

AVefi-how-to

Metadatenlieferung durch Exporte

Erster Anhang zum AVefi-Manual

- Vorabversion 1.1 -

Autoren: Antje Dittmann / Annette Groschke / Jonas Pitz (SDK); Thomas Ochs / Stefan Matijasevic (FMD); Sabrina Silbernagel (TIB)

Was ist AVefi und wo kann ich weitere Informationen zum Projekt finden?

Im DFG-Projekt AVefi (Automatisiertes Verbundsystem für audiovisuelle Bestände über einheitliche Film-Identifikatoren) wird ein – auf persistenten Identifikatoren basierendes, institutionsübergreifendes, web-basiertes – Recherche-System entwickelt und erprobt. Es dient der Identifikation und Vernetzung von Filmwerken, ihren Manifestationen und Exemplaren aus heterogenen Beständen. Jedes Filmwerk, jede Manifestation und jedes Exemplar bekommt vom System einen automatisch erstellten Einheitlichen Film Identifikator (efi)¹ als persistenten Identifikator (PID).

Auf der Projektwebsite <https://projects.tib.eu/av-efi>² stehen detaillierte Informationen zum Projekt.

Die Alpha-Phase für die AVefi-Recherche-Plattform startete am 28.05.2025. Seitdem ist Plattform in der Alpha-Version unter <https://www.av-efi.net/> abrufbar. Da sich jedoch die Website noch in Weiterentwicklung befindet, kann an dieser Stelle nur bedingt dokumentiert werden. Darüber hinaus sind einige Funktionen noch nicht vollständig verfügbar bzw. befinden sich im Erprobungsprozess.

Für wen ist dieses Dokument?

Das Dokument dient als eine Ergänzung des AVefi-Manuals, aktuell in der Vorabversion 2.0: <https://tib.eu/avefi-manual-vv2>. Das How-To stellt eine kurze Anleitung für filmhaltende Institutionen dar, die ihre Bestände und die dazugehörigen Werke persistent registrieren und vernetzen möchten sowie zunächst **KEINE** Schnittstelle entwickeln können. Die Lieferung von Metadaten an das AVefi-System erfolgt bevorzugt per Schnittstelle, um die volle Funktionalität zu nutzen.

Was ist das AVefi-Datenschema, seine Vokabulare und Normdaten?

Das AVefi-Datenschema und die Darstellung auf dem Recherche-Frontend sind dreistufig, wobei die Anzahl der Kernmetadaten bewusst geringgehalten ist. Es gibt nur zwei Pflichtfelder: die Angabe zur datengebenden Institution und den Titel (für Werke und Manifestationen). Daher ist es auch ein- und zweistufig erfassenden Institutionen problemlos möglich, ihre Metadaten auf dem Schema abzubilden.

Das Schema in der aktuellen Version ist im GitHub unter folgendem Link zu finden: <https://github.com/AV-EFI/av-efi-schema>. Die Dokumentation des Schemas mit Hinweisen zur Nutzung sowie einige praktische Beispiele finden sich hier: <https://av-efi.github.io/av-efi-schema>. Die für das Mapping relevanten, kontrollierten Vokabulare von AVefi sind hier beschrieben: <https://av-efi.github.io/av-efi-schema/#enumerations>. Das Schema orientiert sich eng an dem FIAF Manual, der Ontologie FIAFCore, dem Metadatenstandard EN 15744 und dem Standard EBUCore.

¹ Eigenschreibweise

² Diese und alle folgenden Netzquellen wurden am 25.09.2025 auf Zugreifbarkeit geprüft.

Für Datumsangaben wird der EDTF Standard der Library of Congress verwendet:

<https://www.loc.gov/standards/datetime/>.

Für geographische Daten wird der Ghetty Thesaurus of Geographic Names (TGN) empfohlen:

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

Für Personendaten und Schlagworte wird die Verwendung von GND-Daten mit entsprechender Referenzierung empfohlen: <https://lobid.org/gnd>.

Wie ist der Ablauf der Datenlieferung durch Exporte?

An Datenlieferungen interessierte Institutionen können unter der unten genannten E-Mail-Adresse Kontakt mit dem AVefi-Team (Metadatenfachgruppe) aufnehmen, um Fragen zu klären und einen Zeitplan zu erstellen.

Metadaten können entweder direkt als JSON-Format gemäß AVefi-Schema geliefert werden oder es wird ein Mapping zu einem Zwischenformat, wie XML, RDF, CSV, durchgeführt. Auf Anfrage sind weitere Formate möglich. Die Technische Gruppe unterstützt, wenn notwendig, bei einem Mapping. Es ist grundsätzliches Ziel, den Datenexport so niedrigschwellig wie möglich, das heißt mit wenigen Kernfeldern, zu realisieren.

Ein Beispieldatensatz für AVefi konformes JSON findet sich unter:

https://github.com/AV-EFI/av-efi-schema/blob/main/examples/json/menschen_am_sonntag_all_in_one.json

Bei der Entwicklung von Tools zur Erzeugung von AVefi konformen JSON-Datensätzen empfiehlt sich ein Blick auf unser efi-conv Repositorium: <https://github.com/AV-EFI/efi-conv/>. Dieses stellt ein Check-Modul zur Verfügung, mit dessen Hilfe die erzeugten Dateien validiert werden können (siehe README). Außerdem sind Submodule für verschiedene Mappings enthalten, die als Ausgangspunkt für eigene Entwicklungen genutzt werden können, sofern das sinnvoll erscheint.

Wichtig: Um die eindeutige Beziehung zwischen den Ebenen Werk, Manifestation und Exemplar zu gewährleisten, müssen dem Export für jeden Datensatz entweder lokale oder für den Export erzeugte Identifier beiliegen. Details dazu hier: https://av-efi.github.io/av-efi-schema/has_identifier/.

Was sind die Voraussetzungen für eine Datenlieferung?

Werke lassen sich in AVefi nur in Verbindung mit Beständen registrieren. Daher gibt es keine Werke ohne Manifestationen und Exemplare. Die Lieferung von Exporten erfordert zudem die Übernahme der registrierten efi in das lokale System, um bei weiteren Exporten die Zuordnung zu erleichtern. Pflichtfeld für jeden Werksdatensatz sowie aus technischen Gründen auch für jeden Manifestationsdatensatz ist der Titel. Sollten für Manifestationen keine Titel vorliegen, wird der Werkstitel übernommen. Darüber hinaus wird für Werksdatensätze empfohlen, möglichst viele Informationen für die vier Kernfelder Titel, Produktionsort, Jahr und Regie zu übermitteln, da diese die Grundlage für die automatisierte Disambiguierung von Werken bilden.

Titel

Idealerweise wird für Werke der Originaltitel angegeben. Alternativ sind auch andere Titel, wie beispielsweise der Verleihtitel, zulässig. Die Sprache des Titels wird nicht erfasst. Folgennummern, beispielsweise bei Folgen einer Serie, werden im Titel vermerkt. Artikel werden vorangestellt. Zusätzlich kann ein Sortiertitel mit angegeben werden, bei dem der Artikel mit Komma abgetrennt an den Titel hintenangestellt wird.

Wenn kein Titel vorhanden ist, wird ein Archivtitel gebildet. Dazu wird folgende Konvention empfohlen, die sich an den FIAF-Empfehlungen orientiert:

Archivtitel setzen sich zusammen aus:

- Name der Quelle oder Sammlung: geben Sie hier den Namen der Filmsammlung an, falls vorhanden
- Kategorie: Falls gewünscht, kann an zweiter Stelle des Archivtitels eine Kategorie genannt werden, um die Auffindbarkeit z.B. aller Betriebsausflüge zu erleichtern. In Frage kommende Kategorien

wären z.B.: Betriebsausflug, Familienfilm [Home Movie], Filmtagebuch [Institutionsspezifische Bezeichnung (z.B. Haus des Dokumentarfilms) für Aufnahmen aus dem Zweiten Weltkrieg, die von Wehrmachtssoldaten hergestellt wurden.], Fragment, Kompilation, Privataufnahme [Jeder Familienfilm ist eine Privataufnahme, aber nicht jede Privataufnahme ist ein Familienfilm. Beispiel: Katzenvideo], Reisefilm [selbstgedrehte Aufnahmen einer privaten Reise]

- Wer oder was: im Film gezeigte Personen, Ereignisse (Skifahren, Bügeln, Zugfahrt, Einschulung etc.) oder Objekte (Auto, Fahrrad, elektrische Eisenbahn)
- Wo: Ort
- Wann: Zeitspanne - zusätzlich zur Verzeichnung des Produktionszeitraums in den dafür vorgesehenen Datenfeldern kann im Archivtitel z.B. ein genaues Datum (TT-MM-JJJJ) verzeichnet werden oder das Jahr der Aufnahme kann hier wiederholt genannt werden, wenn gewünscht (Weihnachten 1974).

Beispiele für Titelangaben:

- *Originaltitel:* Metropolis
- *Originaltitel:* Das Boot
- *Sortiertitel:* Boot, Das.
- *Episodentitel:* Die Wespe. 2
- *Archivtitel:* Privataufnahmen Vergnügungspark USA, 1940er Jahre
- *Archivtitel:* Sammlung Max Mustermann, Reisetagebuch, Dresden, 1970-1971

Produktionsort

Üblicherweise wird das Land angegeben, in dem die Produktionsfirma des Werks ihren Geschäftssitz hat. Bei Amateurfilmen besteht eine höhere Wahrscheinlichkeit, dass verschiedene Werke beispielsweise gleiche Titel und Jahresangaben aufweisen („Weihnachten, 1974“). Daher kann hier für den Produktionsort eine genauere Angabe, z.B. die Stadt, sinnvoll sein. Die Anreicherung von geographischen Daten mit Identifiern des Getty Thesaurus of Geographic Names (TGN) wird empfohlen.

Beispiele für Ortsangaben:

- *Produktionsort:* Bundesrepublik Deutschland
- *Produktionsort:* Frankreich (TGN ID: <http://vocab.getty.edu/page/tgn/7008038>)
- *Produktionsort:* Dresden
- *Produktionsort:* Deutschland (TGN ID: <http://vocab.getty.edu/page/tgn/7000084>)
- *Produktionsort:* Unbekannt

Jahr

Zeitangaben richten sich nach dem ISO 8601-2019 Standard, speziell nach dem dort integrierten Extended Date/Time Format (EDTF) der Library of Congress. Zugelassen sind alle Schreibweisen der Stufe o sowie die Qualifikatoren ? für ein ungewisses Datum (+ - 5 Jahre) und ~ für einen ungefähren Zeitpunkt (+ - 1 Jahr).

Beispiele für Datumsangaben:

- 2015-04-24: Vollständiges Datum, also Jahr, Monat und Tag
- 2015-04: Jahr und Monat
- 2015: Jahr
- 2015~: Ungefähr 2015, an einem Zeitpunkt zwischen 2014 und 2016
- 2015?: Ungewisses Datum 2015, an einem Zeitpunkt zwischen 2010 und 2020
- 2015/2016: Ein Zeitraum von 2015 bis 2016
- 2015-04-24/2016-06: Ein Zeitraum vom 24.04.2015 bis Juni 2016

Regie

Die Regie wird nach dem Schema Nachname, Vorname eingetragen. Die Verwendung von Identifiern der GND wird empfohlen, da diese bei der Werksdisambiguierung von Vorteil sind, z.B. im Fall von unterschiedlichen Schreibweisen eines Namens im Freitextfeld.

Beispiele für Regieangaben:

- *Regie:* Troller, Georg Stefan
- *Regie:* Petzold, Christian (GND ID: <http://d-nb.info/gnd/134218272>)
- *Regie:* Wachowski, Lana (GND ID: <http://d-nb.info/gnd/122552571>)
- *Regie:* Wachowski, Lilly (GND ID: <http://d-nb.info/gnd/122552601>)
- *Regie:* Unbekannt

Wie bekomme ich die *efi* in meine institutseigene Datenbank?

Für den Fall von Update-Exporten werden die zugeordneten *efi* in der institutionseigenen Datenbank verzeichnet. Dafür werden die ins AVefi-System importierten Datensätze mit den dabei registrierten oder zugeordneten EFIs angereichert und anschließend zurückgegeben.

Gibt es die Möglichkeit einer Redaktion?

Grundsätzlich gibt es auf der AVefi-Rechercheoberfläche mit Anmeldedaten auch eine Redaktionsmöglichkeit, um spezifisch Metadaten ändern zu können. Diese ist jedoch bewusst sehr rudimentär gehalten. Änderungen sollen möglichst in den Lokalsystemen vorgenommen werden und bei einem erneuten Export aktualisiert werden.

Sind die Daten frei verfügbar?

Es können nur Daten eingebracht werden, deren Erfassung, Weitergabe und gemeinfreie Veröffentlichung rechtlich zulässig sind. Die eingebrachten Metadaten stehen allen Interessierten in verschiedenen Formaten und über verschiedene Bezugswege zur Verfügung. Sie sind sowohl als einzelne Datensätze als auch als gesamtes Datenset kostenfrei unter „Creative Commons Zero“-Bedingungen (CCO 1.0) zur freien Nutzung abrufbar.

Wie kann ich Kontakt aufnehmen?

E-Mail: contact@av-efi.net

Mailing-Liste: Registrierung unter <https://listserv.gwdg.de/mailman/listinfo/av-efi-community>