Note könnte man mit Hilfe der Relation „ist abgebildet in“ (<http://terminology.lido-schema.org/lido00998>) verweisen. Darüber hinaus sind Verknüpfungen über weitere sogenannte „Related Work Sets“ (<http://lido-schema.org/schema/v1.1/lido-v1.1.html#relatedWorkSet>) zu anderen Abgüssen des „L'Homme qui marche I“, zu weiteren Versionen des „L'Homme qui marche“, zur geistigen Schöpfung des später verworfenen „Chase-Manhattan-Plaza-Projects“ oder auch zu weiteren schriftlichen Quellen denkbar, die den Kontext der Werk- oder Exemplar-Entstehung näher beschreiben. Alle diese Verweise sind äußerst nützlich für die kunsthistorische Recherche.

Abb. 4 Das Werk und die Exemplare vom „L'Homme qui marche“ von Alberto Giacometti



Die oben erwähnte Beschreibung des Exemplars mit Hilfe verschiedener Ereignisse in der Objektgeschichte führt zu einer deutlichen Präzisierung und Anreicherung der Informationen. Während in einem üblichen analogen Katalogeintrag stets nur das Jahr 1960 als Jahr der Herstellung des „L'Homme qui marche I“ genannt wird, lassen sich mindestens sechs genau datierbare Ereignisse mit dem Objekt verknüpfen, die jeweils an verschiedenen Orten und unter Beteiligung verschiedener Personen stattgefunden haben. Alberto Giacometti gehört zu den am häufigsten gefälschten Künstlern. Je präziser seine Werke beschrieben werden, desto besser ist die Datenlage, aufgrund derer man sein Oeuvre rekonstruieren und Fälschungen erkennen könnte. Die genaue Beschreibung

der einzelnen Ereignisse trägt dazu bei, in einem übergeordneten Portal die Informationen aus verschiedenen Sammlungen sachbezogen miteinander verknüpfbar zu machen. Jede Sammlung tendiert dazu, nur die eigenen Objekte zu betrachten und demzufolge auf das einzelne Exemplar bezogene Informationen bevorzugt zu erfassen. Dabei wird vernachlässigt, dass es in anderen Sammlungen weitere Exemplare desselben Werks geben könnte. In einem Portal sollte man hingegen alle Abgüsse in einem Suchergebnis zusammenfassen können, die auf dasselbe Modell oder dieselbe Idee zurückgehen, um beispielsweise die Patinierung der unterschiedlichen Exemplare zu vergleichen oder um festzustellen, ob die angegebene Auflagenhöhe mit der Anzahl tatsächlich vorhandener Exemplare übereinstimmt. Das gelingt aber nur, wenn die beschreibenden Metadaten die gleiche formale und inhaltliche Struktur aufweisen und inhaltliche Beziehungen zu verwandten Objekten in den Daten bereits angelegt sind.

Um ein einzelnes Exemplar digital in seinen Entstehungskontext einzuordnen, hat man im LIDO-Schema also mehrere Möglichkeiten:

- ▶ die Beschreibung der verschiedenen Ereignisse, die zur Entstehung des Objekts geführt haben, im Rahmen einzelner „Event Sets“ (<https://lido-schema.org/schema/v1.1/lido-v1.1.html#eventSet>) – inklusive vorbereitender Tätigkeiten, die zu heute nicht mehr physisch vorhandenen Vorstufen geführt haben.
- ▶ die Verlinkung auf Datensätze, die andere physische Objekte beschreiben, die in einem engen Zusammenhang mit dem beschriebenen Objekt stehen, z. B. der Verweis auf ein konkretes Modell, das einem Abguss zugrunde lag.
- ▶ die Verknüpfung mit einem Werknormdatensatz, der auf einer konzeptuellen Ebene einen größeren Zusammenhang beschreibt; Verschiedene Arten von Relationen können hier angewendet werden, unter anderem:
 - ▶ hat größeren konzeptuellen Kontext (<http://terminology.lido-schema.org/lido00575>) oder
 - ▶ ist Exemplar von (<http://terminology.lido-schema.org/lido00627>).
- ▶ der Hinweis auf eine ausführliche schriftliche Quelle, die den Kontext näher erläutert, idealerweise in Form einer dauerhaft digital zugänglichen Quelle.
- ▶ die Erläuterung des Kontextes im Rahmen der freitextlichen Objektbeschreibung. Diese ist zwar für den Menschen gut lesbar, kann jedoch schlecht maschinell ausgewertet werden.

Beziehungen zu anderen Werken und Objekten

Oft möchte man den inhaltlichen Zusammenhang zwischen einem Objekt und anderen Objekten, Werken oder Werkteilen in der eigenen oder in fremden Sammlungen deutlich machen, wie z. B. den Bezug einer Kopie zu einem Original, einer druckgrafischen Reproduktion zu einem Gemälde oder den Zusammenhang zwischen der Vorder- und Rückseite eines Objekts. Oder man möchte zeigen, dass ein Objekt Teil einer Objektgruppe ist, die als zusammengehörendes Werk konzipiert wurde. Man denke etwa an die einzelnen Teile eines mehrflügeligen Altarretabels, an Gemälde, die als Pendants geschaffen wurden, oder an eine Skulpturengruppe. Zudem sollte es möglich sein, auf historische Zusammenhänge zu anderen Werken oder Werkteilen aufmerksam zu machen (war früher Teil von, <http://terminology.lido-schema.org/lido01085>). Typische Beispiele für direkte Objektbeziehungen gibt es auch für Objekte, die jeweils einzelne Etappen im Entstehungsprozess einer finalen Realisierung einer Werkidee markieren, wie Modelle, Studien oder Entwürfe. Mittels durchdachter Verknüpfungen von Datensätzen können Informationen zugänglich gemacht werden, die allein durch Suchanfragen nicht offensichtlich werden. Damit wird ein echter wissenschaftlicher Mehrwert erzeugt, indem die Nutzer:innen auf verwandte Werke/Objekte aufmerksam gemacht werden.

Beim Anlegen der Relationen gilt es, folgende Grundregeln zu berücksichtigen:

- ▶ Das Objekt, auf das verwiesen wird, muss stets eindeutig identifiziert werden, idealerweise mit einem URI oder mindestens mit einem eindeutigen numerischen oder alphanumerischen Code (z. B. <https://doi.org/10.16903/ethz-grs-1959.0020> im Datenfeld „**Identifikation des Bezugswerks/-objekts**“).
- ▶ Die Art von Bezügen sollte unbedingt mit einer festen Terminologie benannt werden, idealerweise in Form eines URI aus der LIDO-Terminologie (z. B. ist Kopie nach; <http://terminology.lido-schema.org/lido00262>).
- ▶ Beziehungen zwischen zwei Objekten sind nur dann sicher geknüpft, wenn sie auf beiden Seiten verankert werden. Sofern das Objekt, auf das verwiesen wird, ebenfalls in der eigenen Datenbank beschrieben wird, sollten Bezüge zwischen Datensätzen daher reziprok angegeben werden. In manchen Erfassungssystemen kann dies automatisch beim Datenexport erfolgen.
- ▶ Wenn möglich, sollte man die Quelle nennen, auf der die Aussage über die Art des inhaltlichen Bezugs beruht, da es auch durchaus unterschiedliche Aussagen über diesen Sachverhalt geben kann. Z. B. sagt Forscher A, das Werk sei eine Vorzeichnung zum Gemälde und Forscher B vertritt die Auffassung, es handele sich um eine Nachzeichnung. In einem solchen Fall würden zwei „Related Work Sets“ (<https://lido-schema.org/schema/v1.1/lido-v1.1.html#relatedWorkSet>)

Beziehungen zu anderen Werken und Objekten

mit jeweils unterschiedlichen Arten von Bezügen und verschiedenen Quellen der Aussage angelegt werden. Die Quellen werden in einem Datenfeld abgelegt, das dem LIDO-Element „**Quelle für die Aussage über den Bezug**“ entspricht.

Die häufigste Relation ist die Teil-Ganzes-Beziehung, die bei mehrteiligen Werken zur Anwendung kommt. Hierbei handelt es sich um direkte, intrinsische Beziehungen zwischen Objektteilen, die für das Verständnis des Gesamtwerks unabdingbar sind und daher zwingend angegeben werden müssen. Für ein solches Werk, das möglichst vollständig in der eigenen Sammlung vorhanden ist, wird ein Gruppendatensatz angelegt, der das Gesamtwerk beschreibt. Der **Gruppendatensatz zum Gesamtwerk** im lokalen Sammlungsmanagementsystem sollte dann in einzelnen Informationsgruppen auf die jeweiligen Teile verweisen. Diese Informationsgruppen sollten möglichst aus den folgenden Elementen bestehen:

- ▶ Titel und ggf. Kurzbeschreibung des Einzelteils, z.B. „Katharina von Alexandrien, Statuette, Lindenholz, geschnitzt, polychrom gefasst“
- ▶ URL (= Uniform Resource Locator) -Verweis auf eine Beschreibung des Bezugsobjekts oder -werks, normalerweise eine Ressource im WWW (World Wide Web)
- ▶ ID (identity document) des Einzelteils (Im Erfassungssystem ist das die Nummer des Datensatzes des jeweiligen Teils; im Export ist idealerweise das LIDO-Element „**Publizierte ID für das Objekt**“ des Einzelteils anzugeben.)
- ▶ Art des Bezugs, hier: „hat physischen Teil“ (<http://terminology.lido-schema.org/lido00256>). Achtung: Der Bezug wird ausgehend vom im Datensatz beschriebenen Objekt (z.B. gesamter Altar) in Richtung des Bezugsobjekts (z.B. Element des Altars) angegeben.
- ▶ ggf. Quelle der Aussage über diesen Bezug

Zudem würde man für jeden **Einzelteil** je einen eigenen Datensatz anlegen. Hier hat man die Möglichkeit, die nur auf den Teil zutreffenden Informationen genau anzugeben, z. B. Bildthemen, Maße, Materialangaben. Diese Datensätze sollten jeweils in Bezug auf das Gesamtwerk die folgenden Angaben enthalten:

- ▶ Titel und ggf. Kurzbeschreibung des Gesamtwerks, z.B. „Meister des Molschlebener Retabels (zugeschrieben), Lüderbacher Beweinungsaltar, 1500/1520“
- ▶ URL-Verweis auf eine Beschreibung des Bezugsobjekts oder -werks, normalerweise eine Ressource im WWW
- ▶ ID des Gesamtwerks (im Erfassungssystem ist das die Nummer des Datensatzes des Gesamtwerks; im Export ist idealerweise der Dateninhalt des LIDO-Elements „**Publizierte ID für das Objekt**“ anzugeben.)
- ▶ Art des Bezugs, hier: „ist physischer Teil von“ (<http://terminology.lido-schema.org/lido00255>)
- ▶ ggf. Quelle der Aussage über diesen Bezug

Empfehlungen für die Erfassung

Auf der Präsentationsebene einer Online-Datenbank oder eines sammlungsübergreifenden Portals könnte dieser Werksbezug dann in Form eines hierarchisch gegliederten Datensatzes angezeigt werden, um die Werkstruktur anschaulich zu machen. Der Datensatztyp für den Datensatz, der das Gesamtwerk beschreibt, lautet „Objektgruppe (Katalogisierungsebene)“ (<http://terminology.lido-schema.org/lido00453>). Der Datensatztyp für das einzelne Objekt, das Teil des Gesamtwerks ist, lautet „Bestandteil (Katalogisierungsebene)“ (<http://terminology.lido-schema.org/lido00447>).

Der Verweis auf die Werkverzeichnisnummer oder den Werknormdatensatz für das im Datensatz beschriebene Objekt oder Werk wird im Rahmen des Datenexports als eigenes „Related Work Set“ angelegt und mit der Art des Bezugs „ist Exemplar von“ (<http://terminology.lido-schema.org/lido00627>) gekennzeichnet. Das kann in Portalen anschließend dafür genutzt werden, alle Exemplare eines Werks mit derselben Werkverzeichnisnummer oder demselben Werknormdatensatz in einem Suchergebnis zusammenzuführen. Siehe auch: **Werkverzeichnis + Nr. oder Werknormdatensatz-ID**. Über die oben genannten kontrollierten Werkbezüge hinaus kann man jegliche inhaltliche Bezüge zwischen Objekten angeben, unabhängig davon, ob sie sich in der eigenen oder in externen Sammlungen befinden. Zu allen weiteren definierten Arten von Relationen siehe die LIDO-Terminologie (<http://terminology.lido-schema.org/relatedWorkRelType>).

Folgende indirekte Beziehungen sollten nicht als direkte Verlinkungen zwischen Datensätzen angelegt werden, da diese Bezüge über entsprechende Suchanfragen in einem Portal oder Online-Katalog sichtbar gemacht werden können:

- ▶ Andere Werke desselben Künstlers
- ▶ Andere Werke mit demselben Thema

Ebenso sollten extrinsische Beziehungen zwischen Werken, wie z. B. eine lediglich stilistische Bezugnahme auf ein Vorbild, nicht in Verknüpfungen von Datensätzen abgebildet werden, sondern durch eine freitextliche Beschreibung des Zusammenhangs (siehe auch: https://www.getty.edu/research/publications/electronic_publications/cdwa/21related.html). Wenn es keinen internen oder externen Datensatz gibt, der das verwandte Werk oder Objekt beschreibt, kann nur in einer freitextlichen Beschreibung ein Hinweis auf eine Verwandtschaft zu einem anderen Objekt oder Werk gegeben werden. Dabei sind möglichst genaue und umfassende Informationen über das Bezugsobjekt anzugeben, wie Titel, Objektbezeichnung, Name der Herstellerin oder des Herstellers, Datierung, aktueller Standort und Inventarnummer.

Kontrollierte Vokabulare

Der wissenschaftliche Mehrwert eines Portals im Vergleich zu einer haus-eigenen Online-Datenbank besteht darin, auf mehrere Sammlungen ver-teilte Informationen zusammenzuführen, sie miteinander zu verknüpfen und dadurch neue Aussagen und Erkenntnisse zu gewinnen. Das gelingt umso besser, je mehr Verknüpfungspunkte es gibt. Die Voraussetzung für solche inhaltlichen Verknüpfungen von Daten bilden gemeinsam genutzte kontrollierte Vokabulare. Einige, wie der in diesem Handbuch besonders empfohlene AAT des Getty Research Institute, leisten eine eindeutige Begriffsdefinition ihrer Gegenstände, weisen dem Begriff Be-zeichnungen in verschiedenen Sprachen und Schreibweisen zu und ordnen ihn häufig in eine Hierarchie ein. Neben einer Vorzugsbezeichnung enthalten kontrollierte Vokabulare synonyme Bezeichnungen (= verschie-dene Bezeichnungen für den gleichen Begriff) und leisten die Abgrenzung von Homonymen (= gleiche Bezeichnungen für verschiedene Begriffe).

Wenn eine Sammlung ihre Bestände mit der Bezeichnung „Porträt“ er-schließt, während andere Sammlungen mit den Bezeichnungen „Bildnis“ oder „Portrait“ arbeiten, dann würde es nicht gelingen, alle entsprechend bezeichneten Objekte in einem Suchergebnis innerhalb eines sammlungs-übergreifenden Portals zusammen auffindbar zu machen. Das heißt, Nut-zer:innen würden die Datensätze nicht finden, wenn sie mit einer anderen Bezeichnung suchen, als sie die Erfasser:innen verwendet haben. Kontrol-lierte Vokabulare halten häufig bereits verschiedene Bezeichnungen für ei-nen Begriff vor, oft auch in mehreren Sprachen. Deshalb sollten bereits in dem Sammlungsmanagement-System, zumindest aber beim Datenexport, Referenzen auf sie hergestellt werden. Dafür müssen die persistenten Identifizier für den Begriff verwendet werden, die in Form eines URI in einem pub-lizierten Vokabular vorliegen. Wird in den eigenen Daten der entsprechende URI eingetragen, z. B. <http://vocab.getty.edu/aat/300015637>, so kann man in einer Portalanwendung die im AAT-Datensatz hinterlegten synonymen und auch anderssprachigen Bezeichnungen integrieren. Das bedeutet, Er-fasser:innen können im lokalen Kontext eine der angebotenen synonymen Bezeichnungen für den Begriff bevorzugen, aber sie sollten, wann immer möglich, durch das Ergänzen des URI des passenden Normdatensatzes da-für sorgen, dass dessen Inhalte in den einzelnen Datensätzen nachträglich angereichert werden können. Im genannten Fall kann der externe Portal-betreiber durch die Integration der kontrollierten Vokabulare nachträglich weitere Suchbegriffe als sogenannte „added search terms“ ergänzen. Da-durch können auch Datensätze mit dem Dateninhalt „Bildnis“ gefunden werden, auch wenn der Nutzer oder die Nutzerin nach „Porträt“ gesucht hat. Das erhöht die Genauigkeit und Trefferquote der Suchergebnisse.

Ein grundlegendes Postulat des Linked-Data-Konzepts ist, dass nicht Webseiten, sondern Informationen über Entitäten der realen Welt – be-stimmte Personen, Körperschaften, Orte, Ereignisse, Objekte, Begriffe

Empfehlungen für die Erfassung

etc. – verlinkt werden. Der URI fungiert in diesem Rahmen als eindeutiger, persistenter Identifikator für die jeweilige Entität, der ihre maschinelle Identifizierung erlaubt. Synonyme und Bezeichnungsvarianten, wie unterschiedliche Namensformen oder Sprachäquivalente, stellen kein Hindernis für die Identifikation der Entitäten und den gewünschten Vernetzungsprozess mehr dar. Dieses Handbuch empfiehlt daher Vokabulare, die nicht nur eine gute fachliche Abdeckung bieten, sondern URI aufweisen, weit verbreitet sind, wegen ihrer offenen Lizenzen einfach nachgenutzt werden können, über Schnittstellen verfügbar sind und von vertrauenswürdigen Gedächtnisinstitutionen gepflegt werden. Nutzt man URI der in diesem Handbuch empfohlenen Normdateien, Thesauri und Klassifikationen, wird die webbasierte Vernetzung von verteilten Informationen zu den jeweiligen Entitäten wesentlich erleichtert, oftmals überhaupt erst ermöglicht und die Voraussetzung für eine Nutzung in Linked Data- und Semantic Web-Applikationen geschaffen.

Ein weiterer Vorteil liegt in der eindeutigen Identifizierung der referenzierten Personen, Orte, Objekte oder Begriffe: Auch wenn unterschiedliche Bezeichnungen und Namensvarianten benutzt werden, kann mit Hilfe eines URI eine qualifizierte Aussage darüber getroffen werden, ob es sich z. B. um Lucas Cranach den Älteren oder den Jüngeren handelt. Dies ermöglicht entsprechend differenzierte Datenbankabfragen, etwa nach den Werken von Lucas Cranach dem Älteren (und nicht dem Jüngeren). Grundsätzlich gilt: Je mehr Referenzen auf kontrollierte Vokabulare ein Objektdatensatz enthält, desto mehr Bezeichnungen stehen für den Sucheinstieg zur Verfügung und desto mehr Verknüpfungen zu verwandten Inhalten können generiert werden, wenn die Daten für die Nachnutzung entsprechend aufbereitet werden.

Folgende kontrollierte Vokabulare werden empfohlen:

Art and Architecture Thesaurus® (AAT)

<https://www.getty.edu/research/tools/vocabularies/aat/>

Der AAT ist ein mehrsprachiger, hierarchisch strukturierter Thesaurus mit Allgemeinbegriffen zu Kunst, Architektur und kulturellem Erbe des Getty Research Institute. Er enthält ein umfassendes Fachvokabular zu Werk- und Objektarten, Rollen, Materialien, Stilen, Kulturen und Techniken, die durchweg mit Erläuterungen versehen sind. Eigennamen und ikonografische Begriffe sind nicht enthalten, sie werden durch andere Getty-Vokabulare abgedeckt. Der Auf- und Ausbau des AAT erfolgt konform zu ISO 25964, dem internationalen Standard für Thesauri. Er genießt ein hohes Renommee und wird weltweit in Museen, Bibliotheken und Archiven eingesetzt. Die mit einer Open Data Commons Attribution License (ODC-By) 1.0 versehenen Daten stehen in verschiedenen Formaten als Linked-Open-Data-Publikation über einen SPARQL (SPARQL Protocol And RDF Query Language) Endpoint oder als Datenabzug zur Verfügung.

Die Übertragung des AAT ins Deutsche ist noch unvollständig. Es bietet etwa 8.500 (Stand 2020) deutschsprachige Begriffe zu Werk- und Objektarten an, die auch Getty-Angebot enthalten sind: <http://www.aat-deutsch.de>.

Der AAT wird besonders empfohlen für:

- ▶ Objekttypen (z. B. <http://vocab.getty.edu/aat/300047838> = Bozzetto [bozzetti])
- ▶ Titeltypen (z. B. <http://vocab.getty.edu/aat/300417200> = Sammlungstitel [repository titles])
- ▶ Arten von schriftlichen Bezeichnungen am Objekt (z. B. <http://vocab.getty.edu/aat/300028705> = Signatur)
- ▶ Arten von Objektbeschreibungen oder Ereignisbeschreibungen (z. B. <http://vocab.getty.edu/aat/300435427> = kritischer Kommentar)
- ▶ Rollen von beteiligten Personen oder Organisationen (z. B. <http://vocab.getty.edu/aat/300400903> = Auftraggeberin oder Auftraggeber)
- ▶ Arten von Zuschreibungen (z. B. <http://vocab.getty.edu/aat/300404269> = zugeschrieben an)
- ▶ geografisch/stilistische Einordnungen (z. B. <http://vocab.getty.edu/aat/300111184> = Flämisch)
- ▶ Epochen (z. B. <http://vocab.getty.edu/aat/300021149> = Frühbarock)
- ▶ Materialien (z. B. <http://vocab.getty.edu/aat/300015050> = Ölfarbe)
- ▶ Techniken (z. B. <http://vocab.getty.edu/aat/300054216> = Malen (künstlerisches Verfahren)/gemalt)
- ▶ Werkzeuge (z. B. <http://vocab.getty.edu/aat/300022552> = Pinsel)
- ▶ Objekte nach Art des Bildthemas (z. B. <http://vocab.getty.edu/aat/300117546> = Seestück)

Bitte beachten, dass dieselbe Bezeichnung im AAT in unterschiedlichen Facetten vorkommen kann. Ein Ölgemälde entspricht z. B. dem Begriff „oil paintings (visual works)“ (<http://vocab.getty.edu/aat/300033799>) in der Objektfacette des AAT, weil es sich um eine Bezeichnung für ein Objekt handelt. Die Technik Ölmalerei, die zur Herstellung eines Ölgemäldes dient, entspricht hingegen dem Begriff „oil painting (technique)“ (<http://vocab.getty.edu/aat/300178684>) aus der Facette der Prozesse und Techniken. Der erstgenannte Begriff wird im Datenfeld „Objekttyp“ verwendet, der zweite im Feld „Technik“. Eine Verwechslung der ähnlich lautenden Bezeichnungen ist leicht möglich. Doch durch einen Zusatz (visual works oder technique) wird im AAT auf diesen Unterschied hingewiesen. In jedem Fall sollte durch einen Blick auf die Einordnung des entsprechenden Begriffs in der AAT-Hierarchie sichergestellt werden, dass man den gewünschten Deskriptor aus der richtigen Facette verwendet.

Empfehlungen für die Erfassung

Folgende Hierarchien sind dabei voneinander zu unterscheiden:

- ▶ **für Objekttypen:** Deskriptoren aus der Objects Facet
<http://vocab.getty.edu/hier/aat/300264092>
- ▶ **für Material:** Deskriptoren aus der Materials Facet
<http://vocab.getty.edu/hier/aat/300010357>
- ▶ **für Technik:** Deskriptoren aus der Processes and Techniques hierarchy <http://vocab.getty.edu/hier/aat/300053001>
- ▶ **für Werkzeuge:** Deskriptoren aus der Tools and Equipment hierarchy
<http://vocab.getty.edu/hier/aat/300022238>

Gemeinsame Normdatei (GND)

Die GND bietet den größten domänenübergreifenden Fundus an deutschsprachigen Normdaten zu Individualbegriffen wie Personen, Körperschaften, Veranstaltungen, Geografika und Werken der Literatur, Musik und bildenden Kunst, aber auch zu Allgemeinbegriffen in Form von Sachschlagwörtern an. Für den Ausbau und die Pflege nach definierten Qualitätskriterien ist die GND-Kooperative unter der Leitung der Deutschen Nationalbibliothek verantwortlich. Über die angestammten bibliothekarischen Nutzungskontexte hinaus wird die GND zunehmend von Archiven, Museen, Projekten und in Web-Anwendungen genutzt. Die GND-Kooperative öffnet sich zunehmend für die aktive Beteiligung auch nichtbibliothekarischer Communitys. Sie steht unter einer Creative Commons (CC) 1.0 Lizenz (<https://creativecommons.org/publicdomain/zero/1.0/>) über mehrere Bezugswege bereit, darunter auch verschiedene Schnittstellen: z. B. <https://portal.dnb.de> oder https://gnd.network/Webs/gnd/DE/Home/home_node.html.

Einen Sucheinstieg in die GND bieten auch die Online-GND (OGND) des Bibliotheksservicezentrums Baden-Württemberg (<https://www.bs-zbw.de/ognd.html>) und die lobid-gnd des Hochschulbibliotheksentrums des Landes Nordrhein-Westfalen (<https://lobid.org/gnd>).

Die GND wird empfohlen für:

- ▶ Objekttypen (z. B. <https://d-nb.info/gnd/4172442-2> = Ölskizze)
- ▶ Personen und Organisationen, die mit Ereignissen rund um das beschriebene Objekt in Zusammenhang stehen (Auftraggeber, Hersteller, ausstellende Institutionen etc.) oder die auf dem Objekt dargestellt sind (z. B. <https://d-nb.info/gnd/118594206> = Pablo Picasso)
- ▶ Materialien (z. B. <https://d-nb.info/gnd/4389628-5> = Eichenholz)
- ▶ Techniken (z. B. <https://d-nb.info/gnd/4071463-9> = Freskomalerei)
- ▶ Ortsangaben (z. B. <https://d-nb.info/gnd/4269332-9> = Perleberg)
- ▶ Werknormdaten von Kunst- und Bauwerken: Sowohl für das im Datensatz beschriebene Objekt selbst (sofern ein Werknormdatendatz dafür bereits vorhanden ist) als auch für Werke, zu denen ein inhaltlicher Bezug besteht (z. B. <https://d-nb.info/gnd/4815969-4> = Jacques Louis David, Krönung Napoleons I., Paris, Louvre)

- ▶ dargestellte Ereignisse, die in der Realität stattgefunden haben (z. B. <https://d-nb.info/gnd/4189312-8> = Schlacht bei Waterloo)

Auch bei der Verwendung der GND sollte man darauf achten, den gewünschten Datensatz mit dem richtigen Entitätstyp zu wählen. Besonders übersichtlich ist dies in der Facettennavigation von <https://lobid.org/gnd/> gekennzeichnet.

- ▶ **für Objekttypen, Materialien, Techniken, Werkzeuge:** Deskriptoren aus Datensätzen mit dem Entitätstyp Sachbegriffe – Schlagwort sensu stricto
- ▶ **für Personen:** Deskriptoren aus Datensätzen mit dem Entitätstyp Person
- ▶ **für Institutionen, Organisationen:** Deskriptoren aus Datensätzen mit dem Entitätstyp Organisation/Körperschaft
- ▶ **Ortsangaben:** Deskriptoren aus Datensätzen mit dem Entitätstyp Geografika
- ▶ **für geistige Schöpfungen aller Gattungen auf Werkebene:** Deskriptoren aus Datensätzen mit dem Entitätstyp Werke
- ▶ **für Bauwerke:** Deskriptoren aus Datensätzen mit dem Entitätstyp Geografika – Bauwerk oder Denkmal

Iconclass

<http://iconclass.org>

Iconclass ist ein umfassendes Klassifizierungssystem für die Erschließung und das Auffinden von in Kunstwerken dargestellten Themen, das weltweit in Museen und Bibliotheken angewendet wird. Es handelt sich um ein monohierarchisch geordnetes System von etwa 28.000 Definitionen zu Begriffen, Motiven, Objekten, Personen, Ereignissen und abstrakten Ideen, denen jeweils ein Klassifizierungscode mit Erläuterungen des ikonografischen Motivs und Schlagworten in mehreren Sprachen zugeordnet sind. Iconclass eignet sich besonders für die Erschließung abendländischer Kunst, dabei insbesondere für die Themen der christlichen Ikonografie und der antiken Mythologie. Es wird von der Henri van de Waal Foundation getragen, ist als Linked Open Data unter einer Open Data Commons Open Database License (ODbL) v1.0 verfügbar und vollständig ins Deutsche übersetzt.

- ▶ für ikonografische Themen und Motive (z. B. <http://iconclass.org/71A423>: Versuchung und Sündenfall (Genesis 3:1-7): Adam und Eva halten die Frucht in Händen und schicken sich an, sie zu essen)

Empfehlungen für die Erfassung

Union List of Artist Names® (ULAN)

<https://www.getty.edu/research/tools/vocabularies/ulan>

Die Union List of Artist Names ist ein Vokabular zu über 300.000 Künstlerinnen, Künstlern und anderen mit bildender Kunst und Architektur in Beziehung stehenden Personen und Körperschaften. Es wird vom Getty Research Institute unter einer freien ODC-By-1.0-Lizenz herausgegeben und ist als Linked Open Data einsetzbar.

- ▶ für Personen und Organisationen, die mit Ereignissen rund um das beschriebene Objekt in Zusammenhang stehen oder die auf dem Objekt dargestellt sind (z. B. <http://vocab.getty.edu/ulan/500002921> = Peter Paul Rubens)

Thesaurus of Geographic Names® (TGN)

<https://www.getty.edu/research/tools/vocabularies/tgn/>

Der Thesaurus of Geographic Names, ebenfalls vom Getty Research Institute, verzeichnet georeferenzierte Einträge zu ca. 912.000 Geografika weltweit, darunter auch historische Orte. Verfügbar unter ODC-By-1.0-Lizenz als Linked Open Data.

- ▶ für Ortsangaben (z. B. <http://vocab.getty.edu/tgn/7004333> = München)

GeoNames

<http://www.geonames.org>

Geonames ist eine vor allem aus öffentlichen Datenquellen aggregierte geografische Datenbank mit über 12 Millionen georeferenzierten topografischen Entitäten weltweit, darunter auch zunehmend ortsfeste Artefakte wie Bauwerke und Denkmäler. Die Daten stehen unter einer Creative Commons Attribution 4.0 Lizenz (<https://creativecommons.org/licenses/by/4.0/>) zum Download oder über Schnittstellen für LOD (Linked Open Data) -Anwendungen bereit. Es ist bereits einer der wichtigsten Knotenpunkte im Semantic Web.

- ▶ für Ortsangaben und Bauwerke (z. B. <https://sws.geonames.org/2803138> = Antwerpen)

Cultural Objects Name Authority® (CONA) und Iconography Authority® (IA)

<https://www.getty.edu/research/tools/vocabularies/cona>

Diese beiden Normdateien des Getty Research Institute sind im Aufbau befindlich. CONA enthält als Werknormdatei gegenwärtig etwa 10.000 Datensätze zu visuellen Werken und Architektur. Die IA dient der Erschließung und Referenzierung von Darstellungsinhalten und Bildthemen. Sie verzeichnet Eigennamen und andere Informationen zu benannten Ereignissen, Themen und Erzählungen aus Religion und Mythologie, Literatur und darstellender Kunst sowie zu legendären und fiktiven Figuren und

Orten und strebt dabei auch die Abdeckung der Ikonografie nichtwestlicher Kulturkreise an. Beide stehen unter der Open Data Commons Attribution License (ODC-By) 1.0 und sind über einen Webservice verfügbar. Eine Bereitstellung über SPARQL steht noch aus.

- ▶ CONA für Kunst- und Bauwerke, sowohl für dargestellte sowie für Werke, zu denen ein inhaltlicher Bezug besteht (z. B. <http://vocab.getty.edu/page/cona/700008550> = Auguste Rodin, Höllentor)
- ▶ IA für ikonografische Themen und Motive (z. B. <http://vocab.getty.edu/page/ia/901001323> Legend of Adam and Eve (Abrahamic religion narrative) – Fall of Man (Abrahamic religion narrative))

Wikidata

<https://www.wikidata.org>

Wikidata ist eine seit 2012 bestehende, schnell wachsende Datenbank zu allen Wissensdomänen, die von jeder und jedem Interessierten editiert und durch automatisierte Anreicherung ergänzt werden kann. Für viele der hier relevanten Kontexte, insbesondere bei den Individualbegriffen zu Personen, Körperschaften, Orten und Objekten, liegt bereits eine sehr gute Abdeckung vor. Die redaktionelle Qualität ist streckenweise uneinheitlich, sie verändert sich oft mit großer Dynamik. Der große Wert von Wikidata liegt in seinem beständig wachsenden Netz von Ko-Referenzen der URI seiner Datenobjekte auf die zahlreicher anderer Linked-(Open)-Data-Angebote, so dass Wikidata zu einer der wichtigsten Datendrehscheiben des semantischen Netzes geworden ist. Wikidata wird von der Wikimedia Foundation getragen, die Strukturdaten sind unter einer CC0 1.0 Lizenz über einen SPARQL Endpoint adressierbar. Wegen zahlreicher weiterer URI-gestützter normdatenbasierter Verknüpfungen eignet sich Wikidata auch als guter Einstieg für die Suche nach geeigneten Identifikatoren für Personen, Organisationen, Orte und Ereignisse in den anderen empfohlenen Vokabularen.

- ▶ für Objekttypen (z. B. <http://www.wikidata.org/wiki/Q56676227> = Ölgemälde)
- ▶ für Arten von Maßangaben (Höhe, Breite, cm oder mm, z. B. <http://www.wikidata.org/wiki/Q208826> = Höhe)
- ▶ für Personen und Organisationen, die mit Ereignissen rund um das beschriebene Objekt in Zusammenhang stehen (z. B. <https://www.wikidata.org/wiki/Q5669> = Sandro Botticelli)
- ▶ für Rollen von beteiligten Personen oder Organisationen (z. B. <https://www.wikidata.org/wiki/Q3109381> = Gemmenschneider)
- ▶ für Ortsangaben (z. B. <http://www.wikidata.org/wiki/Q183> = Deutschland)
- ▶ für dargestellte Ereignisse und andere Bildinhalte (z. B. <https://www.wikidata.org/wiki/Q48314> = Schlacht bei Waterloo)
- ▶ für Arten von Rechten (z. B. <http://www.wikidata.org/wiki/Q1206081> = Urheberrecht)

Virtual International Authority File (VIAF)

Gleiches gilt für das VIAF (<https://viaf.org>) des Online Computer Library Center, Inc. (OCLC), ein aggregierter Normdatendienst mit Beständen von aktuell 60 Nationalbibliotheken und ähnlichen Anbietern. VIAF stellt zu einer Entität Konkordanzen über die zu ihr vorhandenen Normdatensätze her und bietet einen URI zu diesem Cluster an. Auf diese Weise kann die Entität unter allen vorliegenden Sprach- und Namensvarianten und den verbundenen Links der Quellsysteme adressiert werden. VIAF steht unter einer Open Data Commons Attribution License v1.0 und bietet Zugang über eine Schnittstelle und verschiedene Download-Formate.

Integration von kontrolliertem Vokabular in LIDO-Indezelemente

Wann immer möglich, wird ein Deskriptor aus einem der empfohlenen Vokabulare zur Indexierung gewählt. Aufgenommen wird nur der URI des gewünschten Indexterms, zum Beispiel der URI „<http://vocab.getty.edu/aat/300177433>“ aus dem AAT für den Objekttyp eines Freskogemäldes. Während in LIDO 1.0 dafür das Element „Concept Identifier“ genutzt wurde, erfolgt die Einbindung der URI von in den Referenzvokabularen verfügbaren Allgemeinbegriffen in LIDO v1.1 im Element „skos:Concept“. In beiden Fällen sind die URI für eine maschinelle Auswertung der Bezeichnungen und die semantischen Relationen, die sich hinter einem URI verbergen, vorgesehen. Die Angabe der Bezeichnungen oder anderer Eigenschaften eines Deskriptors in einem LIDO-Datensatz ist also nicht erforderlich. Eine mit Elementen und Attributen im SKOS (Simple Knowledge Organization System)-Namensraum ergänzte Darstellung, die bestimmte Bezeichnungen aus dem Referenzvokabular in den LIDO-Datensatz übernimmt, hat aber den Vorteil, dass keine Absprachen mit Portalbetreibern über eine automatische URI-Auswertung getroffen werden müssen. Auf diese Weise können auch Publikationsportale ohne LOD-Technologien die Informationen aus dem über skos:Concept adressierten Datensatz für das Auffinden der Informationen und die Anzeige bereitstellen.

Integration von kontrolliertem Vokabular in LIDO-Indexelemente

Beispiel mit mehreren Bezeichnungen aus dem AAT, die im Element skos:Concept ausgedrückt werden:

```
<lido:objectWorkType lido:type="http://terminology.lido-schema.org/
lido00768">
  <skos:Concept rdf:about="http://vocab.getty.edu/aat/300015637">
    <skos:prefLabel xml:lang="en">portraits</skos:prefLabel>
    <skos:prefLabel xml:lang="de">                                </skos:prefLabel>
    <skos:altLabel xml:lang="en">portrait</skos:altLabel>
    <skos:altLabel xml:lang="de">Bildnis (Porträt)</skos:altLabel>
    <skos:altLabel xml:lang="de">Porträts (visuelle Werke)</skos:altLabel>
  </skos:Concept>
</lido:objectWorkType>
```

Diese Darstellung ist die maschinell gewonnene Auswertung des AAT-Datensatzes.

Auch Zuordnungen (Mappings) zu anderen publizierten, über das Web referenzierbaren Vokabularen können in SKOS eingehen, sofern die Begriffe in beiden Vokabularen das Gleiche (exactMatch) oder annähernd das Gleiche (closeMatch) bedeuten. Die Mapping-Relationen stammen entweder aus dem referenzierten Vokabular selbst oder sie werden durch Auswertung der Konkordanz-Angaben in Wikidata gewonnen (siehe Identifier dort). Mappings werden nicht frei ergänzt. Konkordanzen ermöglichen die Suche in Beständen, die mit unterschiedlichen Vokabularen erschlossen sind. Sie ersetzen nicht die terminologische Kontrolle in jedem einzelnen Vokabular.

Beispiel für das Mapping lokaler Bezeichnungen auf verschiedene im Internet referenzierbare kontrollierte Vokabulare:

```
<skos:Concept rdf:about="http://terminology.lido-schema.org/lido00774">
  <skos:prefLabel xml:lang="en">Male</skos:prefLabel>
  <skos:prefLabel xml:lang="de">Männlich</skos:prefLabel>
  <skos:exactMatch>http://vocab.getty.edu/aat/300189559</
skos:exactMatch>
  <skos:exactMatch>https://d-nb.info/standards/vocab/gnd/gender#male</
skos:exactMatch>
  <skos:exactMatch>http://www.wikidata.org/wiki/Q6581097</
skos:exactMatch>
</skos:Concept>
```

Verwendung von lokalen oder nicht als LOD verfügbaren Vokabularen

Obwohl die Nutzung der oben vorgestellten kontrollierten Vokabulare aufgrund der zahlreichen Vorteile empfohlen wird, besteht in vielen Sammlungen die Anforderung, auch hauseigene, nicht publizierte und nicht über einen URI erreichbare Bezeichnungen in den LIDO-Export einbringen zu wollen. Die lokalen Vokabulare sollten vor der Erstellung der LIDO-Daten sorgfältig auf Konsistenz und Fehlerfreiheit geprüft werden. Dies schließt eine Homonymkontrolle ein: Gleichlautende Bezeichnungen für verschiedene Begriffe bekommen zur Unterscheidung einen qualifizierenden Zusatz in Klammern. Auch Begriffe, die im lokalen Zusammenhang eindeutig erscheinen, können in übergreifenden Suchkontexten mehrdeutig sein: Die häufig verwendete Bezeichnung „Öl“ für Malfarben auf Ölbasis kann sich auf eben jene Ölfarbe, aber in größeren Nutzungskontexten auf Pflanzenöle oder Mineralöle beziehen, weshalb es nicht ausreicht, lediglich das Wort „Öl“ als Bezeichnung für ein Material zu verwenden. Hier wäre das am besten durch die Spezifizierung der Bezeichnung selbst – „Ölfarbe“ – gegeben. Im LIDO-Anzeigeelement kann dennoch die in Bestandskatalogen geläufige Form „Öl“ ausgegeben werden, während im entsprechenden für die Suche auszuwertenden Indexelement der spezifischere Begriff „Ölfarbe“ angegeben werden muss, um präzise Suchergebnisse zu ermöglichen. Bezeichnungen sollten auch unabhängig von dem Datenfeld, in dem sie verwendet werden, eindeutig sein. So ist es notwendig, durch Homonymzusätze die Objektbezeichnung Aquarell (Gemälde) von der Materialangabe Aquarell (Farbe) zu unterscheiden.

Die lokalen Bezeichnungen, die anwendungsbezogen für Anzeige und Index ausgewertet werden sollen, werden im LIDO-Element „Term“ erfasst. Der für Anzeige und Suchindex bevorzugte Term erhält kein Typattribut. Wenn eine Präzisierung gewünscht ist, kann der LIDO-Term „Display (Bezeichnung)“ (<http://terminology.lido-schema.org/lido00526>) für das Attribut „pref“ verwendet werden (siehe auch: <http://lido-schema.org/schema/v1.1/lido-v1.1.html#pref>), um den Term für die Anzeige zu kennzeichnen. Für den Sucheinstieg ergänzte Synonyme, alternative Schreibweisen oder Sprachäquivalente erhalten das Typattribut „lido:addedSearchTerm=“yes“. Es darf in einem Element für ein „Concept“ nur einen „Term“ ohne Typattribut geben.

Verwendung von lokalen oder nicht als LOD verfügbaren Vokabularen

Beispiel für einen Begriff aus einem lokalen Thesaurus mit Oberbegriffen als zugesetzten Suchbegriffen:

```
<lido:materialsTech>
  <lido:termMaterialsTech lido:type="http://terminology.lido-schema.org/
    lido00132">
    <lido:term lido:pref="http://terminology.lido-schema.org/lido00526" xml:
      lang="de">Istrischer Kalkstein</lido:term>
    <lido:term lido:addedSearchTerm="yes" xml:lang="de">Kalkstein</
      lido:term>
    <lido:term lido:addedSearchTerm="yes" xml:lang="de">Stein</lido:term>
  </lido:termMaterialsTech>
</lido:materialsTech>
```

Eine besondere Herausforderung besteht darin, einerseits Referenzen auf publiziertes Vokabular mit SKOS zu nutzen, andererseits aber zusätzlich lokales Vokabular ergänzen zu wollen. Viele Begriffe des AAT sind noch nicht ins Deutsche übersetzt: Hier ist es notwendig, lokale deutschsprachige Bezeichnungen vorzuhalten und diese beim LIDO-Export zu ergänzen, um einen durchgängig deutschsprachigen Datensatz anzeigen zu können. Im vorhergehenden Beispiel zur Integration der im AAT vorgehaltenen Bezeichnungen zum Begriff „Porträt“ fehlt ein deutscher Suchterm, nämlich die Schreibvariante „Portrait“. Bei einer Suchanfrage mit „Portrait“ würden deshalb Objekte, die mit dem AAT-Deskriptor indexiert sind, nicht gefunden. Diese fehlende Bezeichnung kann bei der Erstellung eines LIDO-Datensatzes im Element „Term“ im Anschluss an das Element <skos:Concept> ergänzt werden:

```
<lido:objectWorkType lido:type="http://terminology.lido-schema.org/
  lido00768">
  <skos:Concept rdf:about="http://vocab.getty.edu/aat/300015637">
  (...)
  </skos:Concept>
  <lido:term>Portrait</lido:term>
```

Für eine Ergänzung von URI durch lokale Terme kann es weitere Gründe geben:

- ▶ Das Referenzvokabular bietet noch keine deutsche Übersetzung für den Begriff an.
- ▶ Die lokale Vorzugsbezeichnung ist nicht im Referenzvokabular enthalten.
- ▶ Vielleicht hat man auch eine Reihe von Synonymen für einen Begriff gesammelt, die man zusätzlich für die Suche zur Verfügung stellen möchte.

Empfehlungen für die Erfassung

In all diesen Fällen kann die gewünschte deutsche Bezeichnung, wie oben gezeigt, ergänzt werden, indem das Indexelement um ein oder mehrere Termelemente erweitert wird. Der Term wird idealerweise aus einem lokalen Vokabular mit kontrollierten Bezeichnungen bezogen. Wenn es kein lokales kontrolliertes Vokabular gibt, muss er als freies Schlagwort ergänzt werden. Ergänzt werden hier nur Bezeichnungen, die nicht im Referenzvokabular enthalten sind und die genau oder annähernd die gleiche Bedeutung wie der unter `skos:Concept` referenzierte Begriff haben.

Beispiel mit SKOS und weiteren Bezeichnungen aus lokalem Vokabular im LIDO-Element „Term“:

```
<lido:termMaterialsTech lido:type="http://terminology.lido-schema.org/lido00131">
  <skos:Concept rdf:about="http://vocab.getty.edu/aat/300053357">
    <skos:prefLabel xml:lang="en">fresco painting (technique)</skos:prefLabel>
    <skos:altLabel xml:lang="en">fresco (general technique)</skos:altLabel>
  </skos:Concept>
  <lido:term lido:pref="http://terminology.lido-schema.org/lido00526" xml:lang="de">Freskomalerei (Technik)</lido:term>
  <lido:term lido:addedSearchTerm="yes" xml:lang="de">freskiert</lido:term>
  <lido:term lido:addedSearchTerm="yes" xml:lang="de">Freskotechnik</lido:term>
  <lido:term lido:addedSearchTerm="yes" xml:lang="de">Fresko-Technik</lido:term>
</lido:termMaterialsTech>
```

Zuweilen werden für die Erschließung auch fachlich anerkannte, hochwertige Ressourcen genutzt, die zwar publiziert, aber nicht unter freien Lizenzen verfügbar sind und/oder keine URI aufweisen. Alle gewünschten Informationen zur Entität müssen hier bereits beim LIDO-Export ergänzt werden, da eine maschinelle Anreicherung nicht möglich ist. Ein Beispiel ist das „Allgemeine Künstlerlexikon – Internationale Datenbank online“, das als Subskriptionsprodukt zugänglich ist.

Beispiel für einen Eintrag aus einer publizierten, nicht frei lizenzierten Normdatei, mit zitierbarer URL:

```
<lido:actor lido:type="http://terminology.lido-schema.org/lido00163">
<lido:actorID lido:source="Allgemeines Künstlerlexikon – Internationale
Künstlerdatenbank – Online" lido:type="http://terminology.lido-schema.org/
lido00099">https://www.degruyter.com/database/AKL/entry/_30045282/
html</lido:actorID>
<lido:nameActorSet>
<lido:appellationValue>Bonci, Valerio</lido:appellationValue>
</lido:nameActorSet>
<lido:nationalityActor lido:type="http://terminology.lido-schema.org/
lido01027">
<lido:term xml:lang="de">italienisch (Kultur)</lido:term>
</lido:nationalityActor>
<lido:vitalDatesActor lido:type="http://terminology.lido-schema.org/
lido00819">
<lido:earliestDate lido:type="http://terminology.lido-schema.org/
lido00529">1596</lido:earliestDate>
<lido:latestDate lido:type="http://terminology.lido-schema.org/
lido00529">1611</lido:latestDate>
</lido:vitalDatesActor>
<lido:genderActor lido:type="http://terminology.lido-schema.org/
lido00867">
<skos:Concept rdf:about="http://terminology.lido-schema.org/lido00774">
<skos:prefLabel xml:lang="en">Male</skos:prefLabel>
<skos:prefLabel xml:lang="de">Männlich</skos:prefLabel>
</skos:Concept>
</lido:genderActor>
</lido:actor>
```

Urheber- und Verwertungsrechte

Im LIDO-Schema gibt es vier Bereiche, in denen Angaben zu Rechten gemacht werden können, die mit den Informationen im Datensatz in Zusammenhang stehen. Dies betrifft Rechte an:

- ▶ der ausführlichen Objektbeschreibung (<https://lido-schema.org/schema/v1.1/lido-v1.1.html#objectDescriptionRights>).
- ▶ den sonstigen beschreibenden Metadaten (<https://lido-schema.org/schema/v1.1/lido-v1.1.html#recordRights>).
- ▶ den digitalen Reproduktionen (<https://lido-schema.org/schema/v1.1/lido-v1.1.html#rightsResource>).

Empfehlungen für die Erfassung

- ▶ den Kunstwerken, die Gegenstand des Datensatzes sind (<https://lido-schema.org/schema/v1.1/lido-v1.1.html#rightsWorkSet>).

In jedem Set, das Rechteangaben enthält, gibt es wiederum vier verschiedene LIDO-Elemente, die genutzt werden können: Die Art der Rechte (Rights Type), die Geltungsdauer der Rechte (Rights Date), der Name der Rechteinhaberin oder des Rechteinhabers (Rights Holder) und die Creditline, mit der eine vom Datenlieferanten gewünschte und für die Anzeige optimierte Form der Rechte oder allgemein eine Würdigung angegeben werden kann. Diese Elemente sind für jeden der oben genannten Geltungsbereiche jeweils wiederholbar. Bei den Arten von Rechten kann unter anderem unterschieden werden zwischen Urheberrecht, Leistungsschutzrecht, Persönlichkeitsrecht oder Nutzungsrechten. Das Rights-Type-Element kann zudem im LIDO-v1.1-Format mit Hilfe des Typattributs näher spezifiziert werden. Hierbei sollte auf die LIDO-Terminologie verlinkt und zwischen generischen Arten von Rechten (<http://terminology.lido-schema.org/lido00920>) und spezifischen rechtlichen Informationen zur Nutzung (<http://terminology.lido-schema.org/lido00921>) unterschieden werden. Ein Anwendungsfall dafür kann die namentliche Nennung eines Rechteinhabers oder einer Rechteinhaberin für den *generischen* Rechtstyp „Urheberrecht“ sein. Diese Inhaberin oder dieser Inhaber des Urheberrechts gibt dann für bestimmte Objekte und die sie beschreibenden Daten *spezifische* Nutzungsrechte aus, die in Form einer Lizenz genannt werden.

Allein die Urheber:innen bzw. die sie vertretenden Rechtspersonen und die Daten liefernde Institution entscheiden über den Inhalt der Angaben zu den Urheber- und Verwertungsrechten. Damit eine transparente und rechtssichere Nachnutzung der Daten durch Dritte ermöglicht werden kann, ist eine maschinell auswertbare Angabe der Art der Rechte zu bevorzugen, z. B. in Form der Einbettung eines Links auf die entsprechende Creative-Commons-Lizenz (z. B. <https://creativecommons.org/licenses/by-sa/4.0>).

Praktische Hinweise zur Vorbereitung des LIDO-Exports

Vor der Erstellung eines Datenexports ist es notwendig zu entscheiden, welche Informationen überhaupt für den Export freigegeben werden. Die Namen der für den Export auszuwählenden lokalen Datenfelder werden dann in einer Tabelle den ihnen entsprechenden LIDO-Elementen zugeordnet. Dieser Prozess wird als Mapping bezeichnet.

Durch das Ausfüllen der Mapping-Tabelle für das Anwendungsprofil „Malerei und Skulptur“ (siehe: <https://doi.org/10.11588/data/CHEPS6>) gewinnt die Person, die das Projekt innerhalb der Sammlung betreut, die Kontrolle über die Menge und Art der zu exportierenden Informationen und kann z. B. entscheiden, sensible Daten – wie etwa Ankaufspreise oder Namen von Privatsammler:innen – nicht in den Mapping- und Exportprozess mit aufzunehmen. Je mehr die mit der Erfassung beauftragten Personen an diesem Prozess mitwirken, desto eher können sie dafür Sorge tragen, dass die erschlossenen Informationen vollständig und logisch ausgewertet und an Dritte weitergegeben werden können. Auf der Grundlage der Mapping-Tabelle und des vorliegenden Handbuchs, das für jedes Datenfeld das passende LIDO-Element angibt, können die Datenbankadministrator:innen oder externen Dienstleister:innen leicht erkennen, welches Datenfeld beim Export welchem LIDO-Element zugewiesen werden muss. Die für dieses Anwendungsprofil erstellte Tabelle liefert zudem einen schnellen Überblick über die entsprechend farbig markierten Pflichtelemente oder empfohlenen Elemente.

In einigen Fällen ist es möglich, Informationen aus dem Erfassungssystem an verschiedenen Stellen in LIDO auszugeben. Hier entscheiden die Verantwortlichen der Sammlung entsprechend der Maßgaben ihrer gültigen lokalen Erfassungspraxis. Je stärker die inhaltliche Gliederung der Datenbank mit der technischen Struktur der Daten im LIDO-Schema abgestimmt ist, desto leichter lässt sich später ein solcher Datenexport in diesem Format durchführen.

Bei der Auswahl von Datenfeldern für den Mapping-Prozess sind allgemein zu beachten:

- ▶ der Redaktionsstand der Daten
- ▶ die Sensibilität der Daten
- ▶ die wissenschaftliche, semantische und formale Qualität der Daten
- ▶ möglicherweise unsichere Forschungsstände
- ▶ der rechtliche Status der Werke
- ▶ ethische Gesichtspunkte (z. B. Abbildungen menschlicher Überreste)
- ▶ die durchdachte Nutzung der für die menschenlesbare Anzeige (= Display) und für die maschinelle Suche (= Index) optimierten Elemente in LIDO

Vor einem Datenexport ist es generell angezeigt, den Redaktionsstand der eigenen Daten zu prüfen und die Daten ggf. zu korrigieren oder zu ergänzen.

Typische Fehlerquellen bei der Datenerfassung

Im Folgenden werden häufig beobachtete Fehler aufgelistet, die durch die Befolgung interner Schreibanweisungen vermieden werden können. Diese werden in der Regel von den Koordinator:innen eines Digitalisierungsprojekts erstellt, **bevor** mit der Datenerfassung begonnen wird. Dieses Handbuch soll ein Leitfaden für derartige Schreibanweisungen sein. Die beschreibenden Metadaten zu den Sammlungsobjekten sollten bei allen Datensätzen in einheitlicher Form strukturiert erfasst werden. Je feiner aufgliedert die Informationen sind, desto präziser sind die Suchergebnisse und desto eher ist es möglich, durch die Kombination von Suchfiltern besondere Forschungsfragen zu beantworten. Die inhaltliche und semantische Kontrolle der erstellten Datensätze, die begleitende (Nach-)Schulung der Erfasser:innen, die Aktualisierung von Schreibanweisungen in Anpassung an die Alltagspraxis ist eine nicht zu unterschätzende Daueraufgabe. Dabei gilt: Wichtiger als die Menge an erfassten Daten ist die Konsistenz und Qualität dieser Daten.

Zu den möglichen Fehlern bei der Dateneingabe, die sämtlich zu schlechteren Suchergebnissen führen, zählen:

- ▶ **Rechtschreibfehler**
- ▶ **Es gibt mehr als einen Datensatz für dasselbe Objekt innerhalb einer Sammlung.** Oftmals enthalten die Datensätze auch unterschiedliche Informationen über das Objekt. In einem solchen Fall muss im Rahmen einer vollständigen Dubletten-Kontrolle analysiert werden, welcher der Datensätze den veralteten Kenntnisstand enthält. Dieser wird als obsolet markiert und sollte nicht weiterbearbeitet werden dürfen. Besser führt man die Informationen aus beiden Datensätzen in einem Datensatz zusammen und eliminiert die nun redundante Dublette.
- ▶ **Unvollständige oder fehlende Informationen:** Grundsätzlich sollte angestrebt werden, die Sammlungsobjekte nach dem aktuellen Stand der Forschung so vollständig wie möglich zu beschreiben, auch wenn es begründete Fälle geben kann, bei denen manche für diese Sammlung besonders bedeutsame oder repräsentative Objekte sehr tief erschlossen werden, während zu anderen Objekten kaum Informationen vorliegen. Pflichtfelder müssen immer mit Dateninhalt belegt werden.
- ▶ **Daten werden nicht mit der erwünschten wissenschaftlichen Sorgfalt erschlossen.** Beispielsweise wird das Objekt nicht selbst vermessen, sondern die Maße werden aus einer schriftlichen Quelle übernommen, oder die Technik, mit der ein Objekt hergestellt wurde, wird nicht am Original überprüft, sondern anhand digitaler Abbildungen „erraten“. Auch sollte darauf geachtet werden, dass gleiche Maßeinheiten für gleiche Arten von Objekten verwendet werden, z. B. alle Arbeiten auf Papier werden in Millimeter gemessen, alle Gemälde in Zentimetern, jeweils zuerst die Höhe, dann die Breite.

Typische Fehlerquellen bei der Datenerfassung

- ▶ **Uneinheitliches Niveau hinsichtlich der Präzision von Informationen:** z. B. „um 1700“ oder „1705-03-10“. Während sich die zweite Information sehr leicht maschinell auswerten und beispielsweise in einem Zeitstrahl darstellen ließe, braucht man für die erste Information eine Erläuterung bzw. eine Regel für eine maschinenverwertbare Auflösung beim Datenexport: Welcher Zeitraum ist mit „um“ gemeint? Spätestens bei der Transformation der Daten nach LIDO sollten solche Zeitangaben in explizite numerische Datumswerte in den für die Indexierung vorgesehenen Elementen „earliestDate“ und „latestDate“ extrapoliert werden. Wichtig ist auch, dass Bezeichnungen, mit denen die Unsicherheit der Information ausgedrückt wird, im Rahmen der gesamten Sammlungserschließung konsistent angewendet werden. Wird z. B. im selben Datenpool einmal mit „um“, einmal mit „ca.“ und einmal mit „?“ gearbeitet, lassen sich die Daten nicht mehr maschinell auswerten.
- ▶ **Es fehlen Quellenangaben.** Insbesondere bei interpretativen, nicht objektiv messbaren oder auch unsicheren Aussagen, z. B. bezüglich der Zuschreibung eines Objekts an eine bestimmte Urheberin oder einen bestimmten Urheber, sollten die Quellen der Information angegeben und verschiedene Forschungsmeinungen einander gegenübergestellt werden, siehe auch: [Quelle für die Angaben zur Person/Organisation](#). Aber auch bei der oft üblichen Übertragung von Informationen von älteren Karteikarten, bei der Auswertung von Bestandsverzeichnissen oder Ausstellungskatalogen ist es in vielen, nicht zuletzt sammlungs-internen Nachnutzungskontexten wertvoll, einen Beleg für die Quelle der Information zu haben.
- ▶ **Dieselbe Art von Information wird von verschiedenen Personen in unterschiedlichen Datenfeldern erfasst.** Bei einer Suche in einem Portal wäre dann unklar, welche Datenfelder für die Suche bevorzugt auszuwerten wären.
- ▶ **Es wird unterlassen, Bildinhalte zu erschließen,** weil man nicht explizit erfasst, was das Objekt, auch unabhängig von seinem Titel, tatsächlich bildlich darstellt. Auf diese Weise werden die Objekte bei einer textbasierten Suche mit themenbezogenen Suchbegriffen nicht gefunden. Eine Suche mit „Wohnraum“ oder „Gesellschaftsspiel“ findet z. B. nicht Vuillards „Beim Brettspiel (La partie de Dames)“ (<https://www.staedelmuseum.de/go/ds/1809>).
- ▶ **Nicht plausible Informationen:** Durch nachträgliche automatisierte Plausibilitätsprüfungen kann z. B. untersucht werden, ob die Entstehung eines Werks versehentlich außerhalb der Lebenszeit des Urhebers datiert wurde oder das Ankaufsdatum vor dem Entstehungszeitpunkt liegt.
- ▶ **Verwendung synonyme Bezeichnungen** (z. B. „Porträt“ vs. „Bildnis“ oder „im 3. Jhd. n. Chr.“ vs. „im 3. Jahrhundert AD“ vs. „im dritten Jahrhundert nach Christus“); So liefert die Suche nur einen Teil der von den Nutzer:innen gewünschten Treffer, da alle Datensätze mit anders bezeichnetem Dateninhalt nicht gefunden werden können.

Empfehlungen für die Erfassung

- ▶ **Verwendung von Oberbegriffen, obwohl eine genauere Information vorliegt:** Bei der Arbeit mit kontrolliertem Vokabular sollte immer der präziseste Begriff ausgewählt werden, der zur Verfügung steht oder das adressierte Vokabular regelkonform ergänzen kann. Besteht der Bildträger aus Eichenholz, sollte man dieses auch angeben und nicht lediglich den Oberbegriff "Holz" eintragen. Ein hierarchisch strukturierter Thesaurus im Hintergrund leistet diese Einordnung bereits, wie die Verortung des Begriffs "oak (wood)" im AAT zeigt: <http://vocab.getty.edu/aat/300012264>. Über semantische Technologien können auch die Oberbegriffe des AAT für die Indexierung herangezogen werden, sie müssen daher nicht bei der Erfassung eingetragen werden. Auch ein hausinternes Vokabular sollte um Begriffe mit ausreichendem Differenzierungsgrad ergänzt werden, um diese Exaktheit der Beschreibung zu ermöglichen.
- ▶ **Inkonsistente Verwendung kontrollierter Vokabulare:** Hausinterne Vokabulare und Thesauri bedürfen der regelmäßigen redaktionellen Pflege. Die zulässigen Bezeichnungen für bestimmte Felder sollten möglichst vollständig und fehlerfrei angeboten werden – nur dann können sie ein verlässliches Arbeitsinstrument für die Erfassenden sein. Die Redaktion gewährleistet in einem definierten Prozess, wie das Vokabular bei Bedarf um sinnvolle neue Begriffe und Bezeichnungen ergänzt werden kann. Bei Verweisen auf externe kontrollierte Vokabulare muss stets der spezifischste Begriff in der richtigen Hierarchie gefunden und mittels des ihn identifizierenden vollständigen URI zitiert werden, siehe auch: [Kontrollierte Vokabulare](#).
- ▶ **Es fehlen Informationen zum Kontext der Datenerhebung.** Im Rahmen von welchem Projekt, zur Klärung welcher Forschungsfrage, von wie vielen Personen mit welchem Kenntnisstand sind die Daten erhoben worden? Dies sollte idealerweise in Form eines durchsuchbaren Metadatendokuments erläutert werden, das die Umstände der Datenerfassung beschreibt. Das kann beispielsweise in einem Aufsatz in einer Fachzeitschrift oder in einem die digitale Publikation begleitenden Dokument geschehen. Dies erhöht die Nutzbarkeit der Daten in entscheidendem Maß, weil Nachnutzende den wissenschaftlichen Kontext der Entstehung der Daten und die Eignung für ihre eigenen Zwecke besser beurteilen können. Aus den gleichen Gründen wird diese Art der Dokumentation auch für Datenerhebungen im Kontext nicht projektgebundener Bestandserschließung empfohlen.

Siehe auch: Arno Kesper, Markus Matoni, Julia Rössel, Michelle Weidling, & Viola Wenz. (2020). Catalog of Quality Problems in Data, Data Models and Data Transformations. Zenodo. <https://doi.org/10.5281/zenodo.3955500>.